Tutor/s

Dr. Oscar Núñez Burcio
Departament d'Enginyeria Química i Química Analítica

Dr. Javier Saurina Purroy
Departament d'Enginyeria Química i Química Analítica

Grau de Química

Treball Final de Grau

HPLC-UV and HPLC-FLD Fingerprinting for the Detection and Quantitation of Adulterations in the Prevention of Coffee Frauds

Detecció i Quantificació d'Adulteracions en la Prevenció de Fraus en Cafè mitjançant empremtes HPLC-UV i HPLC-FLD

Josep Pons Marquès

January, 2021

Universitat de Barcelona

B:KC Barcelona Knowledge Campus
Campus d'Excel·lència Internacional

Imagination will often carry us to worlds that never were, but without it we go nowhere.

Carl Sagan

Voldria agrair primer de tot, a sa meva família per haver-me ajudat a complir es meus somnis i haver estat amb jo sempre que ho he necessitat.

Agrair també als meus tutors, Oscar i Xavi, per tot el suport, ajuda i paciència que han tingut aquets mesos. També, als companys del CECEM, especialment al Luis i a la Nerea per haver-me fet companyia durant les hores de laboratori i per tot el suport que m'han donat aquets mesos ajudant-me a confiar en mi i en aquet projecte.

Finalment, voldria agrair als meus companys de batalles universitàries Alex, Edu, Javi i Pau per haver fet d'aquets anys una experiència inoblidable, als meus amics ferreriencs i Helena per tot el viscut i, per últim, voldria agrair a n'Ekiñe per haver-me donat suport, comprensió i ànims durant tot aquests mesos.

Gràcies de tot cor.

REPORT

CONTENTS

1. SUMMARY

Globalization has produced a total change of scenario in food industry producing a tough competence to occupy the market share, instigating the reduction of costs by usage of fraudulent practices derived from food adulteration. These practices are performed by substitution of most valuable components for other with less commercial value and/or lower health beneficial properties supposing an economic fraud and a potential health problem. Coffees are sometimes the target of this kind of fraudulent practices due to the high demand of the product where manufacturers adulterate coffee with wheat, corn, and other grains, seeds and plants.

In this work, simultaneous non-targeted HPLC-UV and HPLC-FLD fingerprinting methods were developed to achieve the classification and authentication of different instant coffee, and chicory samples using multivariate chemometric methodologies such as principal component analysis (PCA), partial least squares-discriminant analysis (PLS-DA) and partial least squares (PLS). Both HPLC-UV and HPLC-FLD fingerprints, proved to be excellent chemical descriptors for the discrimination of chicory samples against instant coffee and decaffeinate coffee by PLS-DA. However, better results were obtained with HPLC-UV fingerprints when coffee was discriminated from decaffeinated coffee (94.4% classification rate respect to 83.3% for HPLC-FLD fingerprints). Besides, both methodologies were able to detect and quantify adulterant levels in coffee and decaffeinated samples adulterated with chicory exhibiting good regression linearity ($R^2 \geq 0.996$), and low calibration (0.7-2.1%) and prediction (2.4-3.5%) errors. Overall, both non-targeted HPLC-UV and HPLC-FLD showed to be effective, simple, and trustable to accomplish the characterization, classification and authentication of instant coffee and chicory samples being potential methodologies to prevent food frauds.

Keywords: High Performance Liquid Chromatography, Fingerprinting, Chemometrics, Food fraud, Coffee, Chicory.

2. RESUM

La globalització ha produït un canvi de paradigma total en la indústria alimentària originant una dura competència en la ocupació del nínxol de mercat, instigant la reducció de costos de producció per mitjà de pràctiques fraudulentes derivades de l'adulteració alimentària. Aquestes pràctiques es realitzen per mitjà de la substitució dels components més valuosos per altres amb menor valor comercial i/o menors beneficis per a la salut suposant un frau econòmic i potencials riscs sanitaris. El cafè és, sovint, objecte d'aquet tipus de pràctiques degut a l'alta demanda del producte que pot comportar l'adulteració de cafè per part dels fabricants amb blat, soja i altres grans, llavors i plantes.

