

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## Cocoa smoky off-flavor: Chemical characterization and objective evaluation for quality control

### **This is the author's manuscript**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/1714885> since 2024-04-04T08:53:28Z

*Published version:*

DOI:10.1016/j.foodchem.2019.125561

*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)

**This is the author's final version of the contribution published as:**

[Pamela Perotti, Chiara Cordero, Cristian Bortolini, Patrizia Rubiolo, Carlo Bicchi, Erica Liberto, Food Chemistry, 2019, <https://doi.org/10.1016/j.foodchem.2019.125561>]

**The publisher's version is available at:**

[<https://www.sciencedirect.com/science/article/pii/S0308814619316851?via%3Dihub>]

**When citing, please refer to the published version.**

**Link to this full text:**

[ <http://hdl.handle.net/> ]

1 **Cocoa smoky off-flavor: chemical characterization and objective evaluation for**  
2 **quality control**

3 Pamela Perotti<sup>1</sup>, Chiara Cordero<sup>1</sup>, Cristian Bortolini<sup>2</sup>,-Patrizia Rubiolo<sup>1</sup>, Carlo Bicchi<sup>1</sup>, Erica Liberto\*

4 <sup>1</sup> Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino, Turin, Italy

5 <sup>2</sup> Soremartec Italia S.r.l., P.le P. Ferrero 1, 12051 Alba (CN), Italy

6

7 **ABSTRACT**

8 Cocoa smoky off-flavor is due to inappropriate post-harvest processing and cannot be removed in  
9 the subsequent chocolate-manufacturing steps. To date, no reliable analytical method to detect  
10 key-analytes responsible for smoky off-flavor in incoming raw material is available. This study aims  
11 to develop an analytical method, suitable for quality control, to detect smoky markers. The cocoa  
12 volatilome was first profiled by headspace solid phase microextraction combined with  
13 comprehensive two-dimensional gas chromatography-mass spectrometry from a set of  
14 representative smoky and non-smoky samples; advanced fingerprinting revealed the chemicals  
15 responsible for the off-flavor. The results served to develop a 1D-GC method suitable for routine  
16 application. Ten identified smoky markers were subjected to accurate quantification, thereby  
17 defining operative ranges to accept/reject incoming bean samples. On average, these markers are  
18 present in smoky samples at 7 to 125 fold concentrations vs. those in non-smoky beans, ranging  
19 from 8 ng/g for *p*-ethylguaiacol to 482 ng/g for phenol.

20

21 *Keywords: Cocoa volatilome, smoky off-flavor, HS-SPME, GCxGC-TOF MS, GC-MS, chemometrics*

22

23

24 **Chemical compounds:** 1H-pyrrole-2-carboxaldehyde (PubChem CID: 934729); 2,6-dimethoxyphenol  
25 (PubChem CID: 7041), phenol (PubChem CID: 996), *p*-cresol (PubChem CID: 2879), 2-ethoxy-4-  
26 methylphenol (PubChem CID: 75715), 3-ethylphenol (PubChem CID: 12101), *p*-ethylguaiacol  
27 (PubChem CID: 162465), guaiacol (PubChem CID: 460), naphthalene (PubChem CID: 931).

28

29        **1. INTRODUCTION**

30    Food taints and off-flavors are particularly important in food manufacturing, because they may  
31    impact consumer confidence and quality perception, while influencing the brand image (Ridgway,  
32    Lalljie, & Smith, 2010). A food taint derives from external sources of contamination, e.g. from the  
33    environment, processing or storage, whereas an off-flavor may be due to compounds formed  
34    through chemical or enzymatic reactions undergone by food components: lipid oxidation, hydrolytic  
35    processes or microbiological spoilage (Jelen, 2006; Mottram, 1998; Ridgway et al., 2010).

36    ‘Flavor’ is a multisensory phenomenon involving olfaction, taste, texture and chemestesis, and  
37    provides a distinctive hedonic definition of each food (Auvray & Spence, 2008). In this context,  
38    compounds causing off-flavors are ligands that, even if present at trace levels, may trigger olfactory  
39    and taste perception resulting in unpleasant and/or unexpected flavor notes.

40    Chocolate is a typical comfort foods having a rapid and positive impact on a person’s mood (Macht  
41    & Mueller, 2007); its main ingredient is cocoa, produced from cocoa beans (*Theobroma cacao* L.  
42    Malvaceae), a tree crop native of the South American continent. Top world producers of cocoa in  
43    2017 were African countries (Ivory Coast, Ghana, Cameroon, Nigeria) followed by Indonesia and  
44    South America, with Brazil and Ecuador dominating the market (Eghbal, 2018).

45    The principal climatic factor influencing cocoa yield and quality is rainfall, although temperature and  
46    light exposure are important, affecting pod and bean characteristics (Budiansky, 2018; The  
47    International Cocoa Organization (ICCO), 2018). Fermentation and drying are the two fundamental  
48    post-harvest treatments that impact the final flavor quality of cocoa products. Traditional drying,  
49    where the beans are exposed to the sun for 6-10 days, is to be preferred, not least because it is the  
50    simplest and most common method used, resulting in good-quality beans.

51    To satisfy the ever increasing demand for cocoa, drying is sometimes speeded up by artificial  
52    processes. When carried out using heat generated from burning wood or other fuels, artificial drying

53 requires the smoke originated not to come into contact with the beans, to avoid any transfer of  
54 volatiles and semi-volatiles. However, in small farming communities, correct practices are  
55 sometimes neglected and the sensory quality of beans may be altered. When improperly conducted,  
56 artificial drying can develop a typical smoky off-flavor in cocoa beans; the characteristic note  
57 depends on the drying plant, the fuel (wood, diesel, etc.), the type of wood and, after drying, also  
58 the storage conditions of the beans (CABISCO/ECA/FCC, 2015; Serra Bonvehí & Ventura Coll, 1998).  
59 To date, the occurrence of smoky off-flavor has been found to be limited to African countries, where  
60 cocoa is mainly produced by small family farms, and increasing market demand, together with  
61 climate change, has increased pressure on the producers (Wessel & Quist-Wessel, 2015).

62 The smoky off-flavor is sometimes also described as “hammy” because it is reminiscent of smoke-  
63 cured bacon. Hammy off-flavors can also arise from over-fermentation, although in smoke-  
64 contaminated beans the hammy note is dominant, while in over-fermented cocoa it takes second  
65 place to the predominant putrid, ammoniacal or occasionally soapy/phenolic background  
66 (Aprotosoiaie, Vlad Luca, Miron, 2016; CABISCO, 2015; Serra Bonvehí, 1998). The smoky note has  
67 chiefly been related to the presence of phenolic compounds, which predominantly derive from  
68 lignin degradation by pyrolysis (Janairo & Amalin, 2018; Serra Bonvehí, 1998; Wang, Chambers, &  
69 Kan, 2018). Temperature, one of the principal variables impacting the formation of smoke-  
70 reminiscent odorants, conditions the chemical structure and substitution of the resulting phenol  
71 derivatives. 4-Substituted guaiacols and 4-substituted syringols prevail at lower temperatures, while  
72 at higher temperatures the reaction environment becomes richer in H-donors (H-radicals),  
73 triggering the formation of catechols/pyrogallols and *o*-cresols/xilenols (Janairo, 2018; Kawamoto,  
74 2016). Very interestingly, guaiacols and methylphenols are also cocoa key-aroma compounds, their  
75 presence in high concentrations can affect cocoa’s sensory properties, influencing the native smoky  
76 note (Frauendorfer & Schieberle, 2006). This smoky perception should therefore be considered as

77 a taint, because it mainly derives from exposure to process smoke, or as an off-flavor, when it is due  
78 to the neo-formation of potent odorants in beans exposed to high temperatures. In native cocoa,  
79 the smoky note may differ in intensity; it may persist and/or may be emphasized in finished products  
80 (chocolate or confectionary), partly because of improper manufacturing practices. Moreover, it has  
81 been shown that odorless compounds, when combined with potent odorants, can also contribute  
82 to the sensory profile of a perceived flavor, and increase the perceived intensity of the smoky note  
83 (Chambers & Koppel, 2013; Jaffe, Wang, & Chambers, 2017). This synergistic effect is likewise  
84 possible with the association between two non-smoky phenolic compounds, such as for instance  
85 2,6-dimethylphenol at 100 ppm and eugenol at 1 ppm in propylene glycol and delivered from a  
86 fragrance strip (Wang et al., 2018). Moreover, several flavor compounds responsible for positive  
87 sensory attributes in foods can act as off-flavors when their concentration exceeds a certain  
88 threshold. Known examples are sulfur compounds, such as dimethyl sulphide in beer, which has a  
89 cabbage-like aroma at high concentrations, or 4-vinylphenol and 4-vinylguaiacol in wines (Jelen,  
90 2006; H. Wang et al., 2018). Very few studies are available concerning cocoa smoky off-flavors  
91 (Aprotosoiaie et al., 2016; Lehrian, Keeney, & Lopez, 1978; Serra Bonvehí, 1998).

92 In this context comprehensive two-dimensional gas chromatography (GC×GC) coupled with time-  
93 of-flight mass spectrometry (TOF MS) is an effective approach for detailed characterization of  
94 complex mixtures of volatiles in food (Cordero, Kiefl, Schieberle, Reichenbach, & Bicchi, 2015).  
95 GC×GC exploits the separation power and detection potential of the two dimensions, providing  
96 representative 2D chromatographic patterns, and increasing sensitivity versus trace components.  
97 Despite its potential, GC×GC-TOF MS is not yet routinely used in chemical characterization of foods;  
98 in general, it is considered too complex for quality control laboratories, although new less  
99 sophisticated commercial solutions avoiding thermal modulation and cryogenics make it promising  
100 for routine analysis (Magagna et al., 2017). Conversely, methods based on 1D-GC-MS are well

101 accepted, cost-effective and, when integrated with automatic sample preparation and injection  
102 systems, enables fully-automatic analytical procedures for high-throughput screening to be  
103 developed.

104 Headspace solid phase microextraction sampling (HS-SPME) perfectly meets the above  
105 requirements, and has been widely used to characterize cocoa aroma (Ducki, Miralles-Garcia,  
106 Zumbé, Tornero, & Storey, 2008; Magagna et al., 2017; Phuong et al., 2015). It has also been  
107 adopted to screen off-flavors in several other foods, in particular to identify components responsible  
108 for unpleasant odor(s), such as haloanisoles in wine and cork, or geosmine and methylisoborneol in  
109 water (Jelen, 2006; Ridgway et al., 2010). Food consumption is mainly driven by the pleasure  
110 perceived during its intake; therefore, food sensory features became an integral part of the quality  
111 control (QC) and quality assurance. To date, the approach to detect the smoky aroma is based on a  
112 sensory evaluation by trained panelists resulting in a rather expensive and time consuming process.  
113 In addition, the lack of a reference and objective analytical method to detect and quantify chemical  
114 markers of the smoky off-flavor on incoming raw materials inspired the current research.

