DOI:10.1021/jf901673d

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

Identification of Ethyl Formate as a Quality Marker of the Fermented Off-note in Coffee by a Nontargeted Chemometric Approach

Christian Lindinger,*,† Philippe Pollien,† Ric C. H. de Vos,§,# Yuri Tikunov,§,# Jos A. Hageman,⊥ Charles Lambot,⊗ René Fumeaux,† Elisabeth Voirol-Baliguet,† and Imre Blank△

†Nestlé Research Center, Vers-chez-les-Blanc, 1000 Lausanne 26, Switzerland, §Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands, #Centre for Biosystems Genomics, P.O. Box 98, 6700 AB, Wageningen The Netherlands, ⊥Biometris−Applied Statistics, P.O. Box 100, 6700 AC, Wageningen, The Netherlands, ⊗Nestlé Centre R&D Tours, Avenue Gustave Eiffel, 37097 Tours, France, and △Nestlé Product Technology Center, 1350 Orbe, Switzerland

The quality of coffee is influenced by many factors such as coffee variety, agricultural and postharvest conditions, roasting parameters, and brewing. The pleasure of drinking coffee may be affected by off-notes such as burnt, green, earthy, or fermented. Their presence is related to the variety, fermentation during postharvest processing, or over-roasting of the beans. Sensory expert panels trained for the evaluation of coffee are able to detect off-notes and select coffees by well-defined quality criteria. The application of instrumental approaches detecting quality markers related to the perceived off-notes is shown to be useful to assist sensory panels. This paper describes the discovery of a new marker compound related to the fermented off-note occasionally perceived in coffees. The application of untargeted chemometric methods on volatile compounds revealed correlations between individual compounds and the sensory attribute. The new marker compound was identified as ethyl formate, which can be measured in the headspace of roasted and ground coffee by various analytical techniques including online proton transfer reaction mass spectrometry.

KEYWORDS: Quality; marker; ethyl formate; coffee; fermented; off-note; chemometrics

INTRODUCTION

The perceived aroma and flavor of a roasted and ground (R&G) coffee extract depends on many factors. The green coffee variety and quality with its specific composition of precursors sets the stage for the later flavor development during roasting. However, many other factors influence the quality of the coffee cherries, that is, agricultural, meteorological, and harvest conditions. Postharvest treatment (drying and fermentation) and process parameters (roasting, quenching, and grinding) are additional factors with a high impact on the final cup quality (*1*−*4*).

Monadic sensory evaluation by a trained sensory panel can describe the perceived quality differences of coffee samples. These quality differences are reflected by the chemical composition (*5*−*8*) that may substantially vary depending on the origin and history of coffee treatment. In addition, quality differences may be due to compounds leading to off-notes, which will change the overall quality from the predefined sensory profile. Specific ethyl esters of short-chain fatty acids have been reported as potential key aroma compounds responsible for the overfermented flavor defect of coffee (*9*, *10*). Their individual contribution is a function of the actual concentration in a given sample. It was established that elevated amounts of these off-flavor compounds may occur not only in wet-processed Arabica coffees but also in unwashed Robusta coffees. The occurrence of an overfermented flavor defect is not limited to inappropriate green coffee processing.

In general, approximations by first-principle models for estimating the quality using chemical and process data are difficult because of the high degree of complexity when considering chemical pathways in the formation of aroma and flavor. Therefore, the application of chemometrics to coffee is interesting because it is a purely data driven approach and not based on fundamental chemical theories. It also allows including process parameters, agricultural data, and quality descriptors as well as chemical data as long as the data are reproducible and unbiased (*11*, *12*). First attempts in correlating chemometric data sets with sensory profiles focusing on chemical differences showed promising results when chemometric methods were applied (*13*).

In general, two different chemometric strategies can be chosen, that is, the targeted or nontargeted approach. In the targeted approach, the set of data is predetermined (e.g., nature and quantity of components in the case of chemical data) and multivariate analysis aims at studying correlations between the resulting data and predefined target information. This approach has

*Author to whom correspondence should be addressed (telephone +41 (21) 785-9384; fax +41 (21) 785-8554; e-mail christian.lindinger@rdls.nestle.com).

 Published on Web 10/09/2009

long been applied to analytical chemistry. Its major drawback is its limitation to already known and identified components. Investigations of unknown but possibly relevant chemical markers are not considered, as the data analyses are restricted to predetermined compounds defined from the very beginning.

In contrast, no preselection of data takes place in the untargeted approach: all chemical information present in the data sets is taken into account prior to compound identification (*14*). For instance, information on the relative presence of all compounds present in the coffee headspace and detectable by GC-MS can be used to compare samples and link variation in GC-MS profiles to, for instance, postharvest and processing parameters. Such untargeted approaches require unbiased mass peak extraction and alignment of peaks over all samples, due to slight shifts in retention from one measurement to another. Even though GC separation techniques are already highly developed, usually resulting in stable chromatography, small chromatographic shifts may still be on the order of a full peak width, and peaks therefore need to be aligned. Dedicated metabolomics processing programs such as Metalign (www.metalign.nl) have been shown to be very useful in peak extraction from GC-MS data sets and alignment of chemically separated compounds without the need for prior compound identification (*14*). Multivariate data analysis can be performed using the full set of detected signals from both known and (yet) unknown compounds. Depending on the aim of the study, the identification of compounds can be focused on only those representing significant differences in quality or correlation with a certain quality trait. The main advantage of such essentially untargeted approach is the possibility to find novel markers for a specific target. Finally, the putative markers observed by the untargeted approach might be verified and validated by applying a targeted approach on a selected set of samples.

