IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Non-separative headspace solid phase microextraction-mass spectrometry profile as a marker to monitor coffee roasting degree

This is the author's manuscript

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/122256 since 2017-01-02T15:29:35Z

Published version:

DOI:10.1021/jf303067q

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)

19 January 2022

UNIVERSITÀ DEGLI STUDI DI TORINO

***This is an author version of the contribution published on**:*

[Erica Liberto, Manuela R. Ruosi, Chiara Cordero , Patrizia Rubiolo, Carlo Bicchi, Barbara Sgorbini,

Non-separative headspace solid phase microextraction-mass spectrometry profile as a marker to monitor coffee roasting degree*, JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY, Volume: 61 Issue: 8 Pages: 1652-1660 Published: FEB 27 2013, 10.1021/jf303067q, ACS]*

The definitive version is available at:

[http://pubs.acs.org/doi/full/10.1021/jf303067q]

Non-separative headspace solid phase microextraction-mass spectrometry profile as a marker to monitor coffee roasting degree

Erica Liberto[1]*, Manuela R. Ruosi[1], Chiara Cordero[1] , Patrizia Rubiolo[1], Carlo Bicchi[1], Barbara Sgorbini[1],

[1]Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino, Via P. Giuria 9, I-10125 Torino, Italy

*Address for correspondence:

Dr. Erica Liberto - Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino, Via P. Giuria 9, I-10125 Torino, Italy – e-mail: erica.liberto@unito.it; phone: +39 011 670 7662; fax: +39 011 670 7687

Abstract

This study describes a non-separative Headspace-Solid-Phase-Microextraction-Mass-Spectrometry (HS-SPME-MS) approach, in view of its application to on-line monitoring of a roasting process. The system can quickly provide representative and diagnostic fingerprints of the volatile fraction of samples and, in combination with appropriate chemometric pattern-recognition and regression techniques, can successfully be applied to characterize, discriminate and/or correlate patterns with the roasting process. Eighty coffee samples of different varieties, geographical origins and blends were analyzed. The experimental HS-SPME-MS results show that the TIC fingerprint can be used to discriminate the degree of roasting; diagnostic ion abundance(s) or ratios were closely correlated with the roasting process; both could successfully be used as markers or Analytical Decision Makers, to monitor roasting processes on-line, and to define quality and safety of roasted coffee.

KEYWORDS: HS-SPME-MS; MS-EN; coffee; aroma quality and safety; technological process; roasting indices; on-line monitoring

INTRODUCTION

The roasting process is a significant factor in determining coffee flavor. In particular, coffee aroma depends on the specific quali-quantitative distribution of various components. These are mainly volatile medium-to-high polarity compounds, deriving from drying and from the heat-browning related to the roasting conditions, above all temperature and time. The aroma also depends on the species, variety and blend, as well as on geographical origin[1-4]. Conversely, the roasting process can also produce some compounds that must be monitored because of their toxic properties; this is the case of furan and its derivatives, which have shown carcinogenic and cytotoxic activities[5-8]. The reduction of furan in coffee is therefore highly recommended, and may be achieved by optimizing the roasting process in all its steps (i.e. roasting, cooling, degassing and grinding) while, obviously, keeping the organoleptic properties unaltered.

The effect of the roasting process on coffee beans is generally described in terms of the degree of roasting, and is usually evaluated through several chemical and physical parameters, including external color of the beans, loss of weight during roasting, and variations in chemical composition, as well as through the development of sensory characteristics[9-10]. These parameters concur to define the degree of roasting, although to date a concise, clear and universally-accepted evaluation protocol does not exist. The most widely-used parameter to determine degree of roasting in day-to-day practice is color, measured through the light reflectance of ground beans or, still today, by visual inspection. This latter method is still valid partly because the industry uses a constant quality and variety of green coffee, combined with constant time-temperature conditions and roasting plant. Dry-matter loss is also considered to provide a reliable and complementary evaluation of the degree of roasting, including for in-plant determination[9]. Correct conditions for industrial roasting processes are, in general, obtained by appropriately scaling-up the results of experiments from pilot or laboratory plants monitored by physical methods.

This study describes a non-separative headspace solid phase microextraction-mass spectrometric method (HS-SPME-MS), in view of its possible application to the on-line monitoring of roasting processes. HS-SPME is a high-concentration-capacity technique offering good recoveries and easy to automate, that can be combined directly on-line with mass spectrometry[11]. Non-separative MS methods, better known as mass spectrometry-based electronic nose (MS-EN)[12-16], were introduced by Marsili[17] to study off-flavors in milk; they have since been applied successfully to characterizing several matrices, in particular in the food field[18-20]. They provide a representative, diagnostic and generalized mass spectrometric fingerprint of the volatile fraction of a sample, analyzed directly without prior chromatographic separation, in which each m/z ratio acts as a “sensor” whose intensity derives from the contribution of each compound producing that fragment. These methods, in combination with appropriate chemometric elaboration can be used to quickly characterize and discriminate samples within a set, and to correlate them with a technological process (e.g. coffee roasting). MS-EN can also be used to monitor

target compounds in a group of samples, provided that specific and diagnostic ions are obtained with a compatible ion generation mode (EI, CI, APCI, PTR etc.).
Mass spectral fingerprints, or diagnostic ion abundance(s), were here used both as marker and as analytical decision maker (ADM)[21] to monitor coffee roasting degree, in view of the possibility of combining a mass spectrometer directly on-line to a roasting machine. In particular, the results are reported of a study aimed at correlating HS-SPME-MS profile and coffee color, as a further tool to characterize roasting processes.

MATERIALS AND METHODS

Samples:

Eighty roasted-coffee samples from different geographical origins, of different varieties and blends (100% Arabica, 100% Robusta, 50/50% Arabica-Robusta blend, and several commercial blends) at different degrees of roasting were supplied by Lavazza (Turin – Italy) over a period of 12 months. Colors of samples were measured, by ground-bean light reflectance, with a single-beam Neuhaus Neotec Color Test II instrument, at wavelength 900 nm, on 25-30g of ground coffee. Table 1 gives the list of samples with their color values. A set of 100 coffee pods from the same lot (a 50/50 Arabica and Robusta blend) for espresso machines was stored at -18°C and used as reference to standardize the HS-SPME system performance over time.

Headspace Solid Phase Microextraction (HS-SPME) sampling and GC analysis conditions

The SPME device and fibers were from Supelco (Bellefonte, PA, USA). Divinylbenzene/Carboxen/ Polydimethylsiloxane (DVB/CAR/PDMS) d_f 50/30 μm, 2 cm long fibers were used, conditioning them before use as recommended by the manufacturer[22].

Volatiles were sampled by automated headspace Solid Phase Microextraction (auto-HS-SPME) using a MPS-2 multipurpose sampler (Gerstel, Mülheim a/d Ruhr, Germany) on-line integrated with an Agilent 7890 GC unit coupled to an Agilent 5975 MS detector (Agilent, Little Falls, DE, USA). 500 mg of ground roasted coffee in a 20mL screw-cap vial were sampled by HS-SPME at 50°C for 10 minutes[10]. All samples were analyzed in triplicate. Three fibers were selected from different lots, after preliminary testing to establish their sampling performance, so as to select equivalent fibers for use throughout the entire analysis period (see below: HS-SPME fiber performance evaluation). Fiber performance was monitored throughout the study on an additional set of coffee pod samples, and on 5uL of a 2 mg/mL standard solution of α- and β-thujone in dibutyl phthalate.

Non-separative analysis – GC unit conditions: oven and injector temperature: 250 °C injection mode: split, split ratio: 1/10; carrier gas: helium, flow rate: 0.4 mL/min; fiber desorption time and reconditioning: 3min; transfer column: deactivated fused silica tubing (d_c 0.10mm, length 6.70m) (Mega, Legnano (Milan), Italy).