115 This study aimed at developing an analytical method, suitable for routine quality control, for  
116 volatiles profiling and accurate quantitation of smoky off-flavor key-markers in cocoa beans and  
117 liquors. The final method should afford fast, accurate, objective discrimination between smoky and  
118 non-smoky cocoa products. This goal was pursued with a top-down strategy where HS-SPME-  
119 GC×GC-TOF-MS served as screening platform to identify informative odorants within a subset of  
120 samples characterized as smoky and non-smoky by an internal panel. HS-SPME-GC-MS in a fully  
121 automated set-up was then used to monitor targeted discriminating compounds and to accurately  
122 quantify them by multiple headspace extraction (MHE). Quantitative ranges for targeted  
123 compounds were then fixed as decisive markers to accept or reject incoming cocoa samples.

124

## 125 2. MATERIALS AND METHODS

### 126 2.1 Cocoa samples and Reference compounds

127 The sample set included beans (n= 54) and liquors (n= 31) of cocoa samples (*Theobroma cacao* L.  
128 main crop) of commercial grade (beans size “standard” based on counting test- federation of cocoa  
129 commerce) from different origins and harvested in different years (**Table 1**). Cocoa bean and liquor  
130 samples were provided by Soremartec Italia srl (Alba, Italy).

131 Pure reference standards for identity confirmation (key-aroma compounds and informative  
132 volatiles) were from Millipore (Milan, Italy) (**Table 2**), in particular acetic acid, 3-methyl butanoic  
133 acid, 3-methyl butanal, 2-phenyl ethanol, 2-heptanol, butanoic acid, 2-methyl butanal, linalool,  
134 phenylacetaldehyde, 2-ethyl-3,5-dimethyl pyrazine, 4-hydroxy-2,5-dimethyl-3(2H)-furanone, 2-  
135 ethyl-3,6-dimethyl pyrazine, (E,E)-2,4-nonadienal, dimethyl trisulfide, 2-methyl propanoic acid,  
136 ethyl-2-methyl butanoate, ethylbenzoate, 1,2,4-trimethoxybenzene, 2,6-dimethoxyphenol,  
137 carvacrol, 2-phenoxyethanol, *p*-cresol, *p*-ethylguaiacol, 1-h-pyrrole-2-carboxaldehyde, phenol, 2-  
138 ethoxy-4-methylphenol, guaiacol, isoamylbenzoate naphthalene, 1,2-dimethoxybenzene, and 1,4-  
139 dimethoxybenzene.

140 Normal alkanes (*n*-alkanes *n*-C9 to *n*-C25) for Linear Retention Index ( $I^T_s$ ) determination and Internal  
141 standardization (*n*-heptadecane *n*-C17 - ISTD) were from Millipore (Milan, Italy) (**Table 2**).

142 A standard stock solution of ISTD at 1000 mg/L was prepared in degassed sunflower seed oil and  
143 stored in a sealed vial at -18°C.

144

### 145 2.2 Automated Head Space Solid Phase Micro Extraction: sampling devices and conditions

146 Automated Headspace Solid Phase Microextraction (auto-HS-SPME) was performed using a Combi-  
147 PAL AOC 5000 (Shimadzu, Milan, Italy) on-line integrated with a Shimadzu QP2010 GC–MS system

148 provided with Shimadzu GC–MS Solution 2.51 software (Shimadzu, Milan, Italy). SPME fiber:  
149 Divinylbenzene/Carboxen/Polydimethylsiloxane (DVB/CAR/PDMS) df 50/30  $\mu\text{m}$  - 2 cm length from  
150 Millipore (Bellefonte, PA, USA). Fibers were conditioned before use as recommended by the  
151 manufacturer. The standard-in-fiber procedure was adopted to pre-load the ISTD (*n*-C17) onto the  
152 fiber before sampling (Y. Wang, O’Reilly, Chen, & Pawliszyn, 2005). 5.0  $\mu\text{L}$  of ISTD solution were  
153 placed in a 20 mL glass vial and submitted to HS-SPME at 80°C for 20 min, stirring speed 350 rpm.  
154 Cocoa samples were ground in liquid nitrogen and then stored at -80°C until analyzed. Samples were  
155 ground before headspace analysis to obtain a homogeneous powder. Cocoa powder (1.00 g) was  
156 weighed in the headspace glass vials (20 mL) and submitted to automated HS-SPME sampling. After  
157 ISTD loading, the SPME device was exposed to the headspace of cocoa for 40 min at 80° at a shaking  
158 speed of 350 rpm. Extracted analytes were recovered by thermal desorption from the fiber into the  
159 split/splitless (S/SL) injection port of the GC system at 250°C for 5 min. Each sample was analyzed in  
160 duplicate. Sampling and ISTD standardization for the preliminary screening by GC $\times$ GC-TOF MS  
161 analysis was done at 50°C under the analytical conditions reported by Magagna et al. (2017).

162

### 163 **2.3 Quantitation**

164 An amount of ground material appropriate to achieve headspace linearity for target analytes was  
165 processed. MHE quantification was by the External Standard approach; an aliquot of 0.100 g of  
166 ground beans was sealed in a 20 mL headspace vial and submitted to multiple consecutive  
167 extractions, exposing the fiber to the headspace for 40 minutes at 80°C before analysis. A series of  
168 calibrating solutions of reference compounds in cyclohexane, ranging from 0.1 to 50 mg/L, were  
169 used in full evaporation for MHE external calibration. Suitable volumes of standard solutions at  
170 different concentrations were submitted to multiple consecutive extractions (as for the cocoa

171 samples). All calibration solutions and samples were analyzed in duplicate, by full evaporation MHS-  
172 SPME.

173

#### 174 **2.4 GC-MS and GC×GC-TOF MS instrument set-up and analytical conditions**

175 *GC-MS analysis- Chromatographic conditions:* analyses were run on a Shimadzu QP2010 GC-MS  
176 system, controlled by Shimadzu GC-MS Solution 2.5SU1 software (Shimadzu, Milan, Italy) Injector  
177 temperature: 240°C, injection mode: splitless; carrier gas: helium, flow rate: 1 mL/min; fiber  
178 desorption time and reconditioning: 5 min; column: SolGelwax (100% polyethylene glycol) 30 m x  
179 0.25 mm  $d_c$  x 0.25  $\mu\text{m}$   $d_f$  Trajan Analytical Science (Ringwood, Australia). Temperature program,  
180 from 40°C (2 min) to 200°C at 3,5°C/min, then to 240°C (5 min) at 10°C/min. MSD conditions:  
181 ionization mode: EI (70 eV); temperatures: ion source: 200°C; quadrupole: 150°C; transfer line:  
182 260°C; scan range: 35-350 amu.

183 *GC×GC-TOF MS analysis - Chromatographic conditions:* GC×GC analyses were run on an Agilent 7890  
184 GC unit coupled with a Markes BenchTOF-Select and Select-eV<sup>®</sup> option (Markes International Ltd,  
185 Llantrisant UK) operating in the EI mode at 70eV. The transfer line was set at 250°C. TOF acquisition  
186 was set at m/z 35-350 with 100 Hz sampling frequency. The GC system was equipped with a two-  
187 stage KT 2004 loop thermal modulator (Zoex Corporation, Houston, TX) cooled with liquid nitrogen  
188 and controlled by Optimode™ V.2 (SRA Instruments, Cernusco sul Naviglio, MI, Italy). Hot jet pulse  
189 time was set at 250 ms, modulation time was 3.5 s and cold-jet total flow progressively reduced  
190 with a linear function from 40% of Mass Flow Controller (MFC) at initial conditions to 8% at the end  
191 of the run.

192 SPME thermal desorption into the GC injector port was in split mode, split ratio 1:20. Carrier gas  
193 was helium at a constant flow of 1.3 mL/min. The oven temperature program was: from 40°C (2  
194 min) to 200°C at 3.5°C/min and to 240°C at 10°C/min (10 min). The column set was configured as

195 follows: <sup>1</sup>D SolGel-Wax column (100% polyethylene glycol) (30 m × 0.25 mm dc, 0.25 μm df) from  
196 SGE Analytical Science (Ringwood, Australia) coupled with a <sup>2</sup>D OV1701 column (86%  
197 polydimethylsiloxane, 7% phenyl, 7% cyanopropyl) (2m × 0.1 mm dc, 0.10 μm df), from J&W  
198 (Agilent, Little Falls, DE, USA).

199 The *n*-alkanes solution for  $I^T_s$  determination was analyzed with a split/splitless injector in split mode,  
200 split ratio 1:50, injector temperature 250°C, and injection volume 2 μL.

201

## 202 **2.5 Data acquisition and processing**

203 GC×GC-TOF MS data were acquired by TOF-DS software (Markes International, Llantrisant, UK) and  
204 processed using GC Image GC×GC Software, version 2.8 (GC Image, LLC, Lincoln NE, USA). GC-MS  
205 data were collected by GCMS Solution 2.5SU1 software (Shimadzu, Milan, Italy).

206

## 207 **2.6 Analytes identification and chemometrics**

208 Targeted analysis was focused on about 70 compounds identified by matching their EI-MS  
209 fragmentation patterns (NIST MS Search algorithm, version 2.0, National Institute of Standards and  
210 Technology, Gaithersburg, MD, USA, with Direct Matching threshold 900 and Reverse Matching  
211 threshold 950) with those stored in commercial (NIST2014 and Wiley 7n) and in-house databases.

212 Linear retention indices ( $I^T_s$ ) were taken as a further parameter to support identification, and  
213 experimental values were compared to tabulated units. Principal Component Analysis (PCA), Partial  
214 Least Square Discriminant Analysis (PLS-DA) and regression analysis were performed with  
215 Pirouette® (Comprehensive Chemometrics Modeling Software, version 4.5-2014) (Infometrix, Inc.  
216 Bothell, WA).

217 Heat-map was implemented in Morpheus (<https://software.broadinstitute.org/morpheus/>) while  
218 the Kruskal-Wallis test was performed with XLstat (version 16.05) (Addinsoft, New York, NY USA).

219

### 220 **3 RESULTS AND DISCUSSION**

221 The chemical complexity of the volatilome of cocoa beans and cocoa liquor depends on the many  
222 chemical reactions occurring during the early stages of processing, most of which are catalyzed by  
223 specific enzymes (endogenous or exogenous from moulds, yeasts and bacteria) (Ohene Afoakwa,  
224 Paterson, Fowler, & Ryan, 2008). A top-down approach was here adopted to define the  
225 characteristic markers of smoky off-flavor (Konieczka & Namieśnik, 2018). This approach exploits  
226 the possibility of capturing the necessary information from highly-informative fingerprints of non-  
227 smoky and smoky samples, with sophisticated and powerful techniques. This information is then  
228 exploited to develop a 1D-GC-MS method with suitable informative potential, that offers reliability  
229 appropriate for the needs of a routine laboratory. The differential compositional characteristics  
230 highlighted by GC×GC-MS, between smoky and non-smoky samples, are used in 1D-GC-MS in  
231 combination with chemometrics, to discriminate samples and to obtain robust evidence of the  
232 markers related to the defect. Once their informative role is confirmed, some of the representative  
233 analytes are submitted to accurate quantitation, to set the limits of acceptability for incoming  
234 samples. Quantitation of smoky compounds was then done on cocoa beans.

235 The following sections illustrate: (a) the chemical complexity of the volatile fraction of cocoa  
236 samples and the information deriving from comparative analysis between smoky and non-smoky  
237 samples, as revealed by the untargeted-targeted investigation; (b) the optimization of the sampling  
238 procedure to improve the analytical response from analytes related to defective samples in the 1D-  
239 GC-MS method, (c) the role of unsupervised and supervised approaches in supporting the selection  
240 of informative chemicals, whose quantitation provides a reliable range of sample acceptability  
241 (Sgorbini et al., 2019).