In the present work we discovered a new chemical marker identified as ethyl formate related to the fermented off-note occasionally perceived in coffees. Zlatkis and Sivetz (*15*) already identified ethyl formate in coffee without studying correlations with sensory data. As the fermented off-note is a well-known issue for coffee growers and the coffee trade, the simplicity of analyzing the marker by GC-MS and online instruments such as PTR-MS (*16*) is highlighted.

MATERIALS AND METHODS

Materials. Ethyl formate (CAS 109-94-4, ≥99.5%, mass spectrum (electron ionization) *m*/*z* 43 (100%), 73 (25%), 42, (11%), 59 (12%), retention index DB-Wax of 820) and methyl acetate (CAS 79-20-9, ≥99.9%, mass spectrum (electron ionization) *m*/*z* 31 (100%), *m*/*z* 28 (74%), *m*/*z* 29 (74%), *m*/*z* 27 (47%), *m*/*z* 45 (33%), *m*/*z* 26 (25%), *m*/*z* 74 (10%), retention index DB-Wax of 816) were from Sigma-Aldrich (Buchs, Switzerland). Tenax TA traps (Tenax TA 60−80 mesh, 15 g) were from phase Separation Ltd. (Deeside, U.K.).

Coffee Sample Selection. Robusta coffee (*Coffea canephora*) of different types (Kouillou and Robusta) and groups (Congolese, Guinean and hybrids), grown in identical defined and reproducible conditions in Ecuador, harvested in 2004 and 2005, were selected for their expected sensory differences. The selection was mainly based on quality markers determined for crop 2002 and genetic fingerprints based on molecular markers (RFLP markers). Fifteen samples (samples 1−15) were used for applying chemometric methods on sensory data and chemical fingerprints. A further three samples (sample A−C) of type Robusta and different groups (Congolese, Guinean, and hybrids) with different fermented off-note intensities were selected by an independent coffee expert sensory panel of four coffee experts. Sample C was selected with high fermented intensity, B with medium-high intensity, and A with medium intensity.

Sample Preparation for Technical Sensory Evaluation. For each of the 15 coffee samples, 750 g of green beans was split into five portions of 150 g and roasted to five different roasting degrees using a Neotec roaster (Neuhaus Neotec GmbH, Reinbek, Germany). The roasting temperature was kept constant, and the roasting time was varied to cover the range from low (CTn 109), to medium (CTn 95), to high (CTn 90), to city (CTn 79), to Italian roast (CTn 71) (defined by color table Neuhaus CTn value). The roasting time was adapted to each coffee sample to avoid over-roasting. Therefore, all samples were roasted to two roasting heat-loads at 227 °C/250 s and 227 °C/300 s, and the color of the beans was compared with the color table. On the basis of the results the three missing roasting heat-loads were adjusted depending on the individual roasting characteristic by varying the roasting time.

All coffee samples were ground to Ditting 5.5 (Ditting Maschinen AG, Bachenbülach, Switzerland) and packed into cans of 50 g. Coffee beverages were prepared from 50 g of R&G coffee per liter of water at 80 °C extraction using a mixture of two-thirds Vittel and one-third demineralized tap water and a coffee machine with a Melitta filter. The coffees were kept at 65 °C in Thermos flasks and served in polystyrene cups for tasting by the panelists.

Sample Preparation for Monadic Sensory Evaluation. For each coffee sample, 960 g of green beans was roasted to the individual roasting degrees evaluated during the technical sensory evaluation using a Neotec roaster. All coffee samples were ground to Ditting 5.5 and packed into cans of 50 g. For the monadic sensory sessions, coffees were prepared in the same way as for the technical sensory evaluation.

Sample Preparation for Sensory Evaluation of the Validation Set. The same sample preparation was used for the validation set (A−C) but roasted to medium degree (CTn 100).

Sample Preparation for Chemical Analysis. Two technologies were used to obtain chemical fingerprints of volatile compounds released from R&G coffee extract: online headspace measurement method and offline desorption of Tenax traps to GC. The same sample preparation was used for both analytical methods. Two grams of R&G coffee (same 15 R&G coffees as used for the monadic sensory evaluation and 3 R&G coffees used as validation set) was filled into the filter holder of an espresso machine (FrancisFrancis X1). Extraction was performed at 90 °C using a mixture of two-thirds Vittel water and one-third demineralized tap water. It has to be noted that the small amount of coffee filled into the filter holder does not allow for a high-pressure extraction; it is closer to filter coffee extraction. The extraction volume was set to 28 mL of coffee beverage.

Technical Sensory Evaluation. The appropriate roasting conditions for each coffee sample were investigated in technical tasting, using a panel of nine coffee experts. The five roasting levels of each coffee were presented to the panelists, and they were asked to rank/rate the sensory attributes bitter, acid, roasty, and rubbery on a 12 cm visual analogue scale from 0 to 12. In addition, they were asked to identify the presence of off-notes among the following attributes: burnt, green, earthy/woody, fermented, phenol/medical, and cereal. Finally, the panelists were asked to select one "ideal" roasting level based on a balanced bitterness/acidity, presence of characteristic attributes, and minimal perceived off-notes, that is, green due to a too low roasting level and burnt due to over-roasting. A discussion with the panelists after each evaluation led to a consensus on the "best choice" of roasting degree.