MSD conditions: ionization: EI mode at 70 eV; transfer line: 280°C. Standard tuning was used and the scan range was set at m/z 35-350 with a scanning rate of 1.000 amu/s.

Separative GC-MS analysis - Chromatographic conditions: injector temperature: 230°C, injection mode: split, ratio: 1/10; carrier gas: helium, flow rate: 1 mL/min; fiber desorption time and reconditioning: 5 min; column: MEGAWAX 20M (d_f 0.20 μm, d_c 0.20 mm, length 50 m) (Mega, Legnano (Milan), Italy). Temperature program: from 40°C (1 min) to 200°C at 3°C/min, then to 250 °C (5min) at 10°C/min.

MSD conditions: ionization mode: EI (70 eV), scan range: 35 to 350 amu; ion source temperature: 230°C; quadrupole temperature: 150°C; transfer line temperature: 250°C.

Data were acquired and processed with an Agilent MSD Chem Station ver E.02.01.1177. (Agilent, Little Falls, DE, USA)

Analytes were identified by their linear retention indices and EI-MS spectra, by comparison with authentic standards, or were tentatively identified through their EI-MS fragmentation patterns and retention indices.

Chemometric analyses

Principal Component Analysis (PCA) and Orthogonal Partial Least Square analysis (OPLS) were performed with Pirouette software ver. 4.0 (Infometrix, Inc. Bothell, WA). The software package was used to automatically create ASCII files from Agilent GC ChemStation data, using a post-run macro. The data matrix consisted of as many rows as the number of samples (total objects: 240) and 316 columns (m/z

variables). Samples were randomly divided into a training set (55 samples) and a test set (25 samples). PCA was used in the first step for pattern recognition analysis, to visualize information and sample clusters, in particular as a function of the technological process. OPLS analysis was then carried out to correlate color (as a marker of the degree of roasting) to the chemical fingerprint. Data were pre-treated by baseline correction, through noise subtraction, and by internal normalization of the signal from each sample; they were subsequently pre-processed by mean-centering.

RESULTS AND DISCUSSION

Compared to separative GC-MS profiles, MS-EN patterns provide a fast response; however, the significance of the TIC is low, and less information can apparently be obtained from the MS profile. This is because the intensity of each fragment (m/z) is "built up" from the contributions of each component of the sample involved in generating that ion during its ionization process. Figure 1 gives the HS-SPME-GC-MS (a) and the HS-SPME-MS TIC (b) profiles, together with the mass spectral fingerprint (c) of an Arabica coffee sample. Table 2 lists the components identified in the HS-SPME-GC-MS profile of the same Arabica sample, together with their target (TI) and qualifier ions (Qi).

When applying HS-SPME-MS, it is mandatory to use chemometric techniques to "extract" data from the MS profile that can provide significant information. However, when a mathematical model is generated and used to classify or correlate composition or characteristics of a sample, and to reveal the resources of hidden information, one of the main limits is profile precision over a long period. This variability can reduce the effectiveness of a successful model, generated with a training set, when applied to routine analyses. The most critical points in the case of HS-SPME-MS are SPME fiber performance and MS signal instability; the latter may be due to ion source contamination, aging of electron multiplier, and/or filament electron emission. These limits can be overcome by i) evaluating the response of a suitable number of SPME fibers and monitoring it continually, ii) checking electron multiplier response over time, and iii) standardizing the MS profile through internal normalization.

HS-SPME fiber performance evaluation

SPME fiber effectiveness was evaluated initially in terms of total fingerprint areas: this was aimed at minimizing sampling errors/discriminations, in view of the extended time interval (12 months) covered by this study, and of the large number of samples and replicates expected to be run. As discussed elsewhere[21], the sampling performance of three DVB/CAR/PDMS fibers from different lots was tested on coffee pod samples and on an α- and β-thujone standard solution, to classify them on the basis of the recovery provided, and to monitor any change over time. Normalized spectral fingerprint areas of the coffee volatile fraction of five replicates from the same pod for each fiber (F1, F2, F3) were submitted to Analysis of Variance (ANOVA). The One-Way ANOVA results on the replicates for each fiber are in table S1 (supporting information), and showed that the null hypothesis ("there is no

difference among the fibers in terms of absolute peak areas of the selected target analytes“) was false. Tukey’s test classified F1 and F3 as belonging to the same group, and these were therefore adopted for all experiments. Similar results were obtained with α- and β-thujone standard solution. When additional fibers were necessary, they were submitted to the entire test routine, analyzing reference coffee-pod samples. Normalized spectral fingerprint area values had to fall within 5% variability (expressed as RSD%) as established during performance testing.

Precision and internal normalization

Precision, expressed as repeatability and intermediate precision of HS-SPME-MS on reference coffee-pod samples, was evaluated over the entire period. Repeatability was calculated over five analyses of coffee in the same pod, the MS profile fragments of each sample were normalized *versus* the most intense ions (m/z 43, basic peak) taken as 1; each m/z intensity value is expressed as a percentage of the intensity of the basic m/z fragment[14,24]. Repeatability is expressed as Relative Standard Deviation % (RSD%) on the total areas of the normalized fingerprints, and on some diagnostic m/z ions characteristic of certain components related to roasting, aroma or toxic chemicals. Intermediate precision ~~was calculated as before~~ was calculated as described above~~?// was calculated as described elsewhere?~~ over five analyses of the coffee pod, carried out monthly over a period of one year. Results are in table S2 (supporting information) and show very good intermediate precision, with RSD% values for the total fingerprinting area of 3.16%, and for single ions ranging from 4.65% for m/z 45 to 20.08% for m/z 150.

Unsupervised exploratory analysis

Although color is the most widely-adopted parameter used to monitor the roasting process, on-line monitoring of the development of the volatile fraction by means of an analytical approach (HS-SPME-MS) could be very useful to improve coffee standardization and to optimize its aroma. Fifty five coffee samples of different varieties and blends, and roasted to different degrees, i.e. dark (color range: 35÷42), medium (46÷53) and light (57÷62), were analyzed. Their spectral fingerprints were processed using an unsupervised approach, Principal Component Analysis (PCA), with the aim of finding inter-sample and inter-variable relationships with color (e.g. degree of roasting). This was done by visualizing ~~spots? variables? scores? components?~~ sample distribution on the score plots resulting from exploratory analysis[31-33]. PCA clearly showed that coffee samples are discriminated by their roasting degree (color) on the first principal component (1st PC), while different varieties are separated on the 2nd and 3rd PCs (Figure 2a). The volatile fraction included the well-known Maillard reaction products, derived from non-volatile precursor degradation (Amadori compounds and deoxyosones) namely furanes, furanones and pyranones. Compounds resulting from the Strecker reaction of α-dicarbonyls and aminoacids were also present: 2,3-butandione, 2- and 3-methylbutanal and alkyl-pyrazines, as were sugar degradation products (furfural, 5-methylfurfural and furfuryl alcohol). The loading plot (figure 2b) indicates those m/z

fragments that vary linearly with roasting degree; this indirectly also shows which components change during roasting: m/z 52 and 79 (mainly representative of pyridine and furfuryl acetate), 81 (1-methyl-pyrrole, 1-furfuryl-pyrrole and furfuryl acetate), 98 (furfuryl alcohol), 108 (dimethyl pyrazine groups: 2,3-, 2,5-, 2,6-isomers), 109 (guaiacol), 110 (5-methyl furfural), and 96 (furfural) all have high relevance~~// are all highly significant?~~ on the 1st PC (explained variance 59.90%), and can be taken as markers of roasting, as they increase markedly with the roasting degree. Some of them are also key-aroma components, and are marked with an asterisk in Table S2[25,26]. M/z 43 (2,3-hexandione, acetoxyacetone), 45 (acetoine and 3-hydroxy-2-butanone), 57 (2-oxopropyl propanoate), 60 (acetic acid), 79 (pyridine), and 95-96 (acetyl furan, furfural) on the 2nd PC (exp. var. 23.31%) are markers of variety and, roasting degree ~~being equal????,~~ are more abundant in Arabica than in the blend or Robusta samples; they decrease with roasting ~~time/temperature~~. Conversely, m/z 108 (mainly deriving from dimethyl pyrazines) and 109 (guaiacol) are more intense in Robusta samples. These considerations confirm the findings of a study by Ruosi et al. using separative techniques. [27]