En aquest treball, s'han desenvolupat de forma simultània dos mètodes no dirigits amb perfils d'empremtes per mitjà de la cromatografia de líquids d'alta eficàcia (HPLC-UV i HPLC-FLD) per tal d'aconseguir la classificació i autenticació de diverses mostres de cafè instantani i xicoira utilitzant metodologies quimiomètriques multivariants com l'anàlisi de components principals (PCA), l'anàlisi discriminant amb regressió de mínims quadrats (PLS-DA) i la regressió de mínims quadrats (PLS). Ambdues metodologies no dirigides (HPLC-UV i HPLC-FLD) han proporcionat excel·lents descriptors químics per a la discriminació de mostres de xicoira respecte cafè i cafè descafeïnat usant PLS-DA. Tot i això, al discriminar cafè de cafè descafeïnat s'obtenien millors resultats amb el mètode d'empremta HPLC-UV (amb un 94.4% d'índex de classificació per 83.3% per HPLC-FLD). A més, s'ha comprovat que ambdós mètodes són adequats per a la detecció i quantificació de nivells d'adulterants en mostres de cafè i cafè descafeïnat adulterades amb xicoira exhibint bones regressions lineals ($R^2 \geq 0.996$), i baixos erros de calibratge (0.7-2.1%) i de predicció (2.4-3.5%). En conclusió, ambdós mètodes no dirigits (HPLC-UV i HPLC-FLD) han provat ser efectius, senzills i confiables per aconseguir la caracterització, classificació i autenticació de mostres de cafès instantanis i xicoira convertint-los en potencials mètodes per a la prevenció de fraus alimentaris.

Paraules clau: Cromatografia líquida d'alta eficàcia, Perfil d'empremtes, Quimiometria, Frau alimentari, Cafè, Xicoira.

3. INTRODUCTION

3.1. FOOD FRAUDS

Worldwide massive globalization and urbanization has led to several consequences regarding to alimentary industry producing a total change of focus on the food systems by giving a much larger availability of alimentary products which were totally unbelievable to obtain at certain regions or seasons not far ago. This change of scenario gave the perfect chance to great multinationals with enormous capacities to product, store and distribute to occupy the market niche of traditional small local agents [1]. This alimentary race and the economic interests present on it produces a competition for market share and instigates the reduction of costs to increase benefits offering better prices to customers by usage of various strategies that could end up provoking an alimentary fraud.

The consequence of this alimentary race and competence leads to the necessity to stablish legislation to fight against this new latent problematic: food frauds which includes substitution of aliments, dilution, tampering and mislabeling [2], deriving to serious economic losses and an attacking to the public health (e.g. producing allergy episodes) [3,4].

Increasing concern in population regarding this area leads to the necessity to research on food authentication methods consisting of the confirmation of the stated specifications of the products [5] with traceability registers, labelling tools and emerging analytical methodologies in order to detect or quantify alimentary frauds.

This work will focus on the analysis of coffee and adulterated coffee samples.

3.2. COFFEE

Coffee is one of the most popular beverages around the world. According to the International Coffee Organization (ICO) in 2019, the global coffee consumption has been of 165 million of 60-kg bags [6]. Coffee consists of an infusion of ground roasted beans with characteristic taste and aroma. Coffee plants belongs to Coffea genus from Rubiaceae family plant that contains more than 120 species of coffee being *Coffea Arabica* and *Coffea Canephora* the most important ones.

Arabica coffees, which are the most relevant for consumers due to their better sensorial properties, are cultivated in Colombia, Brazil, Kenya, Ethiopia, and India, principally. On the other hand, Robusta coffee comes mainly from Africa, Indonesia, and Brazil. It presents a lower production cost due its higher resistance to diseases which facilitates its cultivation.

Coffee is known as an appreciated and society-normalized stimulant, property associated mainly to their alkaloids contents, such as caffeine and trigonelline which also contribute to its bitter taste [7].

Apart from alkaloids, recent studies have shown, that coffee is rich in other bioactive compounds like phenolic acids and polyphenols, mainly caffeic and chlorogenic acids[8], giving great antioxidant activity. In fact, some studies have related the coffee intakes with a decrease of prevalent diseases such as cirrhosis, diabetes, cancer and cardiovascular diseases [9–12]. Besides, these recent studies declare coffee as a functional aliment.

According to the manual of Good Clinical Practices (GCP) of the Spanish government, functional aliment is a concept born at 1980 decade in Japan which is defined as specifically developed aliment which is present in diet and act like a health improving and disease risk controller [13].