242

### 243 3.1 Reveiling smoky odorant patterns by GC×GC-TOF MS

244 The sensory description of the smoky off-flavor has driven the search for the compound(s) that may  
245 be related to the smoky note. The smoked flavor is variously described as smoky, ashy, woody,  
246 musty/dusty, musty/earthy, burnt, acrid, pungent, petroleum-like, creosote/tar, cedar, bitter,  
247 metallic and sour (Jaffe et al., 2017). Several phenolic compounds, such as 2,6-dimethoxyphenol, 4-  
248 ethylguaiacol, thymol, guaiacol, and carvacrol, have been indicated as chemicals potentially  
249 associated with smoky aroma in foods. The smoky note is one of the positive sensory attributes for  
250 different foods, such as coffee, cocoa, ham and fish, and it is also used as artificial smoky flavorings  
251 (Frauendorfer & Schieberle, 2006; Janairo & Amalin, 2018; Marušić Radovčić, Vidaček, Janči, &  
252 Medić, 2016; Wang et al., 2018). Conversely, smoky notes may also be considered as negative, e.g.  
253 in wine, where they are associated with volatile and glycoconjugated phenols, and their removal  
254 considerably reduces smoke taint (Krstic, Johnson, & Herderich, 2015).

255 Smoky off-flavor in cocoa and chocolate liquor was studied by Lehrian et al. in 1978; they proposed  
256 a colorimetric method to measure phenols associated with the off-flavor, without offering any  
257 specific chemical speciation (Lehrian et al., 1978). Serra Bonvehí et al. identified 3-methylphenol (*m*-  
258 cresol), 2,3-dimethylphenol (2,3-xyleneol), 3-ethylphenol, and 4-ethylphenol as discriminant markers  
259 of the smoky note, after hydro-distillation followed by solvent extraction and GC-MS of cocoa  
260 powders (Serra Bonvehí et al., 1998). Misnawi et al. in 2011 suggested that the smoky odor of cocoa  
261 liquor analyzed by HS-SPME-GC-O was associated with the presence of  $\alpha$ -ethylidene-  
262 benzeneacetaldehyde, trimethyl pyrazine, and 2,3-dimethyl-*trans*-oxirane (Misnawi & Ariza, 2011).

263 The smoky note has been correlated with several volatiles of different natures and chemical  
264 structures, unlike the case of other sensory defects, such as the musty-earthy note imparted by  
265 haloanisoles, methylisoborneol, or geosmin. This lack of specific information is a challenge that has  
266 here been taken up through a multi-approach strategy.

267 An initial detailed “screening” was applied to reveal compositional differences on the volatile  
268 fractions of smoky and non-smoky beans. Analyses were carried out by HS-SPME-GC×GC-TOF MS  
269 on a sub-set of samples, and resulted in 2D-peak patterns described by an average of 230 peak-  
270 regions, corresponding to both untargeted and targeted analytes. This approach, known as  
271 combined untargeted and targeted (*UT*) fingerprinting, is based on the template matching strategy  
272 and enables direct comparison of peak patterns through their specific metadata (i.e., retention  
273 times and MS fragmentation patterns) (Magagna et al., 2017). Reliable correspondences are  
274 established between the same chemical entities across multiple chromatograms, thanks to analytes’  
275 relative positions (i.e., retention time correspondences) and MS spectral similarity (Direct Match  
276 Factor above 800). The output of *UT* fingerprinting is a data matrix of aligned 2D peaks and peak-  
277 regions, that can be used for comparative purposes (Cordero et al., 2010). **Figures 1 A-D** show the  
278 2D-patterns of volatiles from a non-smoky (**1A**) and a smoky (**1B**) sample, produced by/as they  
279 emerged from a polar × semi-polar column combination. Enlarged areas in **1C** and **1D** show in detail  
280 the region where some aromatic compounds elute. In particular, **Fig. 1C** provides a comparative  
281 visualization of the difference between a non-smoky sample (reference image) and a smoky sample  
282 (analyzed image) rendered as colorized fuzzy ratio. Green areas correspond to analytes (known or  
283 unknown) with a higher relative detector response in the smoky (analyzed) sample. **Fig.1D** shows,  
284 for the smoky sample, the elution region of interest for aromatics and phenol derivatives.

285 The average percent difference between smoky and non-smoky samples was calculated, to locate  
286 the most informative 2D peak-regions (known or unknown) describing the smoky pattern. The  
287 analyte response from non-smoky samples was taken as reference, and an arbitrary cut-off of 100  
288 was set, to include or otherwise the feature in the final list. Of the most informative compounds, 56  
289 were found to be more abundant in smoky samples; however, only a few of them could be  
290 correlated with the smoky note. These were naphthalene, 2-phenylethyl acetate, ethyl-4-ethoxy

291 benzoate, methoxy-4-propylphenol (dihydroeugenol), phenylethyl alcohol, 2-phenoxyethanol, *m*-  
292 cresol, phenol, *p*-cresol and 3-ethylphenol.

293 These results are in agreement with other reports (Serra Bonvehí & Ventura Coll, 1998; H. Wang et  
294 al., 2018); they were therefore taken into account in the next step, in which a 1D-GC approach was  
295 applied to screen these targeted odorants.

296

### 297 **3.2 Transfer to 1D-GC-MS analysis: improving method sensitivity toward phenolic compounds**

298 When the HS-SPME-GC×GC-TOF MS method was transferred to HS-SPME-GC-MS, smoky markers  
299 gave poor signals; in particular, the areas of the analytes with a relatively high boiling point were  
300 very small, in some cases below the method's limit of detection. Moreover, phenolic derivatives  
301 were connoted by high hydrophobicity, thus showing rather high affinity for the highly abundant  
302 cocoa fatty matrix (45-53%), while not being readily releases from ground beans. (Kopjar, Andriot,  
303 Saint-Eve, Souchon, & Guichard, 2010).

304 Careful tuning of the sample preparation step was therefore necessary to improve the recovery of  
305 these compounds, so as to obtain information about the volatiles whose average percentages varied  
306 significantly between the two sets of samples. A compromise was also sought between the need to  
307 increase the extraction rate and the need to adopt a sampling method that is easy to automate and  
308 to combine on-line with the analytical instrumentation for routine controls.

309 The following sampling conditions/variables were investigated in this perspective:

- 310 a) Sample amount (1 – 3g);
- 311 b) Sampling temperature (50 and 80°C);
- 312 c) SPME fiber coatings and composition (PDMS/DVB, DVB/CAR/PDMS, PDMS; and PDMS/DVB  
313 coated with PDMS for in-solution sampling)

314 d) HS enrichment by modifying analytes' solubility, by suspending and salting out the matrix  
315 in water or by direct in-solution sampling (Kolb & Ettre, 2006).

316 Details of the results achieved with the different sampling conditions are included in the  
317 supplementary material (Figure 1S a-d). The results indicated that sampling the cocoa powder at  
318 80°C was mandatory to improve the detectability of phenolic and benzene derivatives, and to  
319 include some components (i.e., 1,2-dimethoxybenzene) not detected before in the GC-MS profiles.  
320 The final sampling conditions were: direct headspace sampling of 1 g of cocoa beans sampled with  
321 DVB/CAR/PDMS polymer coating for 40 min at 80°C.

322

### 323 **3.3 Chemometric-driven approach to select informative markers of the smoky note**

324 The HS-SPME-GC-MS profiles obtained under the optimized sampling conditions (see paragraph 2.2)  
325 on the bean and liquor samples under study, and detailed in **Table 1**, are shown in the heat-map of  
326 **Figure 2**. The rows indicate the investigated samples (beans and liquors) and the columns the  
327 targeted analytes by HS-SPME-GC-MS. The color scale varies from blue (low abundance) to red (high  
328 abundance). Hierarchical cluster analysis (HCA) of both rows and columns, by Spearman rank  
329 correlation through the average linkage method, shows a different distribution of the volatiles based  
330 on their normalized response across samples. HCA shows a clear separation between beans and  
331 liquors. Analytes are ordered according to their inter-class variance.

332 Cocoa beans display higher relative abundances for several targets, as a reflection of quantitative  
333 changes of acids, esters, alcohols and ketones (**Table 2**) and in particular for methyl and ethyl  
334 acetates (*green-fruity*), 2-phenylethylacetate (*flowery*), 2-methyl-1-propanol, 2-heptanol (*citrusy*),  
335 2,3-butandiol (*fruity/creamy*), 3-hydroxy-2-butanone (acetoin-*buttery*) and 2-pentanone (*fruity*)  
336 (**Figure 2**). In particular, acetic acid is the most abundant volatile and, when present in high amounts  
337 in beans, it gives an intense vinegar-like odor that can affect the cocoa aroma quality (Frauendorfer

338 et al., 2006). However, during cocoa processing, undesired volatiles with low boiling points are  
339 removed or drastically decreased in concentration (up to 70% for acetic acid) (E. Ohene Afoakwa et  
340 al., 2008). Ethyl and methyl esters, in particular acetates, derive from amino acids and are typical  
341 aroma components in unroasted beans. 2-Phenylethyl acetate has been found in unroasted and  
342 roasted cocoa, and it can also be formed through yeast metabolism. Alcohols, aldehydes and  
343 ketones result from microbial activity during fermentation but, during roasting, aldehydes (in  
344 particular) are significantly reactive, also taking part in the formation of heterocyclic compounds  
345 (pyrazines), while the alcohol concentration decreases, negatively affecting the aroma because their  
346 presence is desirable to obtain sweet and floral notes in finished cocoa products (Aprotosoai et al.,  
347 2016; Misnawi & Ariza, 2011; Ziegler G., 2009).

348 Conversely, liquors contain higher amounts of 2,3,5-trimethylpyrazines and tetramethylpyrazines  
349 (*cocoa/nutty/musty* notes), acetophenone, benzaldehyde and furfural (*almond/sweet*), 2-  
350 butanone and 2-nonanone (*sweet/fruity*), 4-hydroxybutanoate (*fruity*), guaiacol and phenol  
351 (*phenolic/smoky*) **Table 2** and **Figure 2**. Phenolic compounds are key-odorants formed during  
352 roasting in relatively small amounts (Frauendorfer & Schieberle, 2006; Rychlik, Schieberle, & Grosch,  
353 1998). They are present in both non-smoky and smoky liquors, although large amounts can be  
354 formed because of incorrect drying or storage processes. Their level can also increase during bean  
355 roasting, which is generally between 110°C and 140°C. The clusters and the red and green right-  
356 hand-side bars highlight the smoky and non-smoky bean and liquor samples. The heat-map  
357 highlights the volatiles virtually linked to these clusters, including 1,2-dimethoxybenzene, guaiacol  
358 derivatives (*smoky/phenolic/spice*), *p*-cresol (*phenolic/pungent*), naphthalene (*pungent*), phenol  
359 (*phenolic/rubbery*), 2,6-dimethoxyphenol (*sweet/smoky/medicinal*), 2-methoxy-4-methylphenol  
360 (*sweet/smoky/medicinal*), and 3-ethylphenol (*musty*).