Some m/z variables vary linearly with roasting degree, for example furan and its homologues. Their formation is related linearly to the degree of roasting, as shown in figure 3; determination coefficients were 0.7754 and 0.8917 for furan and 2-methylfuran, respectively. Despite the presence of interference by the same m/z fragment from other origins, this close correlation with degree of roasting means that furan and 2-methyl furan formation can be monitored during roasting via their characterizing ions (target ions), i.e. 68 and 82. If the intensities of these ions are outside fixed limits of acceptance, the relative compounds can be quantified by conventional methods [28,29]. In addition to conventional separative methods, in 2011 Bicchi et al. proposed a quick method to quantify these compounds by HS-SPME-MS; results were comparable to those of conventional methods[30].

"Supervised" multivariate regression

PCA results show that the MS profile of the coffee volatile fraction is closely correlated to the degree of roasting, thanks to some variables (m/z values, and the originating ~~// the relative?~~ components) that are characteristic of this technological process. Orthogonal Partial Least Square analysis (OPLS) was thus used as a measure of correlation, i.e. to evaluate the closeness of association between the fingerprint of the volatile fraction of coffee and its color, as an indicator of technological treatment, rather than for its ability to predict coffee color [33,34]. The OPLS model requires a training set to demonstrate any correlation between coffee color and aroma. Training (55) and test (25) sets of samples, selected randomly and consisting of several commercially-available Arabica (100%) and Robusta (100%) coffees of different origins, plus blends of the two at different percentages, were established.

The OPLS regression model was first calculated on 5 PCs, and internally cross-validated on the training set data relating to the volatile fraction spectral fingerprint and coffee color; the model was then applied to the test set samples. The results showed a highly negative linear correlation between measured (i.e. all m/z fragments

of the volatile fraction of coffee samples) and predicted (color values) variables. The correlation between spectral fingerprint and color is negative due to how color values are expressed, since high color values indicate lightly-roasted samples. ~~//In terms of predictive value, there was a close correlation (rpred 0.9472) with a satisfactory Standard Error of Prediction (SEP 2.53); values were within the range of colors that discriminate the different degree roasting (i.e. light, medium and dark).~~ TEMO DI NON AVER FATTO GIUSTO, CI RIPROVO// _When used in prediction mode, the correlation value was high (rpred 0.9472) with a satisfactory Standard Error of Prediction (SEP 2.53); values were within the range of colors that discriminate the different degree roasting (i.e. light, medium and dark).

Figure 4 shows the association intensity relationship between measured and predicted color values, using the model built with the training set, and gives the parameters of the validated model and of the prediction. As shown by PCA, the m/z ions with good correlation values (> ± 0.7) are related to compounds whose abundance is affected markedly by the roasting process, and that are linked to the aroma developing during this process (Fig 2b). The volatile fraction spectral fingerprint/color correlation is important to monitor the roasting process because, besides establishing a relationship between chemical and physical data, ~~it (sarebbe the correlation: OK?)~~ can give specific chemical information about aroma changes with technological treatments.

Identification of indices of the degree of roasting and possible relation with the utility of roasted coffee color

This part of the study investigated the possibility of identifying reliable chemical indices to be used for roasting process control, as a function of some informative m/z ion ratios. If correctly monitored on the laboratory scale, through objective and measurable indices, the roasting process could be directed so as to obtain the required aroma profile, in particular when new blends are being developed. Starting from the above PCA results, the two-by-two normalized area ratios of ions closely correlated to the roasting process were calculated, using a specific visual basic Excel macro. The resulting indices were multiplied by 1000, to facilitate their handling, and were arbitrarily considered significant only when there was a difference of at least 100 units between light, medium and dark roasting degree, in each variety or blend. Three indices (i.e. m/z ratios) were found to be in common among Arabica, Robusta and blends: 79/110; 97/110; 98/110; these fragments are characteristic, among others, of pyridine, 5-methyl furfural, 1-methyl pyrrole, furfuryl pyrrole and furfuryl alcohol. Each fragment varied linearly with color and, as a consequence, with roasting degree. Table 3 gives correlation equations and coefficients, between indices and color, for each variety/blend, together with the average index values and related standard deviation. These index values can successfully be used to discriminate the degree of roasting with good confidence, as is shown by their standard deviations that, within the same variety, avoid any risk of two index intervals deriving from two different roasting treatments (e.g. medium or dark) overlapping.

These results are particularly interesting for industrial applications, where in general the material being processed comprises blends of green coffees with similar or

unvaried characteristics (variety, origins etc.); they enable a physical parameter (color) to be correlated to chemical markers (indices) for the purpose of monitoring the roasting process. Although the number of samples and varieties tested in the present study is too small to be representative, these results outline the possibility of defining indices of roasting and significant values for on-line process monitoring.

In addition, robust and reliable mathematical models to predict color directly from the total spectral fingerprint and/or index values can successfully be applied on-line, not only to predict color, but also to connect it to the aroma composition. A PLS model equation was built and internally cross-validated with the 55 samples of the training set and verified with the 25 samples of the test set (table 4). The results show a close correlation between indices (i.e. m/z ratios) and color, with 17 of 25 samples having a residual color measure below ± 3, and only two of them (C21 and C24) presenting a difference between measured and predicted color of approximately– 9. Table 4 reports the PLS results, together with the parameters of the model equation, and shows a close correlation within the training set (rval: 0.9399) and a very satisfactory standard error in color validation (SEV: 2.68). When the model equation was applied to the samples of the test set, as expected the correlation coefficient in prediction decreased (rpred: 0.8416) and the standard error in prediction increased (SEP: 4.29), although the values were still satisfactory. These results show a reliable correlation, although a less uniform training set would be necessary to include a wider range of variables (variety, origin, roasting conditions) in a single equation. The results are very satisfactory if the diversity represented is considered, because the samples of the training set were very different and their number was relatively small.

In conclusion, the results show that HS-SPME-MS for on/in-line control of coffee roasting process is a promising approach, through which not only the evolution of the total MS profile can be studied, but also specific ions or ion ratios. The quick non-separative method (HS-SPME-MS) described shows that a correlation between spectral fingerprinting or roasting indices and color can be found, and that chemical parameters can be used reliably to evaluate the degree of roasting. The combination of MS profile and chemical indices with physical parameters affords more reliable off-line monitoring or optimization of the roasting process, in particular to control aroma development when developing new blends, and to detect the formation of toxic compounds. In addition, the reliability of these results may be exploited as a reference to validate those obtained by coupling a laboratory roasting machine to a mass spectrometer directly, for on-line monitoring of the roasting process and marker development[35-38] HS-SPME-MS can also be used as analytical decision maker, i.e. to select those sample(s) that have to be analyzed by conventional separative methods, for instance when the non-separative intensities of some m/z fragments, diagnostic of certain components, are outside the fixed limits of acceptance, or when the aroma profile is not in line with that of the desired final product. Further studies are now under way to extend the applicability of this method from model experiments to real-world samples.