Monadic Sensory Evaluation. Fifteen coffee samples roasted to the individual roasting degrees evaluated during the technical sensory evaluation were used to obtain monadic sensory profiles. A panel of 10 external panelists, already trained in coffee evaluation, participated during four sessions of 90 min each. Even though the panel was already familiar with the glossary and products, training sessions were organized to select the appropriate sensory attributes as described in ref *17*. The coffees were profiled using the quantitative descriptive analysis (QDA) method. Seventeen selected attributes describing odor, flavor, mouthfeel, and persistence were scored on a visual analogue scale (0 = not, 10 = very intense). Samples were tasted monadic according to a complete block design and repeated twice. The statistical significance of each sensory attribute was validated by processing an analysis of variance (ANOVA). Attributes such as bitter, roasted, body, rubbery, coffee odor, coffee flavor, fermented, and chemical were significantly different between the samples with a range/LSD of >1.5 (range = maximum − minimum of evaluated intensity; least significant difference (LSD) obtained by ANOVA). For this approach standard mathematics was used for calculations with the programs NCSS 2007 (Hintze, J. NCSS, PASS, and

Table 1. Monadic Sensory Evaluation of 15 Samples Roasted to the Selected Individual Roasting Degree[a]

sample	coffee odor	coffee flavor	roasted	bitter	rubbery	chemical	fermented	body
1	3.61	4.17	4.11	4.92	1.06	0.47	0.55	3.57
2	4.18	4.05	4.22	5.34	1.22	0.45	1.53	3.39
3	3.76	3.74	5.14	5.65	1.13	0.73	1.73	3.69
4	2.50	2.65	3.52	4.33	1.06	0.93	3.96	3.24
5	5.08	4.81	5.18	5.11	0.97	0.36	0.61	4.03
6	3.69	3.68	3.45	3.79	0.47	0.55	1.49	2.62
7	2.80	3.19	4.10	4.63	0.66	0.95	3.96	3.29
8	4.66	4.63	4.36	4.37	0.45	0.18	0.45	3.49
9	4.17	4.46	5.30	5.50	1.36	0.50	1.15	3.48
10	4.21	4.37	4.19	4.26	0.74	0.16	0.74	3.34
11	4.80	4.52	3.23	3.80	0.49	0.07	0.38	3.54
12	4.49	4.50	5.81	4.90	0.57	0.15	0.63	3.53
13	2.95	3.90	4.30	4.73	0.71	0.35	2.56	3.50
14	3.78	3.37	3.24	3.77	1.13	0.76	2.42	3.47
15	2.43	2.90	1.58	3.74	0.73	0.82	3.63	3.18
range/LSD	2.59	2.16	1.73	2.60	1.68	1.64	3.37	2.33

[a] The attributes rubbery, body, bitter, roasted, coffee odor, coffee flavor, chemical, and fermented were scored on an analogue scale (0 = not, 10 = very intense). Statistical significance expressed by the range/LSD value of each sensory attribute was calculated by processing an analysis of variance (ANOVA) (range = maximum − minimum of evaluated intensity; least significant difference (LSD) obtained by ANOVA).

GESS; NCSS: Kaysville, Utah, 2006) and Excel (Microsoft). The mean values of the sensory evaluation, and range/LSD values are shown in **Table 1**.

Chemical Profiling and Identification. Directly after extraction of 28 g of espresso, the coffee was poured into a double-jacketed, water-heated sample vessel (350 mL glass vessel), mounted inside an oven set to 65 °C with active air circulation. A temperature-stabilized water bath (set to 50 °C) was connected to the double-jacketed cell to keep the sample at constant temperature. The sample cell was connected to the fix-mounted top cover to be easily disconnected and filled with the coffee sample. The coffee headspace was purged continuously with 300 standard cubic centimeters per minutes (sccm) through heated tubes (80 °C) penetrating the fix-mounted top of the cell. Before analysis by online PTR-MS (High Sensitivity PTR-MS, Ionicon Analytik GmbH, Innsbruck, Austria), the sample gas was diluted with 3000 sccm of dry air, preventing saturation of the instrument. Lindinger et al. (*18*) have described the complete setup in detail. After preparation of the coffee sample, the headspace vessel was disconnected, filled with the coffee sample, and connected back to the system. This procedure was carried out quickly (< 10 s) to avoid temperature changes of the sample, oven, and headspace vessel. PTR-MS instrumental parameters were set as follows: drift tube pressure, 2 mbar; drift tube temperature, 80 °C; drift voltage, 550 V; extraction blend, 6 V.

Tenax traps obtained during the online headspace measurement were desorbed at 250 °C for 10 min on the Automatic Thermo Desorption (ATD Turbo Matrix 350, Perkin-Elmer, Boston, MA), flushed with a helium flow of 20 sccm, cryofocused at −30 °C, and injected from the cold trap (250 °C, 3 min) into the GC (Agilent 6890 GC-oven). As column, a 60 m DB-Wax (i.d. = 0.32 mm) with a film thickness of 0.5 μm was kept at 20 °C for 20 min, increased at 4 °C/min to 220 °C, and maintained for 10 min at 220 °C. To identify the separated compounds, a Pegasus III TOF-MS (LECO Corp.) was used. The measurement in high time resolution allows deconvolution of the overlapping peaks as frequently observed in coffee spectra. Identification was performed by comparison with mass spectral databases (internal and Wiley databases) and retention indices of reference compounds.