Considering their beneficial effects and their popularity the market niche becomes more competitive and consequently, an economic cut of the coffee production ends in many cases committing adulteration frauds. Is for that reason that the quality control of the commercial coffee products is very important to ensure their authenticity and to protect the consumer security.

3.3. COFFEE ADULTERANTS

According to the United States Pharmacopeia Convention, coffee is in the top ten of adulterated food products due to its commercial value and the deficiency of coffee beans.

Coffee adulteration is mostly performed by reducing the quality of its beans and the addition of cheaper coffee varieties and spent coffee ground or coffee by-products. Also, a growing tendency is the coffee adulteration by adding non-coffee materials conforming a long list of potential adulterants as roasted barley, wheat corn, maize, soybeans, and other grains, seeds and plants with the aim to resemble coffee beans in terms of color, particle size and texture [14,15].

The presence of adulterants in coffee implicates a direct quality reduction of the product than may cause economic and health issues. Mislabeled coffees could cost perjuries to human health via allergies, intolerances and various offenses to the immunologic human system depending on the adulterant quantities and the types of adulterants used to reduce the cost of the coffee production.

In this work, the adulterant to be analyzed is chicory (*Cichorium intybus*) also known as Belgian endive which is a perennial herb that belongs to the *Asteraceae* family. Chicory dried roots are usually used as coffee substituent due to a less harvesting sensibility permitting its cultivation throughout the world, making chicory "coffee" more affordable. Furthermore, chicory "coffee like" taste is catalogued as bitter and slightly sweet with caramel notes reminding the coffee taste with the particularity of an absence of caffeine [16,17].

Besides, the literature describes recent studies that underlined the positive impact on human health of chicory due to its bioactivity associated to their high polyphenolic and mineral content. In fact, the ancient Egyptian culture already used this herb for medical purposes. Also, it has been stated that chicory is one of the major natural sources of inulin which decreases the risk of any gastrointestinal diseases. The chicory root extract has been catalogued as safe, non-toxic and non-mutagenic administered orally [18] although it has been stated as an allergic sensibilization reaction similar to edibles members of its family-like endive (*Cichorium endivia*) or lettuce (*Lactuca sativa*) with oral, cutaneous and respiratory symptomatology [19].

3.4. METHODOLOGIES

Historically, food fraud analyses have been performed based on single analyte providing scarce information to control the authenticity of an aliment regarding its origin, manufacturing practices and food extension [20].

Globalization has led to the necessity of the development of optimized analytical methods able to return trustable results minimizing time and costs to be able to analyse a large volume of food products. For example, the necessity of trained and skilled analysts when microscopic inspection techniques are used, impossibilities the realization of a large amount of analysis although the technique gives good semi-quantitative trustable results [21]. Furthermore, the challenging necessity of researchers to develop these methods are faced to many problematics such as, the lack of legislation in the validation of methods, difficulty to obtain fully identified

samples to construct models and the absence of validated reference materials. This scenario produces the development of multiple methodologies instead of an specific and validated one [22].

Methods for dealing with authentication issues can be classified into two main groups regarding the nature of the analysis: Targeted and non-targeted methods.

On the one hand, a targeted method is focused on the detection or quantification of predefined analytical markers. These analytical markers are responsible directly (primary marker) when its occurrence directly addresses the authentication issue and indirectly (secondary marker) when provides indirect information of the product authenticity by, for example, the presence of chemical elements, isotopic or metabolites ratios, permitting the indirect authentication of the products of interest. Also, when multiple secondary markers are involved, the approach is known as profiling which is also used to indirectly authenticate samples. Specific targeted methods are generally more complicated although can detect analytes within the concentration range of ng/L in complex matrices [14].

On the other hand, non-targeted methods are based on the study of instrumental responses (fingerprinting) where there are numerous unspecified signals, including both relevant and irrelevant ones to discriminate an adulteration. These techniques generate large amounts of information leading to the necessity to use multivariable statistical programs to recover the valuable information able to detect an adulteration, being very useful when no primary or secondary markers are defined.

For the analysis of coffee samples, some analytical targeted and non-targeted strategies are described in the literature. The use of separation techniques such as liquid chromatography [6,8,23–25], gas chromatography (GC) [26,27], and capillary electrophoresis (CE) [28] coupled to ultraviolet (UV), fluorescence (FL), and mass spectrometry (MS) detection were usually employed to address the classification of different coffee samples.