ACKNOWLEDGMENT

The results here reported are part of the "ITACA" project of the POR-FESR "Competitività regionale e occupazione" 2007/2013, Asse 1, Misura I.1.1, "Piattaforme innovative" of the Piedmont Region (Italy), to which the authors are indebted.

SUPPORTING INFORMATION AVAILABLE

One-way Anova results on fiber performance evaluation of the total volatile fraction, spectral fingerprint areas of coffee pod samples, and method precision over 12 months, are included as supporting material. This material is available free of charge via the internet at http://pubs.acs.org.

LITERATURE CITED

[1] Flament, I. *Coffee Flavor Chemistry*; Wiley: Chichester, U.K., 2002.

[2] Nebesny, E., Budryn, G., Kula, J., Majda, T. The effect of roasting method on headspace composition of Robusta coffee bean aroma. *Eur. Food Res. Technol.* 2007, 225, 9-19.

[3] Moon, J., Shibamoto, T. Role of roasting conditions in the profile of volatile flavour chemicals formed from coffee beans. *J. Agric. Food Chem* 2009, 57, 5823-5831.

[4] Cordero, C., Liberto, E., Bicchi, C., Rubiolo, P., Reichenbach, S.E., Tian, X., Tao, Q. Targeted and non-targeted approaches for complex natural sample profiling by GC×GC–qMS. *J Chromatogr Sci.* 2010, 48, 1-12.

[5] Furan in foods-thermal treatment, US Food and Drug Administration (FDA), Washington, DC, 2004. www.cfsan.fda.gov/<lrd/fr040510.html (27/05/2012).

[6] Report of the scientific panel on contaminants in the food chain on provisional findings on furan in food, European Food Safety Authority (EFSA), EFSA J. 2004 137, 1. Brussels, Belgium; www.efsa.eu.int/science/contam/contam documents/760 en.html (27/05/2012).

[7] Crews, C., Castle, L. A review of the occurence, formation and analysis of furan in heat-processes foods. *Trends Food Sci. Technol.* 2007, 18, 365-372.

[8] IARC, Monographs on the evaluation of carcinogenic risks to human and their supplements 1995, 63, 393.

[9] Clarke, R. J., Macrae, R. In *Coffee: Vol 2 Technology*.; edition 1; R. J. Clarke & R. Macrae Eds., Essex: Elsevier Applied Science, 1987.

[10] Baggenstoss, J., Poisson, L., Perren, R., Escher, F. Coffee roasting and aroma formation: application of different time-temperature conditions. *J. Agric. Food Chem.* 2008, 56, 5836-5846.

[11] Arthur, C. L., Pawliszyn, J. Solid phase microextraction with thermal desorption using fused silica optical fiber. *Anal. Chem.* 1990, 62, 2145-2148.

[12] Fenaille, F., Visani, P., Fumeaux, R., Milo, C., Guy, P.A., Comparison of mass spectrometry-based electronic nose and solid phase microextraction gas chromatographys-mass spectrometry technique to assess infant formula oxidation. *J. Agric. Food Chem.* 2003, 51, 2790-2796.

[13] Pilar Marti, M.; Busto, O.; Guasch, J. Application of headspace mass spectrometry system to the differentiation and classification of wines according to their origin, variety and ageing. *J. Chromatogr. A* 2004, 1057, 211-217.

[14] Pavón, J. L. P., Sanchez, M.; Del Nogal; Pinto, C. G.; Laespada, E. F.; Cordero, B. M.; Pen˜ a, A. G. Strategies for qualitative and quantitative analyses with mass spectrometry-based electronic noses. *TrAC - Trends Anal. Chem.* 2006, 25, 257-266.

[15] Karlshøj, K., Nielsen, P.V., Larsen, T.O. Prediction of Penicillium expansum Spoilage and Patulin concentration in apples used for apple juice production by Electronic Nose Analysis. *J. Agric. Food Chem.* 2007, 55, 4289-4298.

[16] Jelen, H.H., Ziolkowska, A., Kaczmarek, A. Identification of the botanical origin of raw spirits produced from rye, potato, and corn based on volatile compounds analysis using a SPME-MS Method. *J. Agric. Food Chem.* 2010, 58, 12585-12591.

[17] Marsili, R.T. SPME−MS−MVA as an electronic nose for the study of off-flavors in milk. *J. Agric. Food Chem.* 1999, 47, 648-654.

[18] Gardner, J.W., Philip, P.N. In *Electronic Noses: Principles and Applications*, Oxford University Press, New York, 1999.

[19] Marcos Lorenzo, I., Pérez Pavón, J.L., Fernandez Laespada M.E., Garcia Pinto C., Moreno Cordero B. Detection of adulterants in olive oil by headspace-mass spectrometry. *J. Chromatogr. A* 2002, 945, 221-30.

[20] Fisk, I.D., Kettle, A., Hofmeister, S., Virdie, A., Kenny, J.S. Discrimination of roast and ground coffee aroma. *Flavour*, 2012, 1, 1-14.

410 [21] Sandra, P., David, F., Tienpont, B. From CGC-MS to MS-based analytical decision makers.
411 *Chromatographia* 2004, 60, S299–S302.
412 [22] Cordero, C., Bicchi, C., Rubiolo, P. Group-Type and Fingerprint Analysis of Roasted Food Matrices (Coffee
413 and Hazelnut Samples) by Comprehensive Two-Dimensional Gas Chromatography. *J. Agric. Food Chem.*
414 2008, 56, 7655–7666.
415 [23] Bicchi, C., Cordero, C., Liberto, E., Sgorbini, B., Rubiolo, P. Reliability of fibres in solid-phase
416 microextraction for routine analysis of the headspace of aromatic and medicinal plants. *Journal of Chromatogr.*
417 *A*, 2007, 1152, 138-149.
418 [24] Pavon, J. L. P., Sanchez, M. Del N., Pinto, C. G.; Laespada, E. F.; Cordero, B. M., Pena, A. G. A Method for
419 the detection of hydrocarbon pollution in soils by Headspace Mass Spectrometry and Pattern Recognition
420 Techniques. *Anal. Chem.* 2003, 75, 2034-2041.
421 [25] Czerny M., Mayer F., Grosch W. Sensory study on the character impact odorants of roasted *Arabica* coffee.
422 *J.Agric. Food Chem.*, 1999, 47, 695–699.
423 [26] Blank, I.; Sen, A., Grosch, W. Potent odorants of the roasted powder and brew of *Arabica* coffee. *Z.*
424 *Lebensm. Unters Forsch,* 1992, 195, 239-245.
425 [27] Ruosi, M.R., Liberto, E., Cordero, C., Cagliero, C., rubiolo, P., Bicchi, C., Sgorbini, B. A Further Tool To
426 Monitor Coffee Roasting Process: Aroma Composition And Chemical Indices. *J.Agric. Food Chem.*, in revision.
427 [28] Determination of furan in foods, US Food and Drug Administration (FDA), Washington, DC, 2004.
428 www.cfsan.fda.gov/<dms/furan.html.
429 [29] Exploratory data on furan in food, US Food and Drug Administration (FDA), Washington, DC, 2004.
430 www.cfsan.fda.gov/<dms/furandat.html. (27/05/2012)
431 [30] Bicchi, C., Ruosi, M.R., Cagliero, C., Cordero, C., Liberto E., Rubiolo, P., Sgorbini, B. Quantitative analysis
432 of volatiles from solid matrices of vegetable origin by high concentration capacity headspace techniques:
433 Determination of furan in roasted coffee. *J. Chromatogr. A,* 2011, 1218, 753-762.

[31] Cozzolino, D., Smythb, H. E., Cynkara W., Janika, L., Dambergsa, R. G., Gishena, M. Use of direct headspace-mass spectrometry coupled with chemometrics to predict aroma properties in Australian Riesling wine. *Anal. Chim. Acta*, 2008, 621, 2-7.