With only a few exceptions, usually more than one compound contributes to a single ion trace monitored in coffee headspace samples by online PTR-MS. Therefore, the coupling of the GC-TOF-MS technique with PTR-MS was necessary to identify the chemical compounds contributing to specific ion traces measured online by PTR-MS (*18*). Three data sets were obtained for each coffee sample, that is, (i) GC-TOF-MS data of coffee headspace desorbed from Tenax traps (Centrotype notation: "EI" followed by a unique centrotype number); (ii) GC-PTR-MS data of coffee headspace desorbed from Tenax traps (Centrotype notation: "PT" followed by a unique centrotype number); and (iii) PTR-MS online headspace data (time intensity profiles in scan mode m/z 20−160).

GC-PTR-MS analysis allowed detection of 266 aligned peaks (PT centrotypes). A partial identification using simultaneously obtained GC-TOF-MS spectra allowed identification of 166 compounds based on internal databases (notation by molecule name). If the identification was not possible, the notation "U" followed by a unique number was used.

Quantification. For quantification of ethyl formate, 2 mL of coffee extract was filled into 10 mL glass vials closed with Teflon-lined septum metal caps. After 10 min of equilibration at 50 °C, 1 mL of headspace was sampled by using a 2.5 mL headspace syringe and injected into the GC-PTR-MS split/splitless injector (split factor of 2). The same column and oven temperature program was used as described above. The concentration of ethyl formate was determined in duplicate by external calibration with a standard solution of pure ethyl formate in water at concentrations between 0.1 and 0.6 mg/L.

Chemometric Methods. Metalign software (www.metalign.nl) was used to automatically extract all mass signals > 3 times the local noise, and to align the extracted mass peaks across all samples, for both GC-TOF MS and GC-PTR-MS data sets. The data sets were subsequently filtered for redundant signals from the same compound, using the MMSR script (*14*) and a similar program developed in Labview software environment.

Cluster analysis and correlation maps were obtained to visualize correlations between sensory data and volatile compounds. The elimination of redundant information of the data matrix increases the quality of the obtained correlation maps when two data sets originating from two different instruments are fused. The correlation maps are created by plotting all compounds in 2D while keeping the distances between them intact. In a next step, correlation coefficients (Pearson product moment correlation) are calculated between all compounds. Only compounds that showed a significant correlation (corrected for the false discovery rate) were selected, and those that had a coefficient of at least 0.8 were connected by a line. In this way, a network was obtained showing which compounds are closely correlated to each other, because they are connected through a series of lines. This type of analysis can detect groups of functional related compounds and show the interrelatedness.

Olfactometer Method. Study on the validation of the causal relationship between single compound concentration in the headspace of coffee and the sensory perception of individual attributes was performed using an olfactometer (Burghart olfactometer OM6b; Heinrich Burghart Elektro and Feinmechanik GmbH, Wedel, Germany). Two parallel air streams are directed toward the outlet head of the olfactometer (sniffing port). Both air streams are set at the same flow rate, humidity, and temperature. One air stream is made only of pure air, and the other is the diluting air, which can be mixed with odorants filled in 250 mL cell vessels. In total, six cell vessels can be used simultaneously, limiting the experimental design to one coffee and five pure compounds. The switching process takes place just before the outlet of the olfactometer. During the interstimulus interval of time, the air stream containing the odorants is swept to the waste circuit, so that only pure air flows to the nose of the panelists. The vacuum is reversed during odor stimulus, so that air loaded with odorants flows toward the panelist's nose. The total flow was constantly adjusted at 6 L/min, kept at a temperature of 40 °C and moisture between 50 and 60%.

RESULTS AND DISCUSSION

Sample Selection. Coffee sample selection is an essential step ensuring a rich diversity in perceived sensory profiles for successful application of chemometric methods to relate the perceived quality differences to chemical fingerprints. The focus in this study was set on the selection of coffee samples being different on the basis of their genetics, knowing that postharvest and processing parameters play very important roles as well. One approach consists of collecting coffee samples of different origins from different coffee-growing regions all over the world. Besides

differences related to the coffee type/group, this includes variability in soil and growing and harvesting conditions. Another option is to collect seedlings of different coffee varieties growing in different coffee regions all over the world and to grow them at the same place in very well-defined agricultural conditions. Postharvest conditions (dry processed) are easier to standardize in this case, and the remaining quality-determining factors are coffee type/group and roasting parameters. The latter setup was used in the present study using *C. canephora* samples produced at a farm in Ecuador under fully reproducible conditions. Each individual sample was collected on 45 genetically similar (clone) trees.

Sample Treatment. Roasting tests showed that the individual roasting efficiency of the samples was different depending on the different bean size, water content, sugar content, and many more parameters. When using the same roasting parameters for all coffee samples (same heat load set by roasting temperature and duration of roasting), different roasting colors were obtained (measured as CTn value) for each sample. There is no "gold standard" to overcome this problem. Sensory evaluation should show the differences between the coffees but should not be influenced by the different roasting degrees depending on the bean characteristics. Therefore, the individual roasting conditions for each coffee sample had to be determined to avoid effects such as (i) perception of strong green and acid notes resulting from a low roasting degree being insufficient for the full development of the coffee aroma or (ii) perception of strong bitter, roasted, and burnt notes due to over-roasting with the result of inhibiting the perception of subtle differences in caramel, woody, earthy, and rubbery notes. Each individual coffee sample was roasted to five different roasting degrees covering the range from low, to medium, to high, to city, and to Italian roast.

Sensory Evaluation. The appropriate roasting conditions were investigated in technical tasting, using a panel of nine coffee experts. Panelists selected the most appropriate roasting degree by evaluating the sensory attributes bitter, acid, roasty, and rubbery, regarding mainly an acid/bitter balance. Additionally, the presence of off-notes among attributes such as burnt, green, earthy/woody, fermented, phenol/medical, and cereal was identified.