3.5. CHEMOMETRICS

According to the International Chemometric Society (ICS), chemometrics is a chemistry discipline that uses mathematical and statistical methods to design and select optimal measurement procedures and experiments to provide the maximum information coming from the chemical analysis of chemical data. These methodologies permitted the evolution from the

classical analytical calibration (Univariate calibration) to multivariate calibration leading to the simultaneous measurement of several data for each sample with the construction of data vectors.

It is important to distinguish between two types of chemometric methods: supervised and unsupervised, regarding the previous knowledge of the existence of sample classes or groups. Unsupervised methodology will separate into groups only basing on the latent variables estimated from the X-matrix which consists on a response matrix (analytes concentration and chromatograms). Furthermore, the supervised methodology will also require the knowledge of a Y-matrix including the respective sample classes.

Besides, using chemometrics is possible to employ pre-treatments of the data to change the amplitude and the magnitude of the scale of the variables employed. In this work, for instance an autoscale pre-processing was employed.

As previously commented, an interesting approach to chemometrics is the application of multivariate analysis methods to treat all data. Principal component analysis (PCA), partial least squares-discriminant analysis (PLS-DA) and partial least squares (PLS) regression, are the multivariate analysis methods used in this study.

More details of the theoretical background of these statistical methodologies are addressed elsewhere [29].

3.5.1. Principal component analysis (PCA)

Principal component analysis (PCA), is a non-supervised compression technique capable to generate score plots permitting the qualitatively discrimination of class samples. PCA consists of a multivariate statistical method able to reduce the dimensionality of a large set of data without losing relevant information transforming the correlated variables into a new set of independent variables, also known as principal components (PCs). PCA generates a scores matrix (T) that efficiently condenses the experimental information contained in X-matrix [29]. Every PC corresponds to a new orthogonal axis variable which collects a determined variance of the original variables. The enumeration of PC follows the ranking of variance explained being PC1 the axis that collects the highest variance in the data set. The following PCs (PC2, PC3...) can be extracted similarly although providing less information as can be seen in Figure 3.1.

Mathematically, in a PCA model, the data is structured in a matrix X where the columns are studied variables and the rows are the different samples. The original matrix (X) is decomposed into a scores matrix (T) and a loadings matrix (P) as can be shown in Figure 3.2.

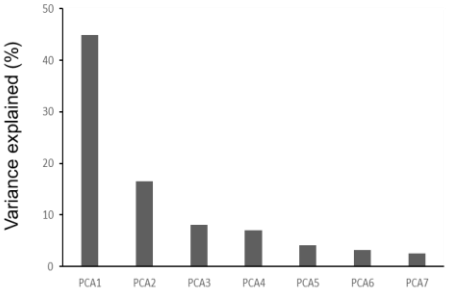

Figure 3.1. Example of the variance explained in function of the PC

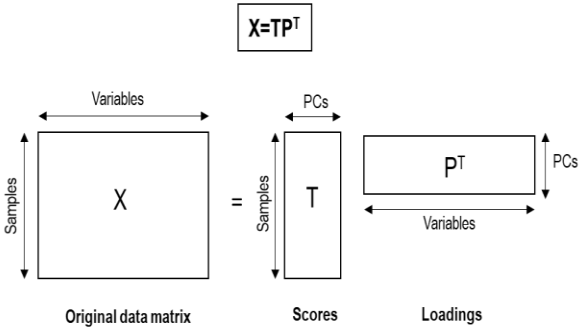

Figure 3.2. PCA decomposition of the original data matrix X into scores and loadings matrices

3.5.2. Partial least squares-discriminant analysis (PLS-DA)

Partial least squares-discriminant analysis (PLS-DA) is a supervised multivariate statistical method that is based on the research of a model capable to discriminate between two or more predefined classes. In other words, PLS-DA is a discriminatory classification method that correlates the X-matrix with the Y-matrix using latent variables (LVs). The X-matrix contains the sample set of signals and Y matrix is a binary data matrix that contains the analyzed samples and their respective class assignation.

As in the PCA model, the PLS-DA model reduces the dimensionality of their data but in this case in latent variables (LVs). After LV selection, it is possible to plot with LVs as axes a score plot to study the formation of clusters indicating the separation between sample classes.

