Volatiles profiling from high quality cocoa samples at early stages of technological treatment by two-dimensional comprehensive gas chromatography - mass spectrometry

Volatiles profiling from high quality cocoa samples at early stages of technological treatment by two-dimensional comprehensive gas chromatography – mass spectrometry

Federico Magagna[1], Erica Liberto[1], Stephen E. Reichenbach[2], Qingping Tao[3] , Carlo Bicchi[1] and Chiara Cordero[1]

[1] Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino, Via Pietro Giuria 9, I-10125 Torino, Italy
[2] Computer Science and Engineering Department, University of Nebraska, 1400 R Street, Lincoln, NE, USA
[3] GC Image LCC, Lincoln NE, USA

Aim and Scope

Cocoa derives from the cocoa tree (*Theobroma Cacao L.*) belonging to the *Malvaceae* family and it is the main ingredient in the chocolate manufacturing. Cocoa is consumed worldwide and its value and quality are related to unique and complex flavour. The sensory quality of cocoa (aroma, taste, texture) is undoubtedly a key-factor in the production of a premium quality product and it can basically affect consumer preferences [1]. Focusing on aroma, about 600 various compounds (alcohols, carboxylic acids, aldehydes, ketones, esters, and pyrazines) [2] have been identified and most of them are also odour-active volatiles. The peculiar cocoa flavour arises from complex biochemical and chemical reactions during the postharvest processing of raw beans and, in particular, fermentation and roasting are considered the key steps in the formation of the characteristic cocoa aroma.

This study investigates the characteristic distribution of technologically informative and sensory-active volatiles included in the unique profile of "high quality" selections of cocoa (*Theobroma Cacao L.*) from different origins.

In particular, it proposes an investigation strategy capable to fully exploit the potential of GC×GC-MS combined with HS-SPME and automated fingerprinting approach, in defining an informative chemical signature of cocoa volatiles [3-4] and explains how effective and automated data elaboration might improve food quality evaluation and authentication process.

[1] Aprotosoaie A., Vlad L. S. and Miron A., Comprehensive Reviews in Food Science and Food Safety (2015)
[2] Granvogl M, Beksan E, Schieberle P. J Agric Food Chem. 2012 Jun 27;60(25):6312-22
[3] Cordero C, Kiefl J, Schieberle P, Reichenbach SE, Bicchi C. Anal Bioanal Chem. 2015 Jan;407(1):169-91
[4] Reichenbach SE, Tian X, Boateng AA, Mullen CA, Cordero C, Tao Q. Anal Chem. 2013 May 21;85(10):4974-81

Experimental

Cocoa samples

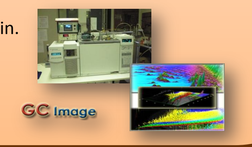
Cocoa samples of different botanical origin (***Trinitario***, ***Forastero***, ***Criollo***), harvested in different countries (**Mexico**, **Ecuador**, **Venezuela**, **Colombia**, **Java**, **Trinidad**, **Sao Tomè**) and at different technological stages (**raw**, **roasted**, **steamed**, **nibs** and **mass**) were analyzed. Each cocoa origin was analyzed in three analytical replicates.

HS-SPME conditions

Sampling: 1.50 g of cocoa powder
Temperature: 50°C
Time: 40 min with the pre-loading of the internal standard (α/β thujone)
Sampling vial volume: 20 mL
Fiber: DVB/CAR/PDMS; 50/30 µm 2 cm Supelco Bellefonte

GC(O)×GC-MS platform

Head Space Solid Phase Microextraction (HS-SPME) was run on a System **MPS-2 multipurpose sampler** (Gerstel, Mülheim a/d Ruhr, Germany).
Agilent 6890-5975C MS: operating in EI mode (70 eV) Scan range 35/250 m/z, Scan rate 12,500 amu/s, 28 Hz
Zoex KT2004 loop modulator: Cryogenic - liquid nitrogen; Modulation period: 3 s; Massflow ctr. Optimode™ by SRA Italy
GC-O sniffing port: designed by Grosh et al at. was installed at the secondary FID outlet and connected by a spliiter device before ^{2}D column . **GC×GC-MS :** injector temperature: 250°C, injection mode: split, ratio: 1/10; carrier gas: helium, initial head pressure 296 kPa. **Column set:** ^{1}D SOLGELWAX (30 m, 0.25 mm ID, 0.25 um d_f) and ^{2}D OV1701 (2 m, 0.1 mm ID, 0.1 um d_f) (Mega, Legnano (Milan), Italy); loop capillary 1 m, 0.1 mm ID, 0.1 um d_f
GC oven : 40°C (1 min) to 240°C (10 min) at 3.0°C/min.
Data Acquisition: Agilent Chem Station vers D.02
Data elaboration : GC-Image ver 2.5.