The monadic sensory evaluation of the samples roasted to the selected individual roasting degree showed differences between the samples when rating the attributes rubbery, body, bitter, roasted, coffee odor, coffee flavor, chemical, and fermented (**Table 1**). In some cases, a strong fermented note was perceived, already evaluated as off-note in the technical sensory evaluation. High, medium-high, and medium fermented off-note intensities were present in three, two, and three different samples, respectively.

As all samples were processed under similar conditions (e.g., careful selection of cherries, sun-drying), it can be assumed that the fermented taste is related to characteristics of the clones. One hypothesis is that some clones have cherries which are more susceptible to fermentation during sun-drying. Strong differences exist between clones in the percentage of pulp in the cherries and their sweetness (sugar content). Therefore, fermentation may be promoted in clones having cherries with high percentage of pulp and/or are more concentrated in sugars. Besides the fermented note, the samples were very different from each other. The highest differentiation can be seen by the attributes bitter, coffee odor, and body, with range/LSD values beyond 2 (**Table 1**).

Analytical Characterization. To obtain a comprehensive view of compounds present in the headspace of coffee, GC-TOF-MS, GC-PTR-MS, and online PTR-MS analyses were used. Tenax traps in combination with automatic thermo desorption (ATD) on a GC column showed good results in obtaining rich chemical fingerprints with high reproducibility (*18*, *19*). GC-TOF-MS and GC-PTR-MS data were processed by using Metalign software (www.metalign.nl). The software includes baseline correction, local noise calculation, and peak picking respecting a limitation in signal-to-noise ratio and alignment of the detected peaks over all samples by an iterative algorithm, which compensates for local slight shifts in retention time. Due to the ionization induced fragmentation by electron impact in GC-TOF-MS and fragmentation by proton transfer and collision-induced fragmentation in GC-PTR-MS, single compounds are represented by an average of 10 mass signals and 3 mass signals, respectively. To reduce the data volume and eliminate redundant information, a "centrotyping" program was applied, comparable to that previously reported (*14*). This program correlates the intensity profiles of individual mass signals across all samples within a predefined retention time window, which can be adjusted according to the shifts in retention time caused by the limitation of instrumental accuracy. Mass signals that correlate are clustered and expressed as single centrotype because they are expected to belong to one and the same compound. This reduces the data volume without loss of compound information, because the fragmentation pattern of each centrotype is stored. Hence, this information can be used for identification via comparison with spectral databases. The aligned and centrotyped GC-PTR-MS data resulted in 266 centrotypes (compounds) indicated with "PT" followed by a compound number, and GC-TOF-MS data in 247 centrotypes (compounds) were obtained, each indicated with "EI" followed by a compound number. Only peaks with a signal-to-noise ratio of >3 were used in these data analyses. Cluster analysis and a correlation map were obtained to visualize correlations between sensory data, GC-TOF-MS data, and GC-PTR-MS data of volatile compounds. The correlation map helped to identify groups of related compounds and shows their interrelation (**Figure 1**). A correlation of the evaluated fermented intensities with attributes such as coffee odor and coffee flavor can be observed because positive and negative correlations are clustered together in the correlation map. Coffees that are perceived with fermented off-note are typically not described as intense in coffee

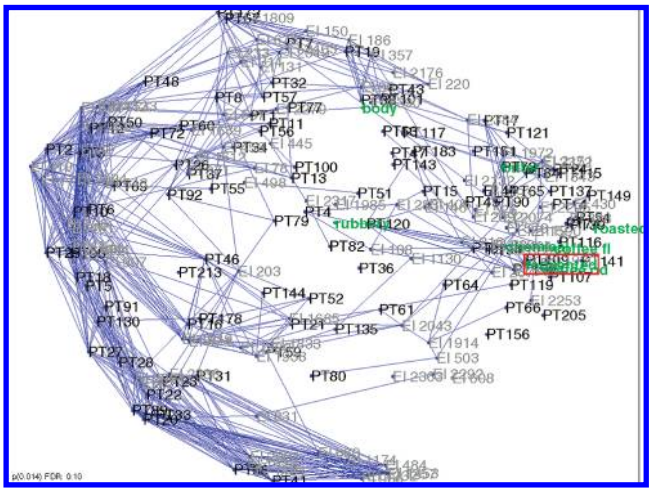

Figure 1. 2D correlation (Pearson product moment correlation) map of GC-TOF-MS data (centrotypes are shown in gray letters indicated with "EI" followed by a centrotype number), GC-PTR-MS data (centrotypes are shown in black letters indicated with "PT" followed by a centrotype number), and sensory data (in green letters). The sensory attribute "fermented" is indicated with a red box. Correlations >0.8 are indicated by a blue connection line.