361 Unsupervised pattern recognition through Principal Component Analysis (PCA) was applied to the  
362 targeted data matrix for beans and liquors, to explore the conformation (groups) of samples and to  
363 localize informative chemicals responsible for discrimination. PCA in **Figure 3**, referred to bean  
364 volatiles, makes a clear distinction between smoky (pink) and non-smoky (blue) samples with an  
365 explained variance of 69.14 % on the first 3 PCs, regardless of the origin of the samples (**Figure 3A**).  
366 The loading plot **Figure 3B** shows that smoky samples are described by most of the volatiles  
367 mentioned above, confirming initial observations derived from the heat-map (**Figure 2**). Within  
368 beans, older samples (\*\_old harvested in 15/16 and 16/17 in table 1) are recognizable on PC1, and  
369 are characterized by a relatively high abundance of hexanoic acid and 1-H-pyrrole-carboxyaldehyde  
370 (**Figure 3 A and 3 B**).

371 PCA obtained by extrapolating only those volatiles related to the discrimination of smoky samples  
372 still shows a coherent distribution by smoky and non-smoky, at 79.44% of explained variance (data  
373 not shown).

374 Supervised pattern recognition via PLS-DA on both beans and liquors provides a coherent  
375 classification by beans or liquors, and by smoky or non-smoky samples **Figure 3 C**. PLS-DA was done  
376 on logarithm (Log10) transformed data, pre-processed by auto scaling and cross validated (5 CV).  
377 The total classification rate was 97%, in particular the classification model built up showed a 100%  
378 ability for beans, and a slightly lower one for liquor (92%). The correlation spectrum is a useful  
379 function to exclude x variables (e.g. volatiles) that correlate weakly with the qualitative y variable  
380 (e.g. smoky/non-smoky liquors and beans) (**Figure 3D**). The correlation spectrum facilitates the  
381 identification of the closest smoky-correlated analytes, i.e. the ten components in the HS-SPME-GC-  
382 MS pattern (highlighted at top right of the graph in **Figure 3D**). These components can discriminate  
383 smoky from non-smoky samples independently of the origin or processing step considered (raw  
384 cocoa beans or liquors). Significance analysis on all analytes was carried out through the non-

385 parametric Kruskal-Wallis test. The *p*-values of bean and liquor volatiles are reported in **Table 2**.  
386 Naphthalene, guaiacol, 2-methoxy-4-methylphenol, phenol, 1H-Pyrrole-2-carboxaldehyde, *p*-  
387 ethylguaiacol, *p*-cresol, 3-ethylphenol, 2,6-dimethoxyphenol, 4-methyl-2,6-dimethoxyphenol  
388 differed significantly between smoky and non-smoky samples, in both beans and liquors, although  
389 to differing extents. These analytes were therefore submitted to accurate quantification (3.4). Other  
390 components show significant variations in beans, but their variation may also be influenced by the  
391 year of harvest, and thus by the “age” of the samples (**Table 1 and 2**). **Table 2** also shows other  
392 volatiles significantly varying in smoky and non-smoky liquors, including 2-butanone, 3-methyl-  
393 butanal, 2 and 3-methyl-ethyl butanoate, hexanal, 1,3-dimethyl-benzene, benzaldehyde, 2,3-  
394 butanediol and hexanoic acid. Further investigations will be required to define their roles.

395

### 396 **3.4 Quantitation of the selected marker compounds**

397 Cross-sample comparisons through relative quantitation, based on Peak Area % or on Internal  
398 Standard normalization, may be inaccurate or misleading if taken as analyte(s) concentration  
399 indicators, in particular when the aim of profiling is to correlate chemical composition with sensory  
400 properties, or process kinetics. However, absolute quantitation of solid matrices is complex, since  
401 they are characterized by a heterogeneous composition and structure, in which volatiles can be  
402 retained and released into the HS in different ways (Sgorbini et al., 2019). Multiple Headspace  
403 Extraction (MHE) is one of the approaches to quantifying solid samples, enabling the matrix effect  
404 to be overcome (Kolb & Ettre, 2006). More recently, its use has successfully been extended to HS-  
405 SPME, also known as MHS-SPME (Bicchi et al., 2011; Sgorbini et al., 2015 and references cited  
406 therein).

407 MHS-SPME is based on stepwise dynamic gas extraction of the investigated analyte/s from a solid  
408 or liquid sample. It comprises three main steps:

409 Step 1. Exhaustive extraction of analytes from samples to define HS linearity boundaries;

410 Step 2. Application of the MHE procedure to the samples of interest;

411 Step 3. Exhaustive extraction of analytes from calibration solutions, in a range of concentrations  
412 matching real-sample concentrations.

413 Steps 1 and 2 define the total peak area obtained from a series of consecutive and exhaustive  
414 extractions; it is directly related to the total amount of analyte in the matrix. The analyte peak area  
415 decreases exponentially with the number of extractions, provided that a suitable amount of matrix  
416 is processed. The cumulative instrumental response is obtained from the following equation:

417

$$418 \quad AT = \sum_{i=1}^{\infty} Ai = A1 \left( \frac{1}{1-e^{-q}} \right) = \frac{A1}{1-\beta} \quad (\text{Eq. 1})$$

419

420 where  $AT$  is the total estimated area,  $A1$  is the area detected with the first extraction, and  $q$  is a  
421 constant describing the exponential decay of the area through successive extractions. Quantitation  
422 is achieved by external standard calibration with a standard solution of the analyte(s) investigated,  
423 subjected to the same MHE conditions as the real sample.

424 Six of the ten markers selected from the above procedure were quantified by MHS-SPME in smoky  
425 and in non-smoky beans; 3-ethylphenol was not quantified at this stage of the study because the  
426 standard was not available commercially, while 2,6-dimethoxyphenol, 4-methyl-2,6-  
427 dimethoxyphenol and 1-H-pyrrole-2-carboxyaldehyde were excluded because they were outside  
428 the HS linearity boundaries related to the sample amount chosen for the MHE procedure.

429 The results indicated that the average amounts of the investigated markers in the smoky samples  
430 were between 7 and 125 times higher than in non-smoky beans ranging from 8 ng/g for *p*-  
431 ethylguaiacol to 482 ng/g for phenol (**Table 3**). 2-Methoxy-4-methylphenol, *p*-ethylguaiacol and *p*-  
432 cresol are detectable but below the limit of quantitation in non-smoky samples. The high standard  
433 deviation for phenol and *p*-cresol is probably due to the different seasonality of the investigated  
434 samples (crop of 2 years). An operative limit below 10 ng/g of the selected smoky compounds can  
435 thus generally be adopted in acceptance of incoming bean samples (**Table 3**).

436

#### 437 **4. Conclusions**

438 ~~The intrinsic information potential of the cocoa volatile fraction has been shown to be diagnostic to~~  
439 ~~discriminate defective from non-defective samples, for both beans and liquors.~~ The top-down  
440 strategy employed successfully defined the cocoa aroma components related to smoky off-flavor,  
441 and led to a routine method for their detection. The informative power of GC×GC-TOF MS analysis,  
442 combined with advanced fingerprinting (i.e., *UT* fingerprinting), was used for a preliminary  
443 investigation on a limited but significant selection of smoky and non-smoky samples. This step  
444 enabled the chromatographic peak-regions (features) discriminating off-flavor samples ~~from the~~  
445 ~~rest to be detected~~. The chief compounds identified related to the smoky note are phenolic in nature,  
446 and are present either as minor components or in larger amounts than in the non-smoky samples.  
447 The results of this first step were used to develop a simple and automatic routine HS-SPME-GC-MS  
448 method combined with multivariate statistics, for discrimination and classification of beans and  
449 liquors. HS-SPME sampling was chosen because of the nature of the cocoa matrix, the low target  
450 analyte concentration, and ~~method's~~ its high concentration capability and reliability in quantitation.  
451 Quantitation of the selected markers allows adopting an operative limit below 10 ng/g for the  
452 acceptance of the incoming cocoa beans.

453 However, some aspects related to quantitation (e.g. HS linearity ranges for markers) merit  
454 investigation in greater depth, because of both the relatively small amounts of components related  
455 to the smoky off-odor, and the rheological complexity of the cocoa matrix. In particular, quantitation  
456 of smoky markers in liquor is still challenging, because the modification of the lipid crystalline  
457 structure during processing can influence their release to the headspace, and affect not only the HS  
458 linearity ranges with MHE (Nicolotti et al., 2013) but also the definition of their chemical limits of  
459 acceptability in compliance with the sensory perception.

460

#### 461 **Aknowledgments**

462 The research was financially supported by Soremartec Italia s.r.l Alba (CN), Italy.

463

464 **Table captions**

465 **Table 1** Cocoa samples analyzed listed by type (beans and liquors), origin and year of harvesting.

466 **Table 2** List of volatiles targeted by HS-SPME-GC-TOF MS, together with their experimental and  
467 reference  $I^T_s$ , volatiles' normalized responses both in beans and liquors,  $p$ -value through the  
468 Kruskal-Wallis test ( $\alpha=0.05$ ), and odor descriptors as reported in the reference literature  
469 (Frauendorfer & Schieberle, 2006; Rychlik et al., 1998; <http://www.thegoodscentcompany.com/>).  
470 Abbreviations: A: target analytes identified by means of authentic standards, MS: analytes  
471 tentatively identified on MS fragmentation patterns available in commercial libraries, and RI: Linear  
472 Retention Indices ( $I^T_s$ ) available in Nist (<https://webbook.nist.gov/>).

473 **Table 3** Amounts of selected smoky markers in cocoa beans (smoky and non-smoky) with standard  
474 deviation, LOQ and LOD.

475

476 **Figure captions**

477 **Figure 1** GC×GC-TOF MS patterns of volatiles for non-smoky A) and smoky samples B); in the  
478 colorized fuzzy differences, brilliant green represents the positive differences in component  
479 abundances of the smoky vs. non-smoky samples; C) enlargement of the aromatic and phenol  
480 region, with tentative identification.

481

482 **Figure 2** Heat-map of the HS-SPME-GC-MS volatile profiling of bean and liquor samples, and  
483 hierarchical cluster analysis of rows and columns by Spearman rank correlation, with the average  
484 linkage method.

485

486 **Figure 3** PCA scores A) and loadings B) plots of bean samples on the first 3 PCs; C) PLS-DA class  
487 prediction: 1 smoky liquors (red), 2 non-smoky liquors (green), 3 smoky beans (cyan), 4 non-smoky  
488 beans (blue); D) PLS-DA correlation spectrum. Data matrix was transformed by Log10 and pre-  
489 processed through auto scaling both in PCA and in PLS-DA.