[32] Armanino, C., Casolino, M.C., Casale, M., Forina, M. Modelling aroma of three Italian red wines by headspace-mass spectrometry and potential functions. *Anal. Chim. Acta*, 2008, 614, 134-142.

[33] Brereton R.G., In Applied Chemometrics for Scientist, Chapter 5-6, John Wiley & Sons Eds, Ltd West Sussex, 2007, 145-220.

[34] Berenson M.L., Levine D.M. In Statistica per le scienze economiche, Chapter 16, Zanichelli Eds, Bologna, 1993, 549-612.

[35] Wieland F., Gloess A.N., Keller M., Wetzel A., Schenker S., Yeretzian C. Online monitoring of coffee roasting by proton transefr reaction time-of-flight mass spectrometry (PTR-ToF-MS) towards a real-time process control for a consistent roast profile . *Anal. Bioanal. Chem,* 2012, 402, 2531-2543.

[36] Wieland F., Gloess A.N., Keller M., Wetzel A., Schenker S., Yeretzian C. Highlights of analytical chemistry in Swizerland. *Chimia,* 2012, 66, 443-443.

[37] Romani S., Cevoli C., Fabbri A., Alessandrini L., Dalla Rosa M. Evaluation of Coffee Roasting Degree by Using Electronic Nose and Artificial Neural Network for Off-line Quality Control. *J. of Food Sci,* 2012*,* 77, 960-965.

[38] Alessandrini L., Romani S., Pinnavaia G., Dalla Rosa M. Near infrared spectroscopy: An analytical tool to predict coffee roasting degree. *J Anal. Chim. Acta*, 2008, 625, 95-102.

Figure captions

Figure 1 HS-SPME-GC-MS (a), HS-SPME-TIC-MS (b) profiles and mass spectral fingerprint (c) of an Arabica coffee sample.

Figure 2 Scores (a) and Loading plots (b) of 55 different roasted coffee fingerprints (first 3PCs exp. var. 94.36%). Pre-processing: mean-center, full-cross validation. Categories: light roasting (solid triangle; color 57÷62), medium roasting (empty diamond; color 46÷53), dark roasting (solid diamond; colour 35÷42). Robusta: solid line; blend 50/50: dotted line; Arabica: dashed line.

Figure 3 Regression model for furan and 2-methylfuran on 55 coffee samples *versus* the degree of roasting, here represented by the experimentally-measured color.

Figure 4 OPLS regression model for color prediction as an association measure between volatile fraction, spectral fingerprint, and color of 25 commercial samples, whose origins, varieties and blends are unknown. [a] Correlation coefficient in Prediction, [b] Standard error in Prediction, [c] Correlation coefficient in Validation, [d] Standard error in Validation

Figure 1 HS-SPME-GC-MS (a), HS-SPME-TIC-MS (b) profiles and mass spectral fingerprint (c) of an Arabica coffee sample.

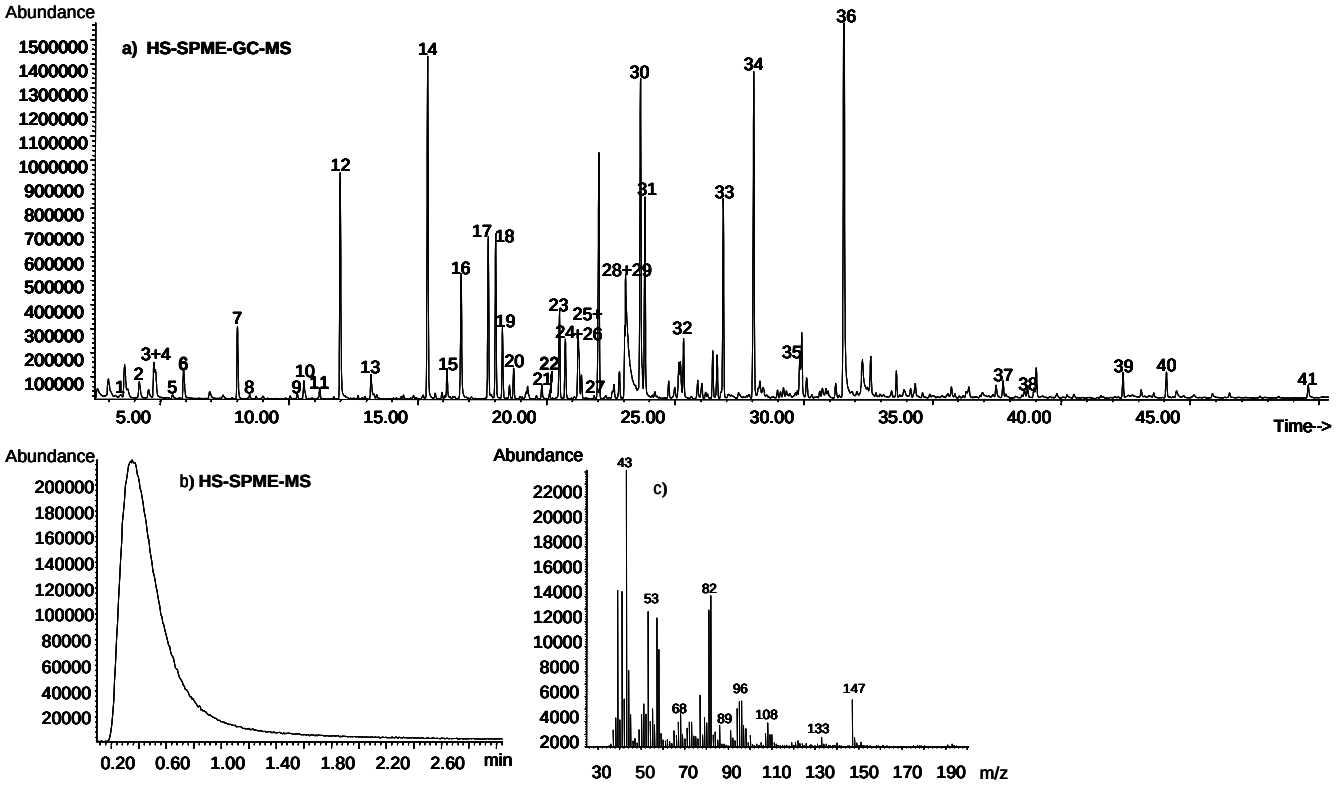
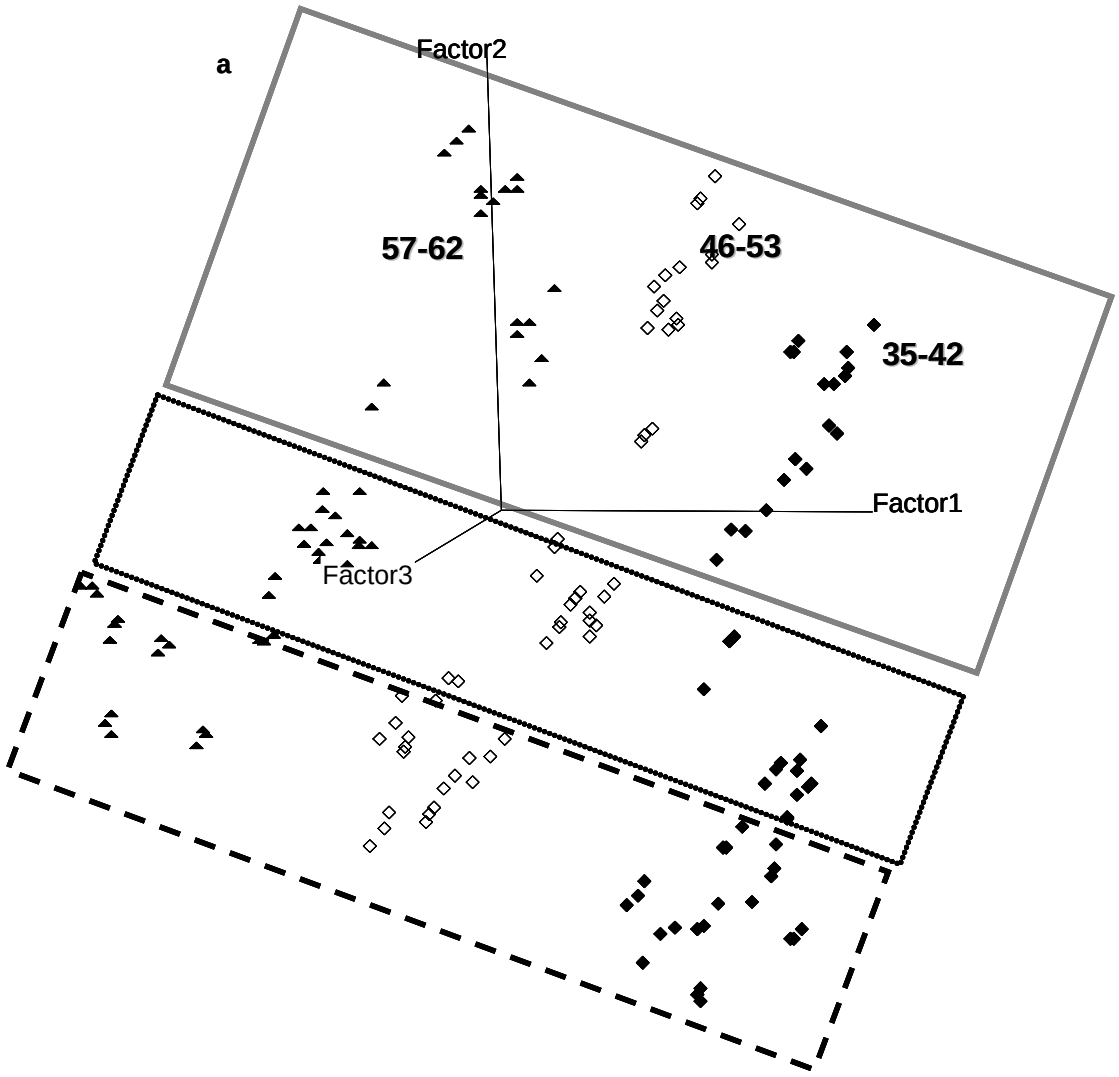
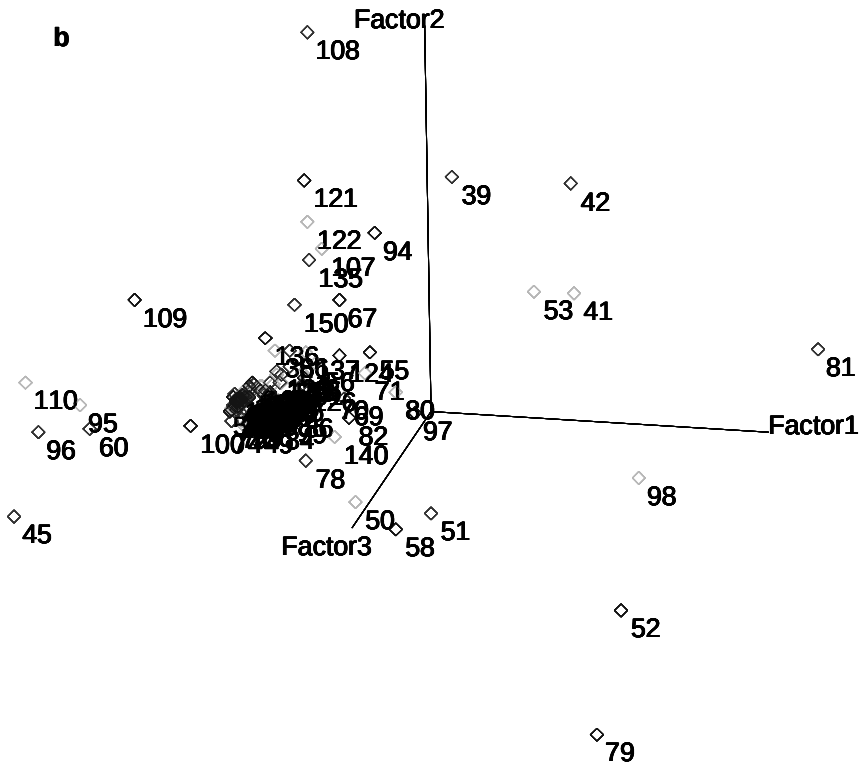