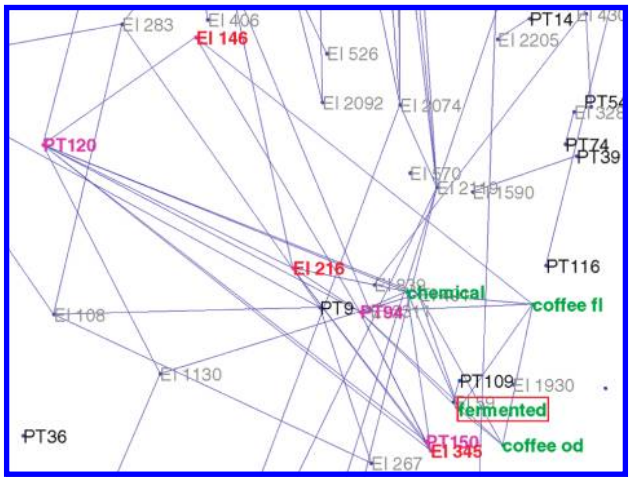

Figure 2. 2D correlation (Pearson product moment correlation) map as shown in **Figure 1** but zoomed into the cluster of fermented, chemical, coffee odor, and coffee flavor attributes. Centrotypes highly correlated with evaluated fermented intensities are indicated with magenta letters in the case of GC-PTR-MS data and with red letters in the case of GC-TOF-MS data.

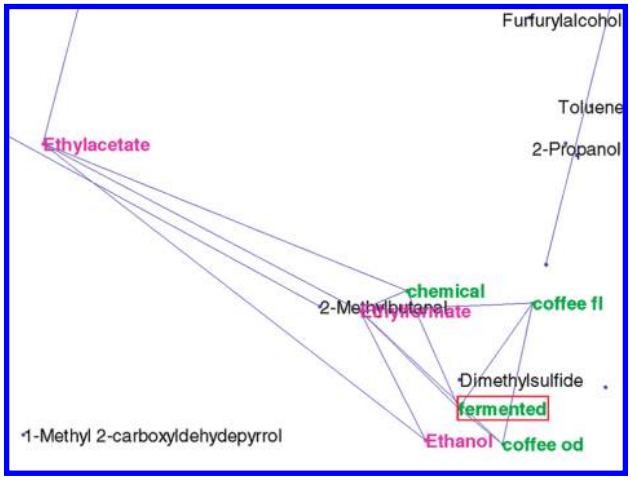

Figure 3. 2D correlation (Pearson product moment correlation) map as shown in **Figure 2** but with the annotation by identified compound name. Only GC-PTR-MS data are shown for better readability.

odor and flavor. Therefore, these attribute intensities anticorrelate with high coefficient. Furthermore, the sensory attribute chemical is often associated with the fermented off-note and therefore shows a high positive correlation.

A good correlation could be observed for the cluster of fermented, chemical, coffee odor, and coffee flavor attributes with several compounds in each GC-TOF-MS and GC-PTR-MS data set (**Figure 2**). Several compounds could be identified by comparison with mass spectral databases and retention indices. These included well-known coffee aroma compounds such as 2-methylbutanal, 3-methylbutanal, 2-methylpropanal, and dimethyl sulfide shown as PT9/EI 267, PT15/EI283, PT4/EI108, and PT109/EI59, respectively, in the correlation maps (**Figures 1** and **2**). Further compounds could be identified as ethyl acetate (PT120/EI216) and ethanol (PT150/EI345), with a positive correlation with the fermented off-note (**Figures 2** and **3**). Both GC-PTR-MS and GC-TOF-MS data indicated that ethyl acetate and ethanol were increased by factors of 10 and 8, respectively, in coffees perceived as fermented. Compound PT134

Table 2. Correlation Matrix of Compounds Detected by GC-PTR-MS Correlating with the Perceived Fermented Off-note

	ethyl formate PT94	ethyl acetate PT120	ethanol PT150	fermented	chemical	coffee odor
ethyl formate PT94	1.00	0.95	0.97	0.79	0.78	−0.77
ethyl acetate PT120	0.95	1.00	0.89	0.77	0.84	−0.75
ethanol PT150	0.97	0.89	1.00	0.69	0.71	−0.66
fermented	0.79	0.77	0.69	1.00	0.86	−0.90
chemical	0.78	0.84	0.71	0.86	1.00	−0.78
coffee odor	−0.77	−0.75	−0.66	−0.90	−0.78	1.00

Table 3. Correlation Matrix of Compounds Detected by GC-TOF-MS Correlating with the Perceived Fermented Off-note

	ethyl formate EI146	ethyl acetate EI216	ethanol EI345	fermented	chemical	coffee odor
ethyl formate EI146	1.00	0.93	0.94	0.77	0.73	−0.71
ethyl acetate EI216	0.93	1.00	0.86	0.81	0.86	−0.78
ethanol EI345	0.94	0.86	1.00	0.68	0.69	−0.66
fermented	0.77	0.81	0.68	1.00	0.86	−0.90
chemical	0.73	0.86	0.69	0.86	1.00	−0.78
coffee odor	−0.71	−0.78	−0.66	−0.90	−0.78	1.00

in the GC-PTR-MS data set showed a positive correlation and an increase by a factor of 3. 2-Methyl-1-butanol and 3-methyl-1-butanol were identified. Both compounds were coeluted when using the described GC method. Due to this coelution, they do not appear in the correlation maps (**Figures 1**−**3**).