490

491

492

493 **References**

- 494 Aprotosoae, A. C., Vlad Luca, S., & Miron, A. (2016). Flavor Chemistry of Cocoa and Cocoa  
495 Products-An Overview. *Comprehensive Reviews in Food Science and Food Safety*, 15(1), 73–  
496 91. <https://doi.org/10.1111/1541-4337.12180>
- 497 Auvray, M., & Spence, C. (2008). The multisensory perception of flavor, 17, 1016–1031.  
498 <https://doi.org/10.1016/j.concog.2007.06.005>
- 499 Bicchi, C., Ruosi, M. R., Cagliero, C., Cordero, C., Liberto, E., Rubiolo, P., & Sgorbini, B. (2011).  
500 Quantitative analysis of volatiles from solid matrices of vegetable origin by high concentration  
501 capacity headspace techniques: Determination of furan in roasted coffee. *Journal of*  
502 *Chromatography A*, 1218(6), 753–762. <https://doi.org/10.1016/j.chroma.2010.12.002>
- 503 Budiansky, E. (2018). Cocoa and Climate Change: Urban Legends and Rural Realities.  
504 [https://www.worldcocoafoundation.org/blog/cocoa-and-climate-change-urban-legends-and-](https://www.worldcocoafoundation.org/blog/cocoa-and-climate-change-urban-legends-and-rural-realities/)  
505 [rural-realities/](https://www.worldcocoafoundation.org/blog/cocoa-and-climate-change-urban-legends-and-rural-realities/). Accessed 08.01.2019.
- 506 CABISCO/ECA/FCC. (2015). *Cocoa Beans : Chocolate & Cocoa Industry Quality Requirements*.  
507 Retrieved from [http://www.cocoaquality.eu/data/Cocoa Beans Industry Quality](http://www.cocoaquality.eu/data/Cocoa%20Beans%20Industry%20Quality%20Requirements%20Apr%202016_En.pdf)  
508 [Requirements Apr 2016\\_En.pdf](http://www.cocoaquality.eu/data/Cocoa Beans Industry Quality Requirements Apr 2016_En.pdf). Accessed 12.2.2018.
- 509 Chambers, E., & Koppel, K. (2013). Associations of Volatile Compounds with Sensory Aroma and  
510 Flavor: The Complex Nature of Flavor. *Molecules*, 18(5), 4887–4905.  
511 <https://doi.org/10.3390/molecules18054887>
- 512 Cordero, C., Kiefl, J., Schieberle, P., Reichenbach, S. E., & Bicchi, C. (2015). Comprehensive two-  
513 dimensional gas chromatography and food sensory properties: potential and challenges.  
514 *Analytical and Bioanalytical Chemistry*, 407(1), 169–191. <https://doi.org/10.1007/s00216->

515 014-8248-z

516 Cordero, C., Liberto, E., Bicchi, C., Rubiolo, P., Reichenbach, S. E., Tian, X., & Tao, Q. (2010).  
517 Targeted and non-targeted approaches for complex natural sample profiling by GCxGC-qMS.  
518 *Journal of Chromatographic Science*, 48(4), 251–261. Retrieved from  
519 <http://www.ncbi.nlm.nih.gov/pubmed/20412646>

520 Ducki, S., Miralles-Garcia, J., Zumbé, A., Tornero, A., & Storey, D. M. (2008). Evaluation of solid-  
521 phase micro-extraction coupled to gas chromatography-mass spectrometry for the  
522 headspace analysis of volatile compounds in cocoa products. *Talanta*, 74, 1166–1174.  
523 <https://doi.org/10.1016/j.talanta.2007.08.034>

524 ~~Ohene Afoakwa E., Paterson A., Fowler, M. & Ryan, A. (2008). Flavor formation and character in~~  
525 ~~cocoa and chocolate: A critical review. *Critical Reviews in Food Science and Nutrition*, 48(9),~~  
526 ~~840–857. <https://doi.org/10.1080/10408390701719272>~~

527 Eghbal, M. (2018). Global Chocolate Industry: From Bean to Bar.  
528 <https://blog.euromonitor.com/global-chocolate-industry/>. Accessed 08.01.2019.

529 Frauendorfer, F., & Schieberle, P. (2006). Identification of the key aroma compounds in cocoa  
530 powder based on molecular sensory correlations. *Journal of Agricultural and Food Chemistry*,  
531 54(15), 5521–5529. <https://doi.org/10.1021/jf060728k>

532 Jaffe, T. R., Wang, H., & Chambers, E. (2017). Determination of a lexicon for the sensory flavor  
533 attributes of smoked food products. *Journal of Sensory Studies*, 32(3).  
534 <https://doi.org/10.1111/joss.12262>

535 Janairo, J. I. B., & Amalin, D. M. (2018). Volatile chemical profile of cacao liquid smoke.  
536 *International Food Research Journal*, 25(1), 213–216.

- 537 Jelen, H. H. (2006). Solid-Phase Microextraction in the Analysis of Food Taints and Off-Flavors.  
538 *Journal of Chromatographic Science*, 44(7), 399–415.  
539 <https://doi.org/10.1093/chromsci/44.7.399>
- 540 Kawamoto, H. (2016). Lignin pyrolysis reactions. *J. Wood Sci.*, 63(2), 117–132.
- 541 Kolb, B., & Ettre, L. S. (2006). *Static headspace-gas chromatography : theory and practice*. Wiley-  
542 Interscience.
- 543 Konieczka, P., Namieśnik, J. (2018). *Quality assurance and quality control in the analytical chemical*  
544 *laboratory: a practical approach, 2nd ed. Analytical Chemistry Series*. NW: CRC Press.  
545 <https://doi.org/10.1007/s00216-018-1461-4>
- 546 Kopjar, M., Andriot, I., Saint-Eve, A., Souchon, I., & Guichard, E. (2010). Retention of aroma  
547 compounds: an interlaboratory study on the effect of the composition of food matrices on  
548 thermodynamic parameters in comparison with water. *Journal of the Science of Food and*  
549 *Agriculture*, 90(8), 1285–1292. <https://doi.org/10.1002/jsfa.3929>
- 550 Krstic, M. P., Johnson, D. L., & Herderich, M. J. (2015). Review of smoke taint in wine: Smoke-  
551 derived volatile phenols and their glycosidic metabolites in grapes and vines as biomarkers  
552 for smoke exposure and their role in the sensory perception of smoke taint. *Australian*  
553 *Journal of Grape and Wine Research*, 21, 537–553. <https://doi.org/10.1111/ajgw.12183>
- 554 Lehrian, D., Keeney, P., & Lopez, A. (1978). Method for measurement of phenols associated to  
555 smoky/hammy flavor defect of cocoa beans and chocolate liquors. *Journal of Food Science*,  
556 43, 743–745. <https://doi.org/10.1111/j.1365-2621.1978.tb02407.x>
- 557 Macht, M., & Mueller, J. (2007). Immediate effects of chocolate on experimentally induced mood  
558 states. *Appetite*, 49, 667–674. <https://doi.org/10.1016/j.appet.2007.05.004>

559 Magagna, F., Guglielmetti, A., Liberto, E., Reichenbach, S. E., Allegrucci, E., Gobino, G., Bicchi, C.,  
560 Cordero, C. (2017). Comprehensive Chemical Fingerprinting of High-Quality Cocoa at Early  
561 Stages of Processing: Effectiveness of Combined Untargeted and Targeted Approaches for  
562 Classification and Discrimination. *Journal of Agricultural and Food Chemistry*, 65(30), 6329–  
563 6341. <https://doi.org/10.1021/acs.jafc.7b02167>

564 Marušić Radovčić, N., Vidaček, S., Janči, T., & Medić, H. (2016). Characterization of volatile  
565 compounds, physico-chemical and sensory characteristics of smoked dry-cured ham. *Journal*  
566 *of Food Science and Technology*, 53(11), 4093–4105. <https://doi.org/10.1007/s13197-016->  
567 2418-2

568 Misnawi, & Ariza, B. T. S. (2011). Use of gas Chromatography-Olfactometry in combination with  
569 solid phase micro extraction for cocoa liquor aroma analysis. *International Food Research*  
570 *Journal*, 18(2), 829–835.

571 Mottram, D. S. (1998). Chemical tainting of foods. *International Journal of Food Science and*  
572 *Technology*, 33(1), 19–29. <https://doi.org/10.1046/j.1365-2621.1998.00155.x>

573 National Institute of Standard and Technology. (n.d.). Nist Chemistry WebBook. Retrieved August  
574 20, 2006, from <https://webbook.nist.gov/chemistry/name-ser/>

575 Nicolotti, L., Cordero, C., Cagliero, C., Liberto, E., Sgorbini, B., Rubiolo, P., & Bicchi, C. (2013).  
576 Quantitative fingerprinting by headspace—Two-dimensional comprehensive gas  
577 chromatography—mass spectrometry of solid matrices: Some challenging aspects of the  
578 exhaustive assessment of food volatiles. *Analytica Chimica Acta*, 798, 115–125.  
579 <https://doi.org/10.1016/j.aca.2013.08.052>

580 Ohene Afoakwa, E., Paterson, A., Fowler, M. & Ryan, A. (2008). Flavor formation and character in

581 cocoa and chocolate: A critical review. *Critical Reviews in Food Science and Nutrition*, 48(9),  
582 840–857. <https://doi.org/10.1080/10408390701719272>

583 Phuong, D. T., Van De Walle, D., De Clercq, N., De Winne, A., Kadow, D., Lieberei, R., Messens, K.,  
584 Tran, D.N., Dewettinck, K., Van Durme, J. (2015). Assessing cocoa aroma quality by multiple  
585 analytical approaches. *Food Res. Int.*, 77, 657–669.

586 ~~Piotr Konieczka, & Namieśnik, J. (2018). *Quality assurance and quality control in the analytical*  
587 *chemical laboratory: a practical approach, 2nd ed. Analytical Chemistry Series. NW: CRC*  
588 *Press.* <https://doi.org/10.1007/s00216-018-1461-4>~~

589 Ridgway, K., Lalljie, S. P. D., & Smith, R. M. (2010). Analysis of food taints and off-flavours: a  
590 review. *Food Additives & Contaminants: Part A*, 27(2), 146–168.  
591 <https://doi.org/10.1080/19440040903296840>

592 Rychlik, M., Schieberle, P., & Grosch, W. (1998). *Compilation of odor thresholds, odor qualities*  
593 *and retention indices of key food odorants*. Dt. Forschungsanst. für Lebensmittelchemie.

594 Serra Bonvehí, J., & Ventura Coll, F. (1998). Evaluation of Smoky Taste in Cocoa Powder. *Journal of*  
595 *Agricultural and Food Chemistry*. <https://doi.org/10.1021/jf9705429>

596 Sgorbini, B., Bicchi, C., Cagliero, C., Cordero, C., Liberto, E., & Rubiolo, P. (2015). Herbs and spices:  
597 Characterization and quantitation of biologically-active markers for routine quality control by  
598 multiple headspace solid-phase microextraction combined with separative or non-separative  
599 analysis. *Journal of Chromatography A*, 1376, 9–17.  
600 <https://doi.org/10.1016/j.chroma.2014.12.007>

601 Sgorbini, B., Cagliero, C., Liberto, E., Rubiolo, P., Bicchi, C., & Cordero, C. (2019). Strategies for  
602 Accurate Quantitation of Volatiles from Foods and Plant-Origin Materials: A Challenging Task.

603 *Journal of Agricultural and Food Chemistry*, acs.jafc.8b06601.  
604 <https://doi.org/10.1021/acs.jafc.8b06601>

605 The Good Scents Company Information System. (n.d.). The Good Scents Company (tgsc).  
606 <http://www.thegoodscentscompany.com/>. Accessed 20.08.2019

607 The International Cocoa Organization (ICCO), C. P. and C. C. C. (2018). About cocoa.  
608 <https://www.icco.org/>. Accessed 08.01.2019.