Figure 2 Scores (a) and Loading plots (b) of 55 different roasted coffee fingerprints (first 3PCs exp. var. 94.36%). Pre-processing: mean-center, full-cross validation. Categories: light roasting (solid triangle; color 57÷62), medium roasting (empty diamond; color 46÷53), dark roasting (solid diamond; colour 35÷42). Robusta: solid line; blend 50/50: dotted line; Arabica: dashed line.

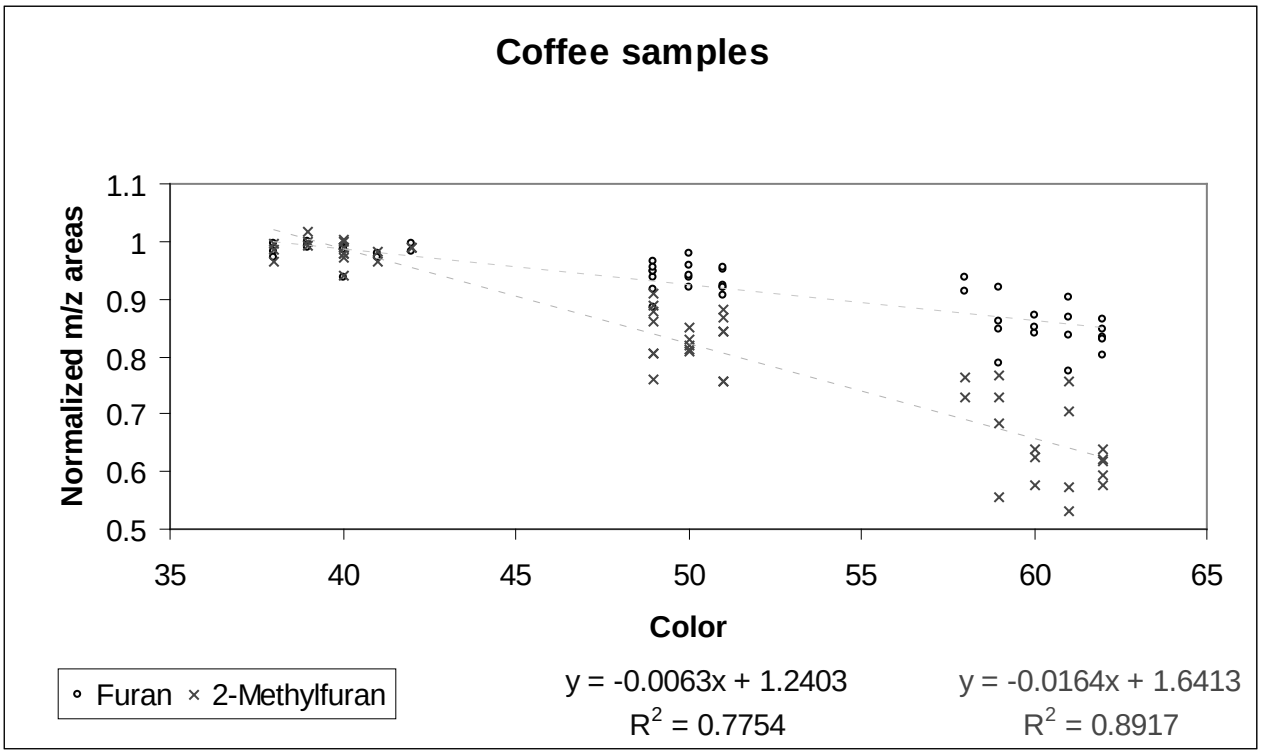



b
Factor2
Factor1
Factor3
108
121
39
42
122
94
107
133
109
150
67
53
41
81
110
95
96
60
71
80
97
82
140
78
98
50
58
51
45
52
79

548 Figure 3 Regression model for furan and 2-methylfuran on 55 coffee samples *versus* the degree of roasting

549 here represented by the experimentally measured color.