3.5.3. Partial least squares (PLS)

Partial least-squares (PLS) is a multivariate regression method based on the construction of a model capable to predict concentration levels. This is achieved by the correlation of an X matrix which contains compositional profiles or chromatographic fingerprints of the samples and an Y-matrix of concentrations using latent variables (LVs), permitting the construction of a regression calibration model. This method possibilities the quantification of external samples making it suitable to predict the adulterant level of samples.

4. OBJECTIVES

The aim of this work is to develop two simultaneous simple and efficient non-targeted HPLC fingerprinting methods with ultraviolet (HPLC-UV) and fluorescence (HPLC-FLD) detection to achieve the characterization, classification and authentication of different instant coffee beverages and chicory samples. Data will be further use to assess the coffee authenticity. In addition, the proposed non-targeted fingerprinting methodologies will be evaluated to detect and quantify the adulteration levels in instant coffee beverages with chicory as adulterant. To achieve this aim, the next steps will be performed:

1. Coffee and chicory samples will be submitted to a simple sample treatment to extract their bioactive compounds.
2. The obtained sample extracts will be analyzed using HPLC-UV and HPLC-FLD methods to obtain characteristic fingerprints of the samples.
3. The obtained HPLC-UV and HPLC-FLD fingerprints will be evaluated as sample chemical descriptors to assess sample characterization and classification by chemometrics.
4. Principal Component Analysis (PCA) in this precise study, will only be employed to study the reproducibility of the method and the robustness of the chemometric results.
5. Partial Least Squares-Discriminant Analysis (PLS-DA) will be employed to study the classification of the samples according to their typology: Coffee, decaffeinated coffee, and chicory.
6. PLS regression model will be employed to quantify the concentration levels in instant coffee and decaffeinated coffee samples adulterated with chicory.

5. EXPERIMENTAL SECTION

5.1. CHEMICALS AND STANDARD SOLUTIONS

For the HPLC-UV and HPLC-FLD employed methods the analytical reagents employed in the mobile phase were:

-Methanol (99.9% from Panreac, Barcelona)

-Formic acid (≥96% from Sigma-Aldrich, Germany)

-Milli-Q water. Water was purified filtering through a 0.22 µm nylon filter integrated into the Milli-Q system using an Elix 3 coupled to the Milli-Q system (Millipore, Bedford, MA, USA).

-Commercial mineral water obtained from Eroski (Barcelona, Spain) was employed for coffee brewing.

5.2. INSTRUMENTATION AND METHODS

An Agilent 1100 Series HPLC instrument from Agilent HPLC 1100 Series (Waldronn, Germany) equipped with a binary pump (G1321A), an autosampler (G1367A), a degasser (G1379A), a diode array detector (G1315B), a fluorescence detector (G1321A) and a PC with the Agilent Chemstation software (Rev. A 10.02) was employed to obtain simultaneously HPLC-UV and HPLC-FLD chromatographic fingerprints.

Chromatographic separation was performed using a Kinetex® C18 reversed phase column (100 x 4.6mm I.D., 2.6 µm particle size) obtained from Phenomenex (Torrance, California, USA). Separation was performed under gradient elution mode, employing 0.1% formic acid in water *(v/v)* as solvent A, and methanol as solvent B. The gradient program employed is shown in Table 5.1. It consisted on a gradient by increasing methanol component from 3 to 75% in 30 min. After that, methanol component increases up to 95% in 2 min. Then, there was an isocratic step of 2 min. Finally, the gradient elution program come back to mobile phase initial conditions in 0.2 min and there was an isocratic step of 5.8 min to guarantee column re-equilibration. The mobile phase flow rate was 0.4 mL/min and the standard injection volume of the sample was 5.0 µL.

HPLC-UV chromatographic fingerprints were registered at 250 nm, 280 nm, 310 nm, 370 nm and 550 nm and HPLC-FLD chromatographic fingerprints were registered with an excitation wavelength of 310 nm and emission wavelengths of 440 nm, 340 nm, 370 nm and 410 nm.

Table 5.1. Elution gradient used for HPLC separation

Time [Minutes]	*%B*	*Elution Mode*
0-30	3→75	Lineal Gradient
30-32	75→95	Lineal Gradient
32-34	95	Isocratic
34-34.2	95→3	Return to initial conditions
34.2→40	3	Isocratic

5.3. SAMPLES

A total of 88 samples of commercial instant coffee and chicory samples purchased from several supermarkets in Barcelona (Spain) were analyzed. The samples were classified in 3 main groups: Coffee (40 samples), decaffeinated coffee (26 samples) and Chicory (22 samples). Commercial brands and number of samples are summarized in Table 5.2.