609 Wang, H., Chambers, E., & Kan, J. (2018). Sensory Characteristics of Combinations of Phenolic  
610 Compounds Potentially Associated with Smoked Aroma in Foods. *Molecules (Basel, Switzerland)*, 23(8). <https://doi.org/10.3390/molecules23081867>

612 Wang, Y., O'Reilly, J., Chen, Y., & Pawliszyn, J. (2005). Equilibrium in-fibre standardisation  
613 technique for solid-phase microextraction. *Journal of Chromatography A*, 1072(1), 13–17.  
614 <https://doi.org/10.1016/j.chroma.2004.12.084>

615 Wessel, M., & Quist-Wessel, P. M. F. (2015). Cocoa production in West Africa, a review and  
616 analysis of recent developments. *NJAS - Wageningen Journal of Life Sciences*, 74–75, 1–7.  
617 <https://doi.org/10.1016/j.njas.2015.09.001>

618 Ziegleder G. (2009). Flavour development in cocoa and chocolate. In Beckett ST (Ed.), *Industrial*  
619 *chocolate manufacture and use*. (pp. 169–91). Oxford: Blackwell Publishing.

620

**Table 1**

		<b>Smoky</b>			<b>Non-Smoky</b>			
		<b>Harvesting Year</b>	<i>15/16</i>	<i>16/17</i>	<i>17/18</i>	<i>15/16</i>	<i>16/17</i>	<i>17/18</i>
		<b>Origin</b>						
<b>Beans</b>	<i>Camerun</i>	2	7	14	4	5	15	
	<i>Ecuador</i>	-	3	-	-	3	-	
	<i>Nigeria</i>	1	-	-	-	-	-	
<b>Liquors</b>	<i>Camerun</i>	1	-	13	-	-	16	
	<i>Nigeria</i>	1	-	-	-	-	-	

Table 2

Identified compounds	Compounds Confirmation	Calc $\bar{I}_s$	Lit $\bar{I}_s$	Odour description	Beans					Liquors				
					Non-smoky	sd	Smoky	sd	p-value	Non-smoky	sd	Smoky	sd	p-value
Methyl Acetate	A; RI; MS	836	832	Green, pungent	19.07	9.10	18.97	9.69	0.404	5.94	1.89	4.42	1.58	0.378
2-Methylfuran	A; RI; MS	866	871	Ethereal, acetone, chocolate	0.07	0.05	0.11	0.09	<b>0.038</b>	0.06	0.03	0.07	0.02	0.950
Ethyl Acetate	A; RI; MS	878	895	Fruity, aromatic	136.58	82.69	93.81	58.68	0.060	3.65	4.12	4.12	1.68	0.801
2-Butanone	A; RI; MS	886	908	Ethereal, slightly fruity, balsamic	0.75	3.94	1.19	3.98	0.297	8.30	6.26	5.25	1.49	<b>0.006</b>
2-Methyl-butanal	A; RI; MS	898	942	Musty, chocolate, nutty, malty	44.41	23.69	57.31	34.10	<b>0.042</b>	20.09	8.58	13.58	7.34	0.078
3-Methyl-butanal	RI; MS	902	917	Ethereal, aldehydic, chocolate, peach, fatty	95.07	48.43	115.27	69.33	0.054	48.63	17.19	29.96	14.46	<b>0.044</b>
2,4,5-Trimethyl-1,3-Dioxolane	RI; MS	927	967	-	6.25	3.43	5.57	3.67	0.602	0.51	0.18	0.37	0.19	0.166
2-Pentanone	A; RI; MS	957	988	Fruity	282.10	152.13	296.29	167.70	0.465	26.86	13.88	25.49	8.02	0.850
Isobutyl Acetate	RI; MS	992	1029	Sweet, fruity, ethereal	31.95	16.24	31.74	18.11	0.335	7.28	3.43	3.50	2.00	0.345
$\alpha$ -Pinene	A; RI; MS	1006	1016	Harsh, terpene-like, minty	0.05	0.02	0.06	0.03	<b>0.028</b>	1.26	1.22	0.69	0.47	0.208
2-methyl-ethyl butanoate	RI; MS	1034	1062	-	1.56	0.72	1.41	0.78	0.434	1.41	0.38	0.78	0.41	<b>0.014</b>
3-methyl-ethyl butanoate	RI; MS	1050	1064	-	1.83	0.89	1.57	0.87	0.137	1.04	0.34	0.58	0.33	<b>0.012</b>
2-Pentanol acetate	RI; MS	1055	1075	-	26.84	10.04	25.57	11.42	0.835	37.16	12.64	17.64	7.51	0.051
Hexanal	A; RI; MS	1060	1095	Tallowy, leaf-like	2.06	0.89	1.13	0.57	<b>0.000</b>	2.25	0.66	1.26	0.37	<b>0.001</b>
2-Methyl-1-propanol	A; RI; MS	1081	1101	-	16.23	9.27	17.16	9.98	0.676	1.53	0.73	1.04	0.55	0.313
3-Methyl-2-butanol	RI; MS	1107	1125	Fruity	87.04	42.67	80.31	42.81	0.794	21.39	26.16	26.72	11.24	0.703
1-Butanol-3-methyl-acetate	A; RI; MS	1107	1125	-	136.39	57.24	144.49	69.14	0.639	89.84	36.40	46.48	20.05	0.115

1,3-Dimethyl-benzene	RI; MS	1118	1140	-	13.10	20.68	2.81	7.61	<b>0.008</b>	1.19	1.06	3.76	5.62	<b>0.003</b>
3-Carene	A; RI; MS	1130	1143	Sweet citrus terpenic fir needle	7.18	6.84	8.15	4.89	0.657	1.84	1.77	0.18	2.20	0.059
$\beta$ -Myrcene	A; RI; MS	1137	1150	Sharp, terpenic, citrus, floreal	7.02	6.77	7.98	4.96	0.657	3.93	1.71	4.42	2.04	0.614
2-Heptanone	A; RI; MS	1157	1174	Sweet, fruity	18.97	12.64	16.06	9.91	0.191	16.61	5.37	11.64	7.65	0.101
Limonene	A; RI; MS	1166	-	Citrus, mint	0.99	0.60	0.44	0.98	0.853	2.49	0.55	2.10	0.37	0.753
3-Ethyl-1-butanol	RI; MS	1187	1318	Pungent, fusel, wine, cocoa	72.93	41.23	6.28	52.70	0.123	10.59	5.15	7.26	3.28	0.488
<i>cis</i> $\beta$ -Ocimene	RI; MS	1218	-	-	1.51	1.37	2.08	1.45	0.620	0.31	0.51	0.29	2.05	0.488
3-Hydroxy-2-butanone	A; RI; MS	1256	1259	Butter	1400.26	761.60	93.77	788.66	0.531	79.03	31.44	73.54	19.92	0.705
2-Heptanol	A; RI; MS	1294	1294	Citrus	67.25	47.76	63.70	33.11	0.896	30.73	11.60	27.95	12.99	0.488
2-Nonanone	RI; MS	1363	1385	Fruity, waxy, soapy, cheese, coconut like	11.94	9.89	11.61	9.46	0.167	31.91	17.44	41.93	18.39	0.166
2,3,5-Trimethylpyrazine	RI; MS	1380	-	Nutty, musty, powdery cocoa, potato and musty	5.14	3.00	5.36	2.40	0.549	74.55	21.26	93.05	25.98	0.313
Acetic Acid	A; RI;MS	1410	1408	sharp, pungent, sour	4213.46	1873.64	3478.30	1698.88	0.273	1894.40	842.18	1321.88	632.09	0.950
Furfural	A; RI; MS	1432	1448	Sweet, almond, woody	3.10	2.18	6.72	2.55	<b>0.000</b>	38.78	18.34	25.17	7.21	<b>0.003</b>
Tetramethyl-Pyrazine	RI; MS	1442	1466	Nutty, musty, cocoa	103.00	45.83	79.90	42.54	0.167	279.30	119.07	285.96	87.19	0.801
Benzaldehyde	A; RI; MS	1483	1508	Almond, burnt sugar	132.88	48.46	140.87	83.03	0.620	210.36	58.04	129.62	46.05	<b>0.007</b>
2,3-Butanediol	A; RI; MS	1507	1537	Fruity creamy buttery	760.13	377.03	638.99	371.64	0.855	5.73	96.72	274.20	127.04	<b>0.000</b>
1-Methoxy-2-propyl acetate	RI; MS	1532	1238	-	157.16	68.93	142.26	70.37	0.917	93.72	26.42	60.93	21.17	<b>0.010</b>
2-Ethoxy-propane	RI; MS	1544	-	-	210.52	217.09	238.27	260.01	0.958	41.90	62.40	155.71	66.81	<b>0.001</b>
4-Hydroxy-butanoate	RI; MS	1581	-	-	11.53	3.62	20.56	8.31	0.085	41.79	18.12	38.17	9.20	0.801
Acetophenone	A; RI; MS	1606	1627	-	39.31	15.38	37.86	20.28	0.715	94.19	24.57	74.41	25.65	0.231
Pentanoic acid	A; RI; MS	1642	1712	Sweaty	144.18	46.53	136.96	61.24	0.774	260.91	60.96	152.75	50.75	<b>0.002</b>

2-Phenylethyl acetate	A; RI; MS	1773	1785	Flowery	319.15	135.65	341.95	167.35	0.159	113.12	75.24	99.95	31.91	0.753
Hexanoic acid	RI; MS	1814	1816	Pungent, sweat	1.00	3.27	1.90	1.56	0.656	2.08	6.07	12.00	7.80	<b>0.043</b>
Phenylethyl Alcohol	A; RI; MS	1863	1912	Honey-like	811.52	374.55	813.07	436.74	0.938	415.25	172.26	339.70	110.91	0.900
Ethylbenzoate	A; RI; MS	1626	1641	Sweet, wintery, fruity, medicinal, cherry, grape	17.66	8.47	19.69	8.79	0.498	17.67	4.64	19.69	7.51	0.231
1,2-Dimethoxybenzene	A; RI; MS	1687	1699	-	0.14	0.08	0.08	0.44	0.916	1.05	2.34	7.21	9.83	<b>0.002</b>
Naphthalene	A; RI; MS	1689	1707	Pungent, dry, tarry	0.77	0.29	10.17	7.75	<b>&lt; 0.0001</b>	2.76	1.64	10.68	5.88	<b>0.000</b>
Isoamylbenzoate	A; RI; MS	1798	1921	Sweet, fruity, green and waxy	29.47	21.35	27.69	24.63	0.876	24.85	11.41	35.70	12.16	0.051
Guaiacol	A; RI; MS	1812	1823	Phenolic, smoke, spice, vanilla, woody	0.67	1.11	10.26	4.84	<b>&lt; 0.0001</b>	8.00	3.74	35.91	24.77	<b>0.000</b>
2-Methoxy-4-methylphenol	A; RI; MS	1906	1956	Sweet, candy, spice, eugenol, vanilla, leather, spicy, smoky	0.28	0.19	18.30	12.35	<b>&lt; 0.0001</b>	2.15	1.69	23.35	13.32	<b>&lt; 0.0001</b>
Phenol	A; RI; MS	1958	1965	Phenolic, plastic, rubber	1.26	0.69	86.79	57.07	<b>&lt; 0.0001</b>	9.75	7.35	51.12	33.11	<b>0.000</b>
1H-Pyrrole-2-carboxaldehyde	A; RI; MS	1968	-	-	0.68	1.79	2.67	1.09	<b>0.006</b>	18.19	4.81	16.68	5.56	0.614
<i>p</i> -Ethylguaiacol	A; RI; MS	1979	2032	Spicy, smoky, bacon, phenolic clove	0.25	0.10	18.51	14.05	<b>&lt; 0.0001</b>	3.07	2.08	8.33	14.65	0.378
<i>p</i> -Cresol	A; RI; MS	2040	-	Phenolic, narcissus, animal, mimosa	0.08	0.06	7.51	4.77	<b>&lt; 0.0001</b>	1.89	1.51	17.48	10.74	<b>&lt; 0.0001</b>
2-Phenoxyethanol	A; RI; MS	2085	2087	Weak, mild, rosy, balsamic, cinnamon-like	6.30	5.05	6.34	6.78	0.602	8.63	3.64	7.31	2.75	0.753
3-Ethylphenol	RI; MS	2123	-	Musty	0.13	0.06	2.19	1.33	<b>&lt; 0.0001</b>	0.40	0.94	4.11	2.44	<b>0.000</b>
2,6-Dimethoxyphenol	A; RI; MS	2204	2269	Sweet, phenol, smoky, medicinal, balsamic	0.03	0.10	27.66	22.37	<b>&lt; 0.0001</b>	1.76	1.58	19.77	12.66	<b>&lt; 0.0001</b>
4-Methyl-2,6-dimethoxyphenol	A; RI; MS	2289	-	Phenolic, smoky, woody, spicy, eugenol-like	n.d.	n.d.	12.45	10.85	<b>&lt; 0.0001</b>	0.44	0.44	7.26	3.89	<b>&lt; 0.0001</b>