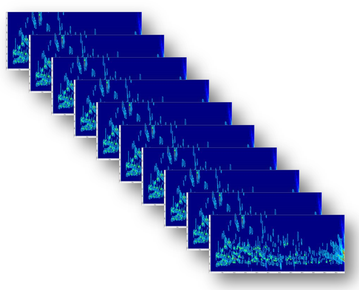


1. *Peak-region feature* fingerprinting

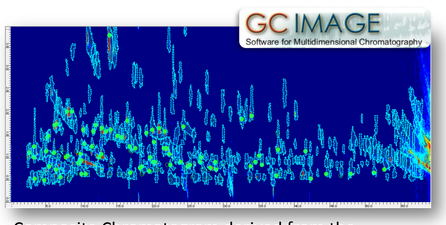
Untargeted analysis was performed to extend the comparative process to the entire pattern of detected volatiles from the headspace of cocoa samples. The unsupervised fingerprinting was based on the ***"peak-region feature"*** approach and implemented by Image Investigator® in the GC Image® software. This data elaboration step was made more informative by considering targeted 2D peaks, included in a ***targeted template.*** The strategy enables to preserve all information about known analytes within the fingerprinting.

The fully automated procedure of ***peak-regions fingerprinting*** delineates a small 2D retention-times window (or *region*) per 2D *peak* over the chromatographic plane and it approaches the "one-feature-to-one-analyte" selectivity typical of peak features methods, with all the advantages of regional features matching

These advantages include unambiguous cross-detection/matching of trace peaks that may be detected in some samples but not in others and co-eluting analytes that may be resolved in some chromatograms but not in others.



Briefly the unsupervised procedure:
(1) detects and records 2D peaks in individual chromatograms;
(2) locate *registration* peaks, i.e. peaks that reliably match across all chromatograms;
(3) aligns and combines the sample chromatograms to create a ***composite chromatogram;***
(4) defines a pattern of *region features* from the peaks detected in the cumulative chromatogram.
(5) create a combined **targeted** and **untargeted** template from the registration peaks from step 2, the peak-regions from Step 4, and the targeted peaks.
(*6*) the feature vector (*peak-region reliable template*) is used for cross-sample analysis.



Composite Chromatogram derived from the automated processing of 105 cocoa analyses

Once the resulting template is matched to a target chromatogram, the analysis includes peak-regions (light blue graphics), targeted peaks (green circles) and registration peaks (red circles). Feature regions are aligned to corresponding peaks, and the characteristics of those features including all metadata (**retention times in both chromatographic dimensions**, **detector response**, **relative/absolute intensity**, **peaks' EI-MS fragmentation pattern**, **response factors**, etc.) are computed to create the *peak-region reliable template* for the target chromatogram to be adopted for comparative purposes within cocoa samples.
The final output is a data matrix where *peak-regions* and template peaks are cross-aligned within all samples' chromatograms and the response data are available for further chemometrics approaches.

2. Cocoa volatiles profiling/fingerprinting

The resulting 2D Peak Volumes derived from GC×GC analysis, referred to the 450 **untargeted peak features (Untargeted profiling)** detected from the headspace of different cocoa samples, are visualized as **Heat map** in **Figure 1.** Columns follow the Linear Retention Indices (from left to right) ordered coherently with the polar x apolar column combination. Peak Volumes were normalized by dividing by row standard deviation.

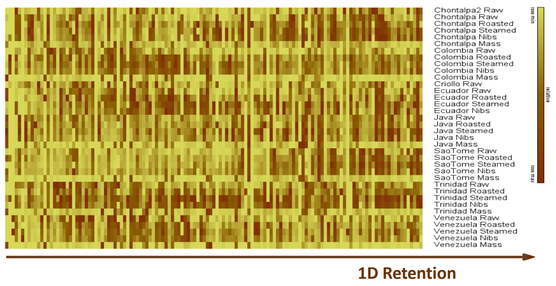
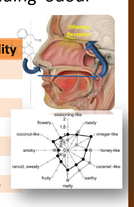


Figure 1: Heat map obtained on data set of all cocoa samples.