550



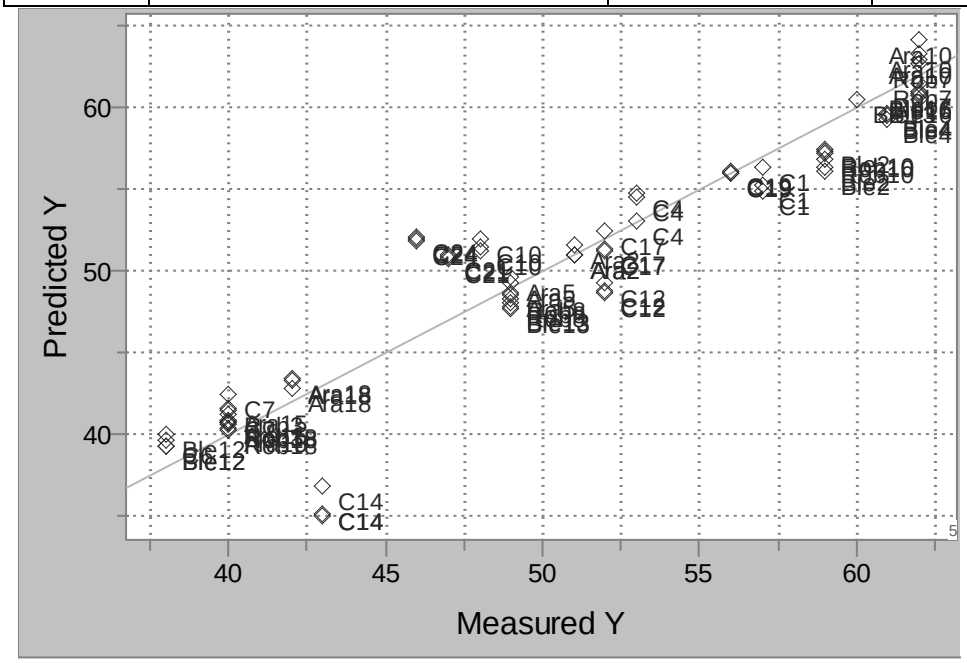
551

552

553 Figure 4 OPLS regression model for color prediction as an association measure between volatile fraction
554 spectral fingerprint and color of 25 commercial samples, whose origins, varieties and blends are unknown. [a]
555 Correlation coefficient in Prediction, [b] Standard error in Prediction, [c] Correlation coefficient in Validation, [d]
556 Standard error in Validation

557

	Slope	Intercept	rpred[a]	SEP[b]	Model validation	rval[c]	SEV[d]
Prediction	0.91	4.32	0.9472	2.53	error on 5 Factors	0.9916	1.0024



560

561

562

563

564 Table 1 List of samples together with varieties and blends, color values and degrees of roasting .

Sample Code	Varieties and blends	Color Values	Roasting degree	Sample Code	Varieties and blends	Color Values	Roasting degree
Ara1	Arabica	61	Light	Rob13	Robusta	58	Light
Ara2	Arabica	51	Medium	Rob14	Robusta	50	Medium
Ara3	Arabica	40	Dark	Rob15	Robusta	40	Dark
Rob1	Robusta	61	Light	Ble13	Blend 50/50	60	Light
Rob2	Robusta	50	Medium	Ble14	Blend 50/50	49	Medium
Rob3	Robusta	40	Dark	Ble15	Blend 50/50	38	Dark
Ble1	Blend 50/50	59	Light	Ara16	Arabica	61	Light
Ble2	Blend 50/50	49	Medium	Ara17	Arabica	51	Medium
Ble3	Blend 50/50	39	Dark	Ara18	Arabica	42	Dark
Ara4	Arabica	59	Light	Rob16	Robusta	62	Light
Ara5	Arabica	49	Medium	Rob17	Robusta	50	Medium
Ara6	Arabica	40	Dark	Rob18	Robusta	40	Dark
Rob4	Robusta	60	Light	Ble16	Blend 50/50	62	Light
Rob5	Robusta	50	Medium	Ble17	Blend 50/50	49	Medium
Rob6	Robusta	41	Dark	Ble18	Blend 50/50	39	Dark
Ble4	Blend 50/50	61	Light	C1	Commercial	57	Light
Ble5	Blend 50/50	51	Medium	C2	Commercial	53	Medium
Ble6	Blend 50/50	42	Dark	C3	Commercial	47	Medium
Ara7	Arabica	62	Light	C4	Commercial	53	Medium
Ara8	Arabica	51	Medium	C5	Commercial	58	Light
Ara9	Arabica	39	Dark	C6	Commercial	39	Dark
Rob7	Robusta	62	Light	C7	Commercial	56	Light
Rob8	Robusta	49	Medium	C8	Commercial	58	Light
Rob9	Robusta	40	Dark	C9	Commercial	43	Dark
Ble7	Blend 50/50	60	Light	C10	Commercial	48	Medium
Ble8	Blend 50/50	51	Medium	C11	Commercial	39	Dark
Ble9	Blend 50/50	40	Dark	C12	Commercial	52	Medium
Ara10	Arabica	62	Light	C13	Commercial	50	Medium
Ara11	Arabica	50	Medium	C14	Commercial	43	Dark
Ara12	Arabica	41	Dark	C15	Commercial	48	Medium
Rob10	Robusta	59	Light	C16	Commercial	61	Light
Rob11	Robusta	51	Medium	C17	Commercial	52	Medium
Rob12	Robusta	38	Dark	C18	Commercial	47	Medium
Ble10	Blend 50/50	58	Light	C19	Commercial	56	Light
Ble11	Blend 50/50	49	Medium	C20	Commercial	40	Dark
Ble12	Blend 50/50	38	Dark	C21	Commercial	47	Medium
Ara13	Arabica	59	Light	C22	Commercial	57	Light
Ara14	Arabica	49	Medium	C23	Commercial	48	Medium
Ara15	Arabica	40	Dark	C24	Commercial	46	Medium
				C25	Commercial	52	Medium

569 Table 2 List of markers identified in HS-SPME-GC-MS profile of an Arabica roasted sample together with their

570 target (TI) and qualifier (Qi) ions. *Markers tentatively identified through their MS-EI fragmentation patterns