For the sample treatment 0.5L of water was heated in a water heater, until water boiling temperature. After temperature stabilization, 25mL of water were added to 0.5g of soluble coffee. In order to ensure a complete dissolution and homogenization, samples were shaken for 2 min using a Vortex (Stuart, Stone, United Kingdom) and then, the extracts were sonicated during 20 min on an ultrasound bath. After that, samples were filtered with a 0.45 μm filter in order to remove suspension particles, if present, and transferred into 2 mL injection vials. Finally, samples were kept in fridge at -4ºC until HPLC analysis.

Some ground Chicory samples (not instant ones) were prepared using an Italian coffee maker. In order to ensure the reproducibility, a stablished methodology on previous studies for the research group was employed [8]. Briefly, sample treatment consisted of compress 35 g of the chicory sample in the Italian coffee maker, adding 400 mL of mineral water and brewing for 15 minutes using a Bunsen burner. After achieving chicory lixiviation, samples were filtered and stored at the same conditions explained for the rest of the analyzed samples.

Table 5.2. Description of the commercially available coffee and chicory samples analyzed

Sample	*Commercial brand*	*Time*	*Number of samples**
Coffee	Marcilla Classic Natural	Instant	4
	Marcilla Crème Natural	Instant	4
	Nescafé Classic	Instant	6
	Nescafé Classic Crème	Instant	4
	Nescafé Black Roast	Instant	4
	Nescafé Gold	Instant	6
	Nescafé Origins	Instant	6
	Eroski	Instant	6
Decaf Coffee	Marcilla Classic	Instant	4
	Marcilla Crème	Instant	4
	Nescafé Classic	Instant	6
	Nescafé Gold	Instant	6
	Eroski	Instant	6
Chicory	Chicorée lima original	Ground	8
	Chicorée Biocop	Instant	8
	Chicorée La niña	Ground	6

*Number of samples collected from different containers

Coffee samples were analyzed randomly with the proposed HPLC method. Furthermore, a quality control (QC) was prepared by mixing 50 µL of each sample extract in order to evaluate the reproducibility of the method and the robustness of the chemometric results. For that purpose, QCs were analyzed every 10 samples.

5.4. DATA ANALYSIS

The capacity of HPLC-UV and HPLC-FLD fingerprints as chemical descriptors of the analyzed samples was evaluated employing chemometrics. The obtained HPLC chromatograms from the Chemstation software (Agilent) were exported to a spreadsheet using Unichrom® to create data matrices. After working with all the wavelengths proposed in both methods, the best results were obtained for 280 nm in HPLC-UV and 310 nm of emission and 410 nm of excitation for HPLC-FLD, so all the results explained in this work are referred to these selected wavelengths.

Stand Alone Chemometrics Software® (SOLO) from Eigenvector Research was employed for PCA, PLS-DA and PLS calculations.

PCA was only employed in order to study the reproducibility and robustness during the analysis although its characteristics would also permit the discrimination between class samples using score plots. To classify the samples, we used a supervised chemometric method such as

PLS-DA, and to quantify the adulteration levels in adulterated samples, PLS regression was performed.

The X data matrices contained HPLC-UV and HPLC-FLD signals and their dimensions were 101 samples x 6001 UV absorbance signals and 101 samples x 5555 FLD intensity signals, respectively. Instead, Y-data matrices defined each sample class in PLS-DA whereas for PLS, Y-data matrices consisted of chicory adulteration percentages. All fingerprints were autoscaled to provide the same weight to each variable. The number of LVs in PLS-DA and PLS was stablished by considering the first relevant minimum of the CV error using a Venetian blind approach.

The applicability of the built PLS-DA models was evaluated by external validation. For that purpose, 70% of a sample group (randomly selected) was used as the training set (data set used for model generation and optimization), and the remaining 30% of the samples, used as unknown samples, constituted an independent validation set (used for model prediction).

6. RESULTS AND DISCUSSION

6.1. HPLC FINGERPRINTS

As previously commented, the intention of the project was to classify coffee and chicory samples with the aim to prevent alimentary frauds. In order to achieve that objective, non-targeted HPLC-UV and HPLC-FLD fingerprintings were employed.