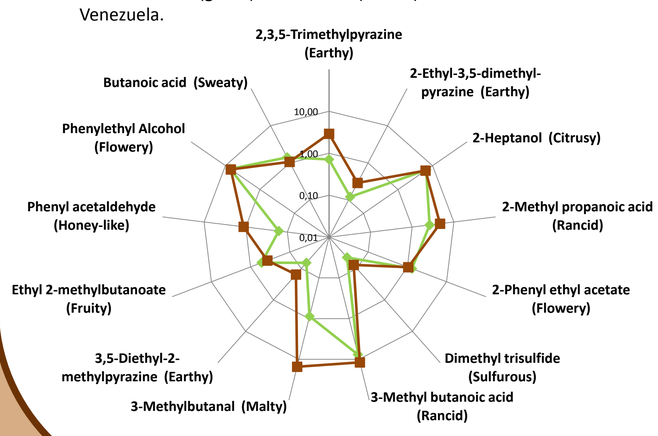
As illustrated by coloured spots distribution, cocoa of different geographical/botanical origin and analyzed at different technological stages show significant differences in the volatile fraction and also in the distribution of **sensorially-active compounds** (***Key-aroma compounds***), especially after the roasting process.
The Table below shows the most potent odour compounds of raw/roasted Cocoa each one associated with the corresponding odour descriptor [1]

Compound Name	Odour quality	Compound Name	Odour quality
2,3,5-Trimethylpyrazine	Earthy	3-Methylbutanal	Malty
2-Ethyl-3,5-dimethyl-pyrazine	Earthy	Ethyl 2-methylbutanoate	Fruity
3,5-Diethyl-2-methylpyrazine	Earthy	2-Heptanol	Citrusy
2-Methyl propanoic acid	Rancid	Dimethyl trisulfide	Sulfurous
3-Methyl butanoic acid	Rancid	2-Phenyl ethyl acetate	Flowery
Acetic acid	Sour	Phenylethyl Alcohol	Flowery
Butanoic acid	Sweaty	Phenyl acetaldehyde	Honey-like



[1] Frauendorfer F., Schieberle P., (2008) J. Agric. Food Chem., 56, 10244-10251

Figure 2: Spider diagram with the distribution of most potent odorant in raw (green) and roasted (brown) cocoa from Venezuela.



Raw cocoa (**Figure 2**, green line) shows a peculiar qualitative distribution of *key-aroma compounds* (reported as relative abundances) that entails a specific sensory profile and, as a consequence, a different sensory impact.
Among the manufacturing steps, **fermentation** and **roasting** are the most important for flavour formation.
During **fermentation** of raw beans, non-volatile aroma precursors, such as free amino acids, peptides, reducing sugars, are generated while some volatiles (alcohols, esters, aldehydes, organic acids) are formed by microorganisms .
Thermal treatment during **roasting** (brown line) leads to an increase of specific (technological) markers: alkyl pyrazines, Strecker aldehydes (*3-Methyl butanal, phenyl acetaldehyde*) and acids (*3-Methyl butanoic acid* - rancid).

Results and Discussion

3. Volatiles distribution across samples of different origin and technological stages

An unsupervised multivariate approach (**Principal Component Analysis** - PCA), was adopted to map the natural conformation of samples' groups and sub-groups and to localize informative chemicals responsible of samples differentiation – fingerprinting.
The PCA was carried out on 180 target analytes identified within the volatile fraction of cocoa samples of seven different origins and at five different technological stages (**Figure 3**); a total of 105 GC×GC runs (three analytical replicates for each sample) .

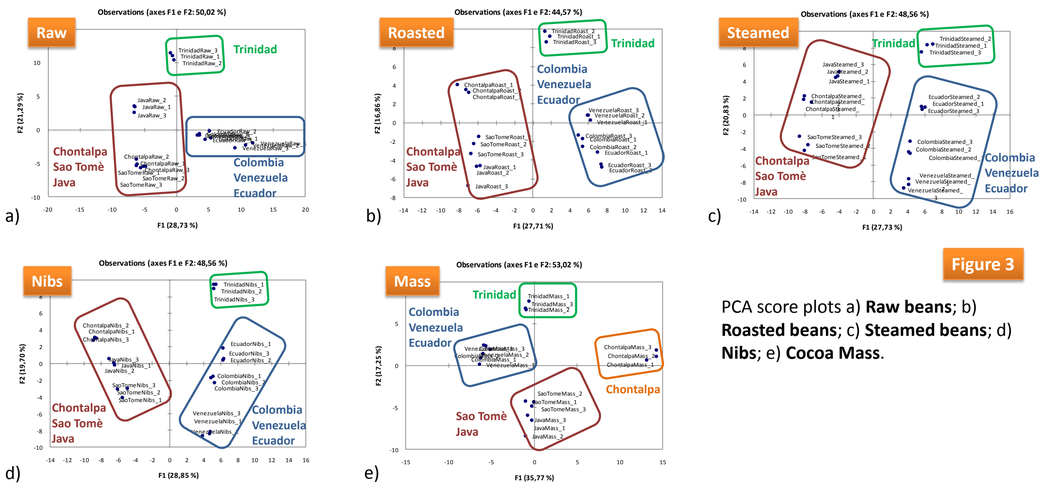


Figure 3

PCA score plots a) **Raw beans**; b) **Roasted beans**; c) **Steamed beans**; d) **Nibs**; e) **Cocoa Mass**.