**Table 3**

Compounds	smoky		non-smoky		LOQ (ng/g)	LOD (ng/g)
	Average (ng/g)	s.d.	Average (ng/g)	s.d.		
Naphthalene	32.5	10.7	4.8	3.3	3.0	0.9
Guaiacol	68.7	25.0	8.2	3.8	3.1	0.9
2-Methoxy-4-methylphenol	63.8	20.1	-	-	5.8	1.7
Phenol	721.8	482.2	5.7	6.2	1.0	0.3
<i>p</i> -Ethylguaiacol	83.0	8.0	-	-	32.5	9.7
<i>p</i> -Cresol	143.0	47.9	-	-	24.1	7.2

Figure 1  
Perotti P.

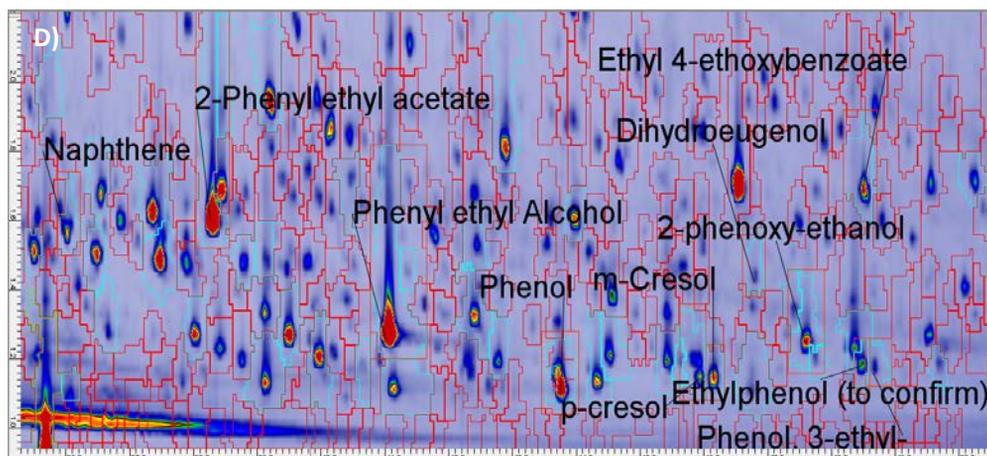
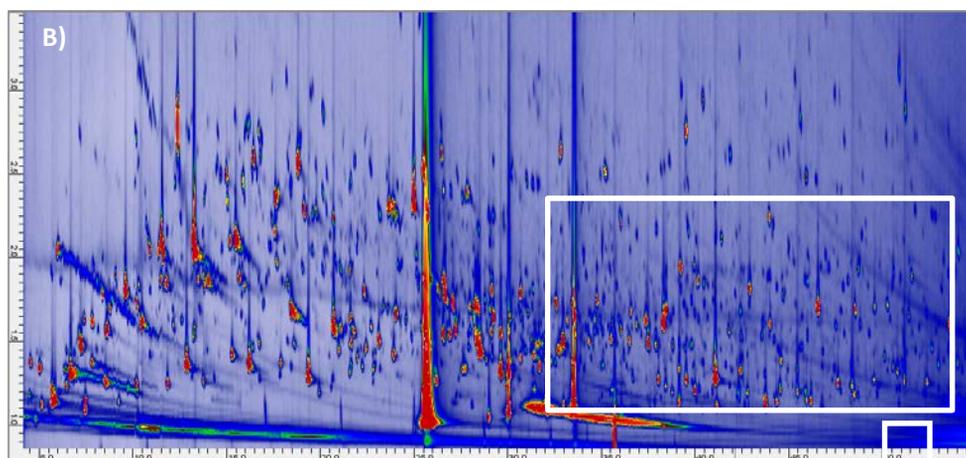
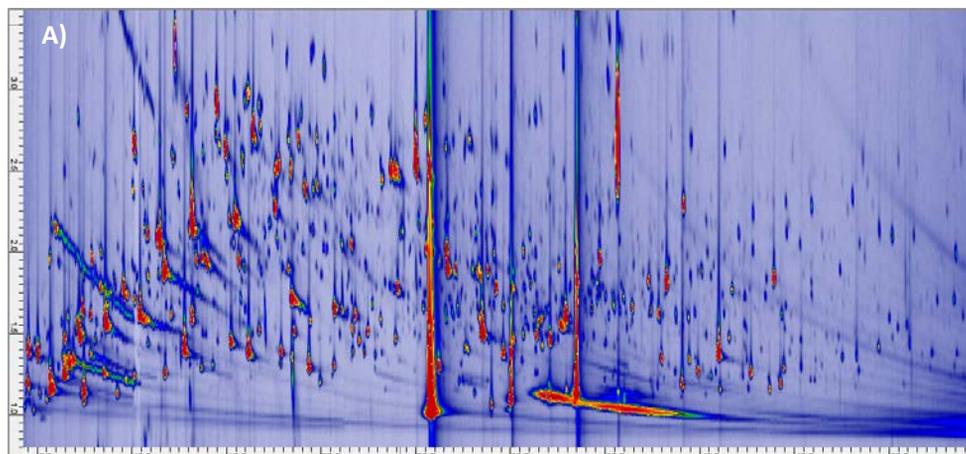
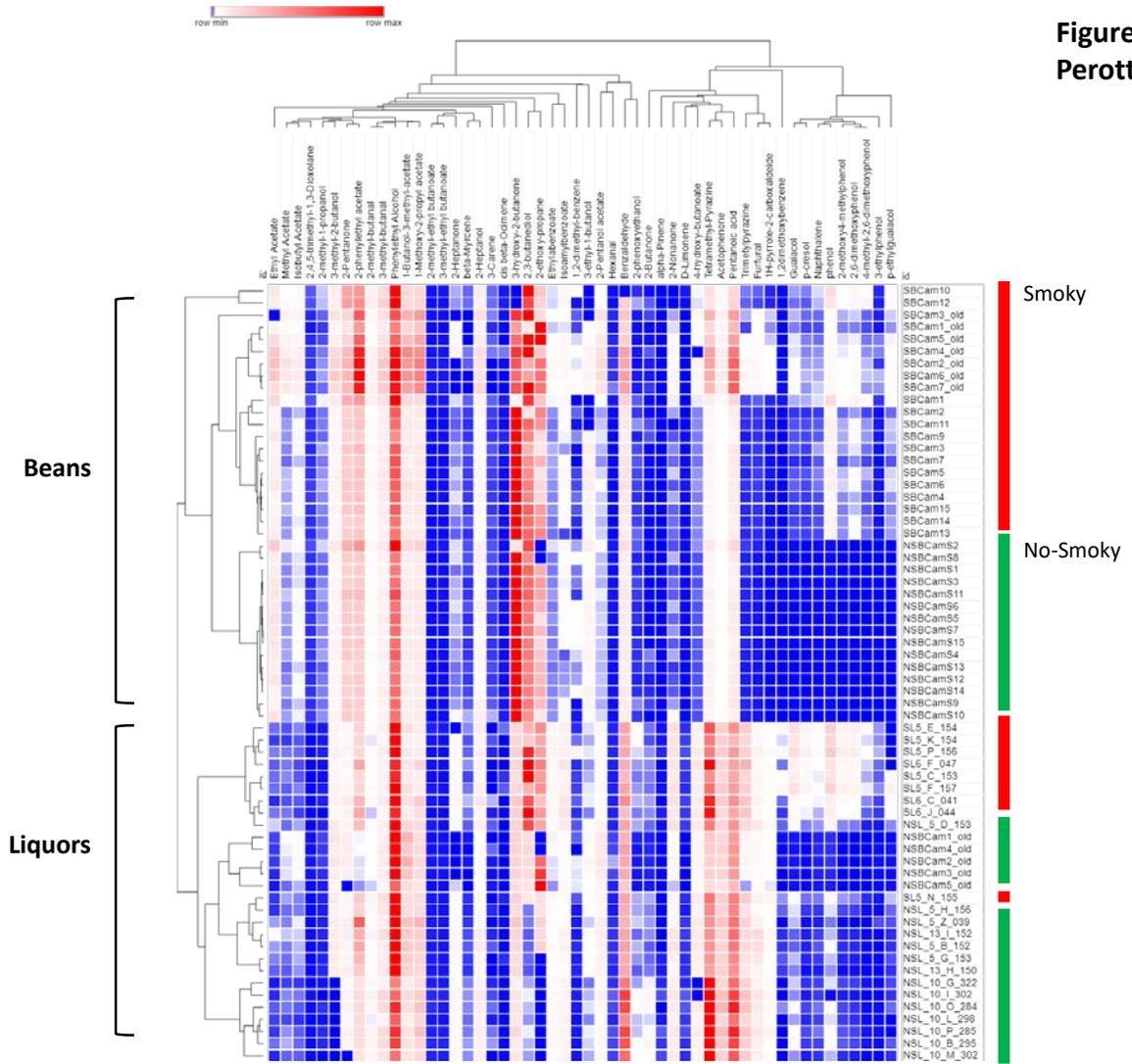
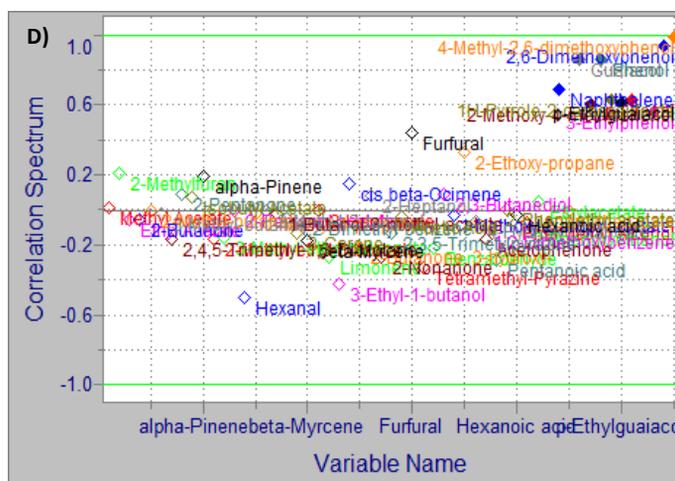
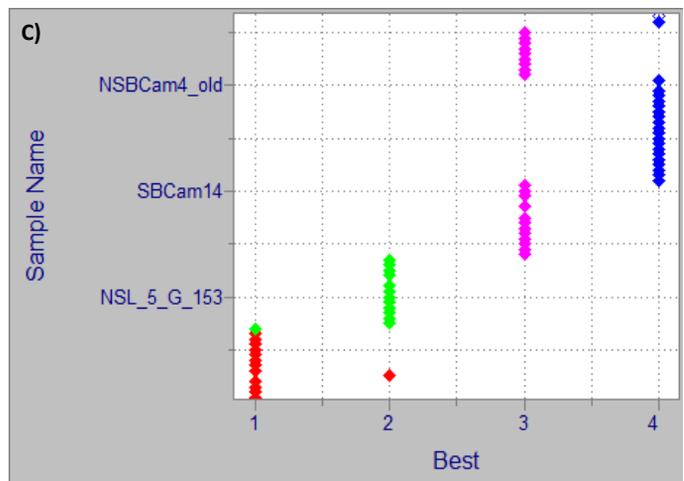
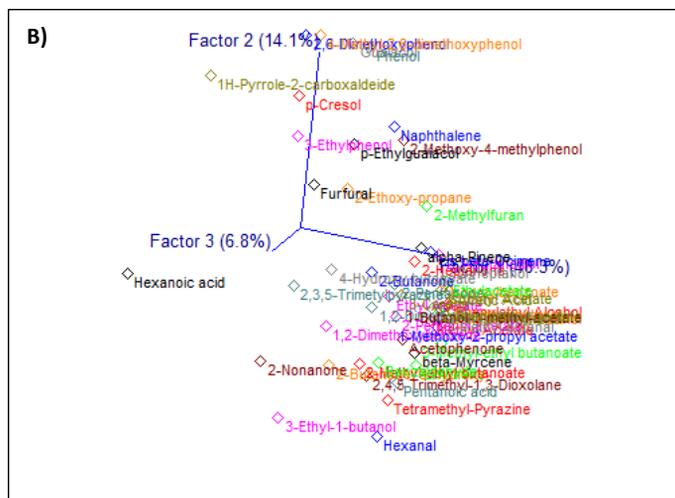
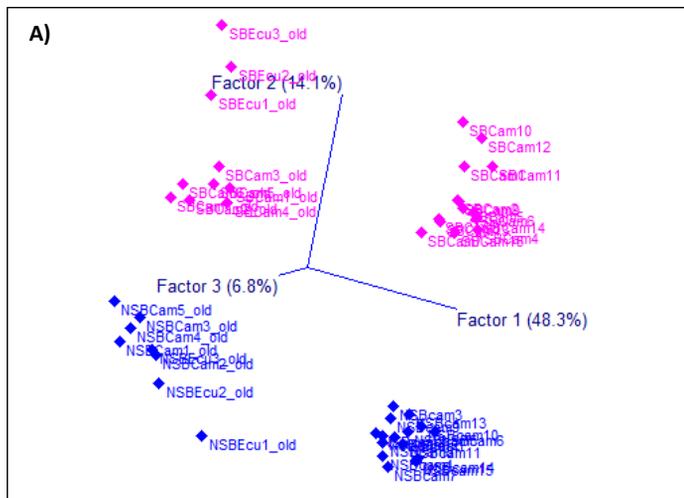