Identification of Ethyl Formate. An increase in concentration by a factor 20 was observed in the case of a distinct compound (PT94/EI146) when the raw data of GC-TOF-MS were analyzed and GC-PTR-MS simultaneously recorded. Due to coelution with a second compound, the centrotypes include only one unique ion signal at m/z 47 in the case of GC-PTR-MS data and m/z 31, 47, 56, and 73 ion signals in the case of GC-TOF-MS data. The coeluting compound was identified as methyl acetate (PT7/EI150) with a GC-PTR-MS spectrum of m/z 75, 43, 76, 61, and 77 with decreasing ion intensities. A small shoulder on the methyl acetate peak at m/z 75 (GC-PTR-MS) at a slightly earlier retention time equal to the retention time of the compound PT94 with m/z 47 (GC-PTR-MS) was observed. In combination with the centrotype spectrum obtained by GC-TOF-MS at the same retention time (EI146), the second compound was tentatively identified as ethyl formate. Pure and mixed injection of methyl acetate and ethyl formate validated the identification of both compounds by comparison of the mass spectrum from both detectors (TOF-MS and PTR-MS; PTR-MS unique mass for ethyl formate m/z 47) and the expected retention indices. As a result, ethyl formate was found as a new marker for coffees perceived as fermented. **Figure 3** shows the zoomed correlation map including identified compounds. **Tables 2** and **3** show the correlation matrix of the mentioned compounds correlating with the fermented off-note. Until now, ethyl formate has not been reported in the literature to be correlated with the fermented defect of coffee samples.

Quantification of Ethyl Formate. Partition coefficient measurements in water solution and coffee solution showed no significant difference for compounds such as 2-methylpropanal and 3-methylbutanal (*20*). Even though this might not be valid for all volatile organic compounds, we assumed that using the solution of pure ethyl formate in water should be comparable with a solution in coffee. Therefore, we compared concentrations in the headspace of pure ethyl formate in water with that in coffee.

A concentration of 280 ± 10 μg/L of ethyl formate was quantified in coffee extracts with the highest concentration. Coffees with ethyl formate concentration of > 60 μg/L were perceived with a fermented off-note. Coffees with ethyl formate concentration of < 20 μg/L were described as not fermented.

Validation Trials. Therefore, three additional samples (A−C) were selected by an independent sensory panel of four coffee experts evaluating only the fermented attribute of the coffee brews (coffees were roasted to CTn 100). One sample was perceived with high fermented intensity (C), one medium-high (B), and one medium-intense (A). The results were compared with the intensity differences of ethyl formate in the headspace of each of the coffees and, indeed, validated the findings. Studying the variation in intensity of ethyl formate depending on the roasting degree, the coffee with the highest intensity in the fermented attribute was analyzed at five different roasting degrees, CTn 109 (low roasted), 95, 90, 79, and 71 (Italian roast) and CTn 90 with 10 replicates. Using the described GC methods, the results showed no significant change in concentration of ethyl formate depending on the five different roasting degrees (analytical repeatability of 6% and sample reproducibility of 6%).

Online PTR-MS time−intensity profiles of coffee headspace as described in ref *18* were measured for all coffee samples. A strong increase in the maximum intensity of the ion signal at m/z 47 was observed in case of coffees perceived as fermented. By analyzing the headspace trapped on a Tenax cartridge by GC-PTR-MS, four compounds were identified contributing to the ion signal at m/z 47. The average contribution of the compounds to m/z 47 was measured for three coffees with the highest perceived fermented intensity. A fragment of ethyl formate (56%), the protonated molecular ion of ethanol (18%), the protonated molecular ion of formic acid (24%), and an isotope of acetaldehyde (2%) were identified. Ethanol and ethyl formate concentrations in the headspace of coffee are highly correlated with the perceived fermented intensity, whereas the formic acid concentration is not. The major contribution of ethanol and ethyl formate to the ion trace at m/z 47 makes the online PTR-MS measurements suitable for detecting this off-note in R&G coffee independent of the roasting degree in the range from low roasting to Italian roasting degree.

The causality between ethyl formate concentration and the perceived off-note was tested by using an olfactometer. The extract of a coffee without a fermented off-note was placed in one cell vessel and pure ethyl formate diluted in water in a second cell vessel. The volatiles were purged to the sniffing port at various dilutions to obtain the same ethyl formate concentration as obtained from a coffee extract of coffee with fermented off-note (highest concentration = 1 ppmv or 3.3 mg/m^3). Four panelists were asked to evaluate the difference in the perceived coffee profile while the concentration of ethyl formate was changed. No conclusive change in the overall profile was perceived. When pure ethyl formate was sniffed at the same concentration as present in the coffee with fermented off-note, only a slight intense smoky odor was perceived by the panelists. By the same approach but with ethyl formate replaced with 3-methyl-1-butanol, the overall profile changed and the fermented off-note was perceived as expected from the results obtained when the pure compound was sniffed individually (described as chemical and fermented). However, the overall perception was still different compared to the profile of a coffee with the fermented off note, indicating that probably a misbalance of several compounds leads to the perceived off-note.

Formation. Ethyl formate is most likely formed by esterification of formic acid with ethanol. Formic acid is a well-known sugar degradation product, along with acetic acid, formed in the course of the Maillard reaction cascade. Its high abundance has been shown in model studies (*21*) and, thus, it is also expected to be generated upon coffee roasting. Ethanol, on the other hand, is a well-known fermentation product that can lead to various ethyl esters during coffee processing. Indeed, several of these ethyl esters have been reported by Bade-Wegener et al. (*9*) in fermented coffee samples. It seems that in ethyl formate formation, the ethanol concentration is the limiting factor, not formic acid, when the correlation data obtained in this study are considered.

Application. The untargeted chemometric approach allowed detection of a positive correlation between ethyl formate and the perceived fermented off-note in coffee and showed that ethyl formate can be used as a new marker for this important quality trait. However, it could be shown by olfactometer tests that this is not a causal effect. The marker can easily be detected by online PTR-MS or GC-PTR-MS measurements of R&G coffee headspace, whereby the coffee can be roasted over a wide range of roasting degrees (low roasted to Italian roasted) without influencing the result. Ethyl formate should be seen as an easily detectable quality marker correlating with the fermented off-note.