PCA results show coherent clustering for cocoa samples: geographical and botanical origin dominate this group conformation while the effect of technological trasformation is more evident along manufacturing process. In general, cocoa from **Ecuador**, **Venezuela** and **Colombia** (all *Trinitario*) are clusterized together at all stages also in accordance with their sensory profiles.
Cocoa from **Trinidad** has a distinctive fingerprint that enables its independent clustering at all stages; **Chontalpa** and **Java** origins, despite their different geographical provenience, show similar chemical fingerprint.
Cocoa samples clustering is described by different variables: for example **Chontalpa** and **Sao Tomè** raw beans resulted higher in *Acetic acid* (Sour), *phenyl acetaldehyde* (honey-like) and *esters* while cocoa from South America is characterized by a pool of short chain primary alcohols (*1-Butanol*, *1-Pentanol*, *1-Hexanol*). In addition, roasted samples of **Chontalpa**, **Java** and **Sao Tomè** have a distinctive and intense fingerprint of *pyrazines.* South America cocoa is connoted by higher amounts of aromatic *ketones* such as *1-Hydroxy-2-propanone*, *2,3-Pentanedione*, *2,3-Butanedione.*

Figure 4 (PCA scores plot) shows the **influence of the processing stage** on the sub-classification of the samples (**Chontalpa** example).

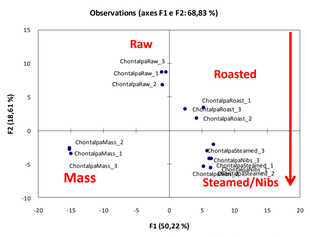


Figure 4: PCA score plots of Chontalpa (Mexico) cocoa .

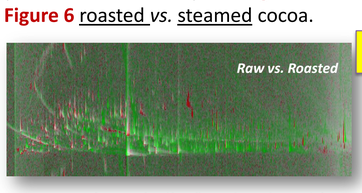
PCA carried out on **Chontalpa** data set shows a sub-classification of the stages: raw, roasted, steamed/nibs and mass. **Steamed beans** and **nibs** have a similar fingerprint of volatiles; this data is also confirmed by others cocoa origins with the only exception of **Colombia** (data not shown).

Interestingly, from raw to steamed beans, there is a costant increase of some technological markers such as *alkyl pyrazines*, some ketones (*2,3-pentandione*, *2,3-octandione*), *furfural, furfuryl alcohol, dimethyl trisulfide* etc (red arrow).

The cocoa **mass** sample shows a less informative chemical profile due to the matrix effect that influence the release of volatile compounds. By sampling volatiles in the same conditions as for the other stages this behaviour is magnified and opens interesting future perspectives in terms of release kinetics.

The ***visual* features fingerprinting** named **image comparison** allows to immediately reveal compositional differences in the chromatographic patterns within samples. **Figure 5** shows the comparative visualization of fermented *vs.* roasted beans (Chontalpa) while **Figure 6** roasted *vs.* steamed cocoa.

Figure 5: Raw vs Roasted Chontalpa beans.



Chontalpa

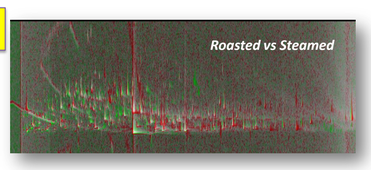


Figure 6: Roasted vs Steamed Chontalpa beans.

The effect of roasting on the volatile fraction is evident: Green colorized regions indicate those analytes that are more abundant in the roasted sample, in particular *alkyl pyrazines*, *strecker aldehydes* and *methyl ketones*.
Red colorized regions indicate compounds more abundant in the unroasted sample; these volatiles are formed during fermentation and includes linear aldehydes (*hexanal*, *decanal*), *Acetic acid*, *esters* and *short chain alcohols*.

It is interesting to monitor the effect of steaming on the volatile fraction: green colorized regions indicate those analytes that are more abundant in the steamed sample. Some markers, diagnostic of thermal treatment, show a costant increase, in particular *alkyl pyrazines*: **this result is important for the sensory perception because an increase of these markers (Earthy note) could lead to a product with undesidered sensory property.**
Red colorized regions indicate compounds more abundant in the roasted sample, especially *2-Methyl ketones*.

Conclusions

The volatile fraction of cocoa of different geographical/botanical origin and at different technological stages has been characterized showing its evolution along manufacturing processing (fermentation, roasting, steaming). This multidimensional approach (sampling-GC×GC-MS-data mining) enables an effective and objective chemical characterization of the product including data on sensory profile. The results here shown demonstrate that the combination of the separation power of GC×GC-MS with automated data elaboration can be successful in defining an informative volatiles chemical signature for a valuable product as cocoa, which sensory profile is fundamental for high quality food products and economical value.

Acknowledgments