Figure 2  
Perotti P



**Figure 3**  
**Perotti P**



## Cocoa smoky off-flavor: chemical characterization and objective evaluation for quality control

Pamela Perotti<sup>1</sup>, Chiara Cordero<sup>1</sup>, Cristian Bortolini<sup>2</sup>, Patrizia Rubiolo<sup>1</sup>, Carlo Bicchi<sup>1</sup>, Erica Liberto\*

<sup>1</sup> Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino, Turin, Italy

<sup>2</sup> Soremartec Italia S.r.l., P.le P. Ferrero 1, 12051 Alba (CN), Italy

A careful tuning of the sample preparation conditions was required to improve informative analytes recovery. Results of the experimental design for HS-SPME sampling optimization are reported below together with the experimented parameters.

The following sampling conditions were tested (Kolb & Ettre, 2006):

- a) Sampling temperature (50 and 80°C)
- b) Evaluation of different sample amount (1-3 g) and polymer coatings (PDMS/DVB, DVB/CAR/PDMS, PDMS, and PDMS/DVB coated with PDMS for in-solution sampling)
- c) Modifying analytes' solubility

*Increase of the sampling temperature:* the sampling temperature was increased from 50°C to 80°C keeping constant the other conditions (1 g sampled for 40 min) to improve the transfer of the low volatility compounds to the headspace. The bar chart (Figure 1S a) shows the increase of the phenolic derivatives signals due to the effect of the temperature on the analytes.

Evaluation of different sample amounts and polymer coatings:

different coating fibers and amount of sample were tested in consideration of the nature of the investigated markers. As expected (Figure 1S b), DVB/CAR/PDMS fiber shows the highest analyte recovery. On the other hand, an increase of sample amount does not improve signals for PDMS/DVB coating, as already observed by Mejias et al. with other SPME coatings (Castro Mejías et al., 2003; Jelen, 2006).

c) *Changing analytes solubility:* analyte solubility in a food matrix can also influence their matrix/headspace partition coefficient, as well as their recoveries. Cocoa beans are solid where analytes are adsorbed, the strength of their interaction can therefore be modified through:

1) *Salting out applied to the suspended matrix:* suspension was made by adding a displacer (2 mL of water solution with 30% of NaCl) to 1 g of grinded beans. Suspension in a high ionic strength solution should afford both the displacement of the retained analytes in the liquid phase and a better partition between the displacer and headspace (Figure 1S c); this approach does not result in an increased abundance of the investigated analytes with the DVB/CAR/PDMS fiber.

2) *Salting out applied to water suspended samples in combination with direct immersion SPME*: 1 g of beans was suspended in 20 mL of a saturated NaCl water solution (Figure 1S d) and sampled with different polymeric coatings (PDMS/DVB, PDMS/DVB coated with PDMS).

Salting out of the suspended sample displays recoveries not comparable to those with untreated grounded beans at 80°C for almost all compounds investigated. Direct immersion shows a very poor recovery. Figure 1S d shows a comparison of the recovery of analytes under investigation with the different sampling approaches using PDMS/DVB coating.

## References

- Castro Mejías, R., Natera Marín, R., García Moreno, M. de V., & García Barroso, C. (2003). Optimisation of headspace solid-phase microextraction for the analysis of volatile phenols in wine. *Journal of Chromatography. A*, 995(1–2), 11–20. [https://doi.org/10.1016/S0021-9673\(03\)00541-7](https://doi.org/10.1016/S0021-9673(03)00541-7)
- Jelen, H. H. (2006). Solid-Phase Microextraction in the Analysis of Food Taints and Off-Flavors. *Journal of Chromatographic Science*, 44(7), 399–415. <https://doi.org/10.1093/chromsci/44.7.399>
- Kolb, B., & Ettre, L. S. (2006). *Static headspace-gas chromatography : theory and practice*. Wiley-Interscience.

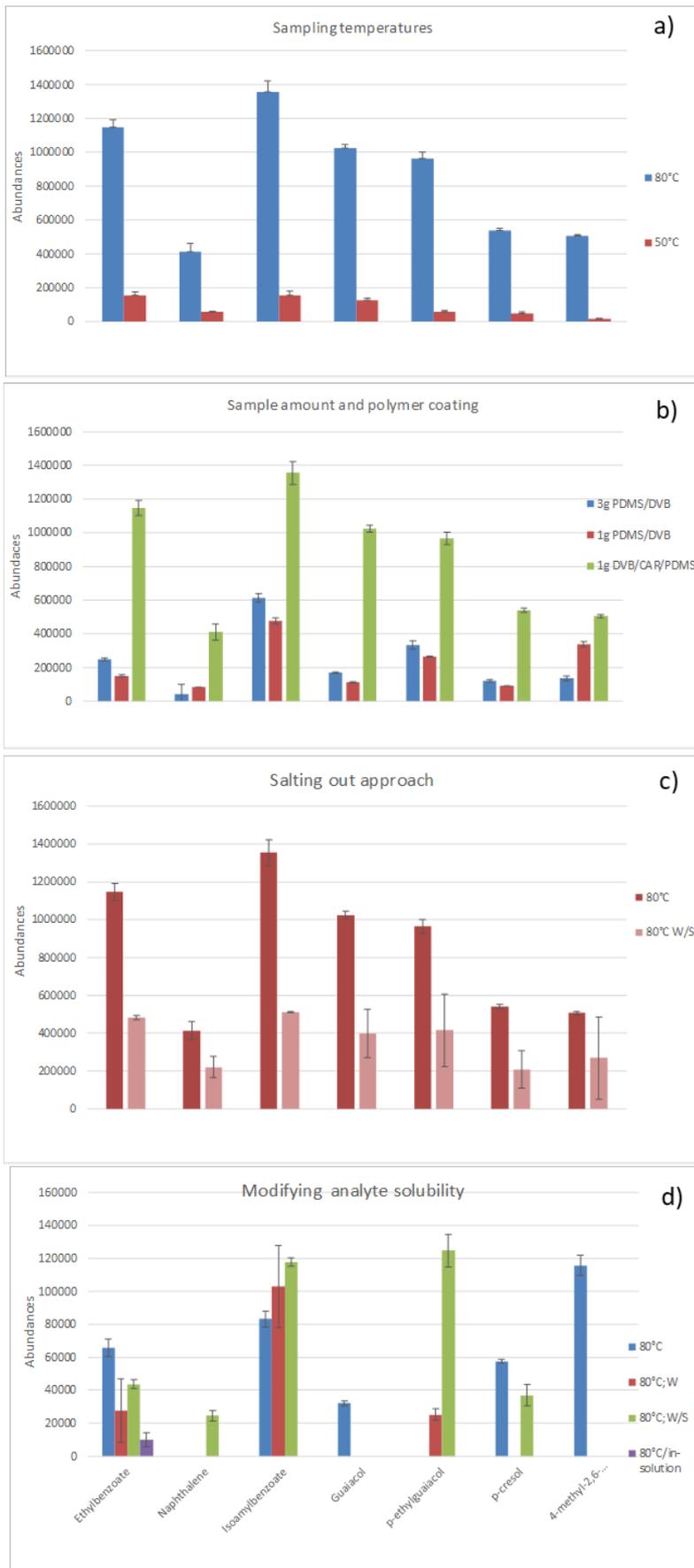


Figure S1.