LITERATURE CITED

(1) Viani, R. Quality of coffee. *Riv. Ital. EPPOS* **1993**, *4*, 55–80.

(2) Selmar, D.; Bytof, G.; Knopp, S. E.; Bradbury, A.; Wilkens, J.; Becker, R. Biochemical insights into coffee processing: quality and nature of green coffees are interconnected with an active seed metabolism. *Colloque Sci. Int. Cafe* **2004**, *20*, 111–119.

(3) Selmar, D.; Bytof, G. Green coffee is alive! A review on the metabolic processes taking place in coffee beans during processing and their implication for modern coffee research. *Colloque Sci. Int. Cafe* **2006**, *21*, 423–433.

(4) Bytof, G.; Selmar, D.; Schieberle, P. New aspects of coffee processing. How do the different post harvest treatments influence the formation of potential flavour precursors? *J. Appl. Bot.* **2000**, *74*, 131–136.

(5) Grosch, W. Warum riecht Kaffee so gut? *Chem.−Ztg.* **1996**, *30*, 126–33.

(6) Vitzthum, O. G. Thirty years of coffee chemistry research. In *Flavor Chemistry: Thirty Years of Progress*; Teranishi, R., Wick, E. L., Hornstein, I., Eds.; Kluwer Academic: New York, 1999; pp 117−133.

(7) Grosch, W. Chemistry III: Volatile compounds. In *Coffee: Recent Developments*; Clarke, R. J., Vitzthum, O. G., Eds.; Blackwell Science: London, U.K., 2001; pp 68−89.

(8) Shankaranarayana, M. L.; Abraham, K. O. Evaluation of coffee quality using chemical and instrumental methods. *J. Coffee Res.* **1986**, *16*, 14–22.

(9) Bade-Wegner, H.; Bendig, I.; Holscher, W.; Wollmann, R. Volatile compounds associated with the over-fermented flavor defect. *Colloque Sci. Int. Cafe* **1997**, *17th*, 176–182.

(10) Scheidig, C.; Czerny, M.; Schieberle, P. Changes in key odorants of raw coffee beans during storage under defined conditions. *J. Agric. Food Chem.* **2007**, *55*, 5768–5775.

(11) Tavares, L. A.; Ferreira, A. G.; Ferreira, M. M. C.; Correa, A.; Mattoso, L. H. Analysis of blends of raw coffees of Arabica and Robusta varieties by ^{1}H NMR and chemometric methods. *Magn. Reson. Food Sci.* **2005**, *299*, 80–88.

(12) Alonso-Salces, R. M.; Serra, F.; Reniero, F.; Heberger, K. Botanical and geographical characterization of green coffee (*Coffea arabica* and *Coffea canephora*): chemometric evaluation of phenolic and methylxanthine contents. *J. Agric. Food Chem.* **2009**, *57*, 4224−4235.

(13) Lindinger, C.; Labbe, D.; Pollien, P.; Rytz, A.; Juillerat, M. A.; Yeretzian, C.; Blank, I. When machine tastes coffee: instrumental approach to predict the sensory profile of espresso coffee. *Anal. Chem.* **2008**, *80*, 1574–1581.

(14) Tikunov, Y.; Lommen, A.; de Vos, C. H. R.; Verhoeven, H. A.; Bino, R. J.; Hall, R. D.; Bovy, A. G. A novel approach for nontargeted

data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles. *Plant Physiol*. **2005**, *139*, 1125–1137.
(15) Zlatkis, A.; Sivetz, M. Analysis of coffee volatiles by gas chromatography. *J. Food Sci*. **1960**, *25*, 395–398.
(16) Lindinger, W.; Hansel, A.; Jordan, A. On-line monitoring of volatile organic compounds at pptv levels by means of proton-transfer-reaction mass spectrometry (PTR-MS): medical applications, food control and environmental research. *Int. J. Mass Spectrom. Ion Processes* **1998**, *173*, 191–241.
(17) Labbe, D.; Rytz, A.; Morgenegg, C.; Ali, S.; Martin, N. Subthreshold olfactory stimulation can enhance sweetness. *Chem. Senses* **2007**, *32*, 205–214.
(18) Lindinger, C.; Pollien, P.; Ali, S.; Yeretzian, C.; Blank, I.; Maerk, T. Unambiguous identification of volatile organic compounds by proton-transfer reaction mass spectrometry coupled with GC/MS. *Anal. Chem*. **2005**, *77*, 4117–4124.
(19) Pollien, P.; Lindinger, C.; Yeretzian, C.; Blank, I. Proton-transfer reaction mass spectrometry, a tool for on-line monitoring of acrylamide in food and Maillard systems. *Anal. Chem*. **2003**, *75*, 5488–5494.
(20) Pollien, P.; Yeretzian, C. Liquid−air partition coefficients of coffee flavour compounds: a novel approach using proton-transfer-reaction mass-spectrometry. *Conference Proceedings: ASIC-19eme Colloque Scientifique Int. sur le Café*, 2002.
(21) Davidek, T.; Devaud, S.; Robert, F.; Blank, I. Sugar fragmentation in the Maillard reaction cascade: isotope labeling studies on the formation of acetic acid by a hydrolytic β-dicarbonyl cleavage mechanism. *J. Agric. Food Chem*. **2006**, *54*, 6667–6676.

Received for review May 19, 2009. Revised manuscript received July 30, 2009. Accepted September 10, 2009.