ID	Compounds	Ret Time (min)	I^T_{CW}	I^T_{OV1}	TI	Qi1	Qi2
1	Furan	3.74	837	500	68	58	39
2	2-methylfuran	4.49	873	586	82	81	53
3	2-methyl butanal	5.09	903	641	86	57	41
4	3-methyl butanal	5.09	904	635	86	71	57
5	2,5-dimethylfuran	5.86	938	691	96	95	81
6	2,3-butanedione	6.31	960	555	86	57	43
7	2,3-pentandione	8.49	1043	668	100	57	43
8	2-vinylfuran*	9.00	1059	-	94	65	66
9	2,3-hexanedione	10.88	1117	756	43	71	43
10	1-methyl-pyrrole	11.18	1124	715	81	80	66
11	2-vinyl-5-methylfuran*	11.79	1139	-	108	107	79
12	Pyridine	12.62	1165	720	79	52	39
13	Pyrazine	13.85	1195	709	80	53	70
14	methylpyrazine	16.08	1249	802	94	67	53
15	3-hydroxy-2-butanone	16.84	1265	681	88	73	45
16	1-hydroxy-2-propanone	17.42	1278	626	74	43	41
17	2,5-dimethylpyrazine	18.45	1306	893	108	81	42
18	2,6-dimethylpyrazine	18.74	1313	889	108	81	42
19	ethylpyrazine	19.01	1318	895	107	108	80
20	2,3-dimethylpyrazine	19.45	1329	904	108	67	93
21	1-hydroxy-2-butanone	20.57	1353	732	88	57	42
22	3-ethyl-pyridine	20.85	1364	934	107	92	79
23	2-ethyl-6-methylpyrazine	21.23	1371	977	121	122	94
24	2-ethyl-5-methylpyrazine	21.46	1376	981	121	122	94
25	trimethylpyrazine	21.96	1389	984	122	81	42
26	2-ethyl-3-methylpyrazine	22.01	1391	985	121	122	80
27	2-propylpyrazine	22.60	1402	985	94	107	122
28	2-ethyl-3,6-dimethyl pyrazine	23.78	1427	1059	135	136	108
29	acetic acid	23.79	1432	547	60	43	45
30	furfural	24.45	1443	801	96	95	39
31	1-acetoxy-2-propanone	24.57	1448	825	43	86	116
32	2-acetylfuran	26.12	1483	882	95	110	39
33	furfuryl acetate	27.64	1521	963	81	98	140
34	5-methyl furfural	28.86	1551	933	110	109	81
35	1-methyl-2-carboxaldehyde pyrrole	30.67	1596	974	109	108	80
36	furfuryl alcohol	32.37	1640	823	98	81	69
37	1-furfuryl pyrrole	38.58	1805	1152	81	147	53
38	guaiacol	39.54	1832	1064	109	124	81
39	2-acetyl pyrrole	43.27	1941	1030	94	109	66
40	2-carboxaldehyde pyrrole	44.98	1991	976	95	94	66
41	p-vinylguaiacol	50.47	2163	1289	150	135	107

571

572

573 Table S1 One-way ANOVA results on Total Volatile fraction spectral fingerprint areas of coffee pod samples

574 (n=5; α=0.05).

Groups	*Count*	*Sum*	*Mean*	*Variance*
Fibre 1 (F1)	5	103.41	20.68	2.38
Fibre 2 (F2)	5	134.87	26.97	0.22
Fibre 3 (F3)	5	103.14	20.63	0.52

One-Way Anova

Source of variation	*SQ*	*gdl*	*MQ*	*F*	*Significant Value*	*F crit*
Between groups	133.08	2	66.54	63.86	4.01E-07	3.89
Within groups	12.50	12	1.04			
Total	145.59	14				

575

576

577 Table S2 Method precision (repeatability and intermediate precision) of total fingerprint and some
578 characteristic ions expressed as Relative Standard Deviation % calculated over replicates collected across the
579 entire period (12 months). * m/z ions characterizing key-aroma compounds according to Czerny[23] and Blank[24].
580

m/z	*Characterizing properties*	*Repeatability*	*Intermediate precision*
37	Roasting	1.30	14.15
38	Roasting	1.10	9.32
39	Aroma and roasting	0.38	5.06
41	Aroma and roasting	0.84	16.27
45	Aroma and varieties	2.25	4.65
50	Aroma and roasting	0.36	9.29
51	Aroma and roasting	1.06	8.04
52	Aroma and roasting	1.72	8.50
53	Aroma and roasting	2.27	6.50
60	Aroma and varieties	1.70	8.00
62	Roasting	1.63	8.00
66	Roasting and varieties	1.95	9.28
67	Roasting	0.99	8.48
68	Toxic	1.92	7.84
69	Aroma and roasting	1.70	10.37
70	Aroma and roasting	2.19	22.77
78	Roasting	*1.67*	9.29
79	Roasting	1.22	12.67
80	Roasting	0.58	10.35
81	Roasting and varieties	1.60	9.10
82	Toxic	0.58	6.50
86*	Aroma	2.62	7.62
88	Aroma and varieties	4.62	16.12
94	Aroma and roasting	1.14	13.56
95	Roasting and varieties	1.29	7.68
96	Roasting	2.22	9.82
97	Roasting	2.27	5.26
98	Roasting	2.31	7.67
100*	Aroma	4.47	13.74
108	Aroma and roasting	0.50	11.96
109*	Aroma and roasting	1.34	6.98
110	Roasting	1.70	7.86
121	Aroma and roasting	2.49	9.98
122	Aroma and roasting	2.54	11.32
135*	Aroma	2.44	5.19
150*	Aroma	11.24	20.08
Total Fingerprint Area		1.27	3.16

581

582

583 Table 3 Indices (m/z ratios) and related correlation equation and coefficients with the color for each variety
584 and/or blend analyzed together with each index value and its standard deviation.
585

Index m/z ratios	Variety or blend	Equation	r2	Index value	Index Standard deviation	Mean color value
79/110	100% ARABICA	y = -119x + 8018	0.8754	1051	226	60
				1765	358	50
				3426	600	40
	BLEND	y = -110x + 7604	0.9872	1161	207	60
				1796	310	50
				3361	343	40
	100% ROBUSTA	y = -63x + 4874	0.9550	1148	111	60
				1621	135	50
				2407	123	40
97/110	100% ARABICA	y=-51x+3884	0.8946	879	76	60
				1239	143	50
				1898	139	40
	BLEND	y=-48x+3872	0.9912	1013	101	60
				1382	154	50
				1979	139	40
	100% ROBUSTA	y=-41x+3571	0.9925	1148	118	60
				1529	90	50
				1958	107	40
98/110	100% ARABICA	y=-108x+8043	0.8896	1692	172	60
				2438	303	50
				3846	325	40
	BLEND	y=-101x+7913	0.9918	1960	206	60
				2725	333	50
				3971	285	40
	100% ROBUSTA	y=-81x+7034	0.9939	2203	252	60
				2967	171	50
				3818	229	40

586
587

588 Table 4 Color measured with the three indices through PLS elaboration and model parameters

sample	Measured Y	Predicted Y	Residual Y	Upper Limit	Lower Limit
Ara2	51	52	-0.71	57	46
Rob3	40	45	-4.59	50	39
Ble1	59	56	2.73	61	51
Ara5	49	49	0.02	54	44
Ble4	61	56	4.82	61	51
Rob7	62	57	5.13	62	52
Rob8	49	48	1.12	53	43
Ara10	62	59	2.74	65	54
Rob10	59	54	5.18	59	49
Ble12	38	38	-0.43	44	33
Ara15	40	40	0.32	45	34
Ble13	60	57	3.21	62	52
Ble14	49	48	0.80	53	43
Ara18	42	42	0.08	47	37
Rob18	40	43	-2.90	48	38
Ble16	62	58	4.25	63	53
C19	56	62	-5.84	67	57
C21	47	57	-9.79	62	52
C24	46	55	-9.26	60	50
C14	43	36	6.66	42	31
C17	52	52	-0.43	58	47
C1	57	56	1.27	61	50
C4	53	54	-0.91	59	49
C6	39	41	-2.05	47	36
C10	48	55	-6.76	60	50
C12	52	53	-1.22	58	48
Prediction Equation	rpred	SEP	Model validation error on 2 Factors	rval	SEV
y=0.754x+12.71	0.8416	4.29		0.9399	2.68

589