A non-targeted method is the combination between emerging analytical methods and statistic software to treat the high amount of obtained data. It is based on the analysis of the instrumental responses (fingerprinting) without any previous information of the relevant or irrelevant components present in the samples. They are very useful to analyse complex matrices, such as foods, from which normally there is not a complete information of their chemical composition, avoiding the necessity of using standards for quantitation purposes such as in the case of targeted methodologies [14].

In this study, a Kinetex® C18 reversed-phase column (100 x 4.6 mm I.D., 2.6 µm particle size) was employed to obtain the chromatographic separation under gradient elution conditions as indicated in the experimental section.

Once the chromatographic fingerprints were obtained, a first visual comparison between the different types of analyzed samples, as well as between HPLC-UV and HPLC-FLD fingerprints was performed. As an example, Figure 6.1 shows the (a) HPLC.FLD chromatographic fingerprints obtained at 310 (excitation) and 410 nm (emission) and (b) the HPLC-UV chromatographic fingerprints obtained at 280 nm for a selected sample of each class.

As can be seen in the figure, HPLC-FLD fingerprints are richer regarding the number of chemicals detected among those extracted from the analyzed samples. In contrast, HPLC-UV fingerprints seem to detect only some specific, although quite intense, compounds.

On the other hand, when coffee samples are compared between them, in general the two types of coffees in both proposed methods had a similar chromatographic fingerprint although differing in the peak intensities. However, these fingerprints are clearly different from those

obtained for chicory samples, which were characterized by different peak signal distributions as well as peaks intensities. Besides, when samples are compared in UV detection, it can be noticed that chicory fingerprints disrupt with the general profile tendency of the rest of the samples by having different chemical descriptors where it highlights an intense peak at 9 min.

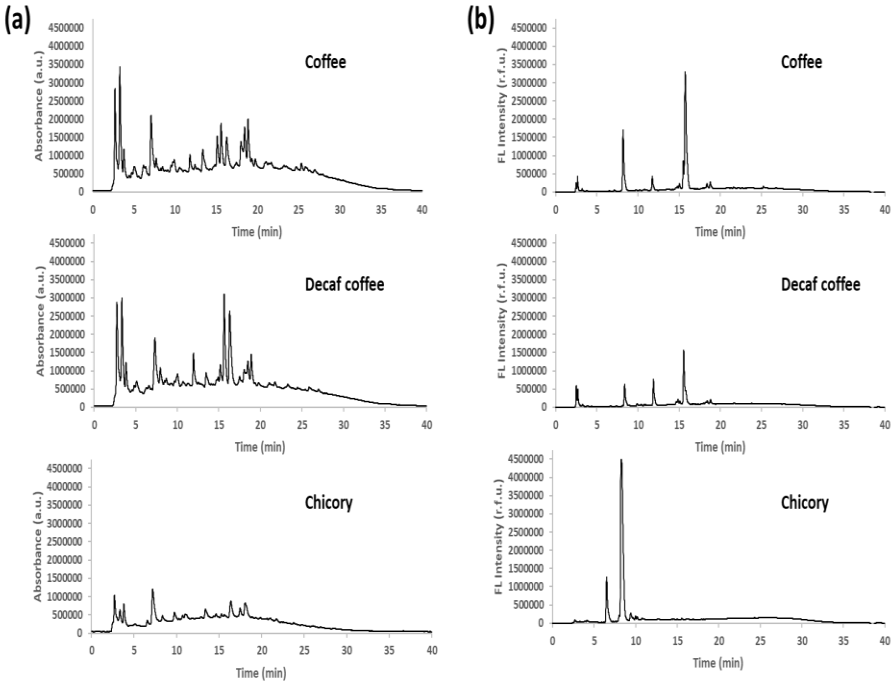

Figure 6.1. (a) HPLC-FLD chromatographic fingerprints registered at 310 nm (excitation) and 410 nm (emission) and (b) HPLC-UV fingerprints acquired at 280 nm for selected coffee and chicory samples

Furthermore, another highlighted feature when comparing both coffees using FL detection is an important intense signal at around 15 minutes in the decaffeinate coffee. In contrast, when comparing both coffees using UV detection is important to consider the important difference in peak intensity of two remarkable signals (at around 8 min and 16 min) that have a significant higher intensity in the coffee samples in comparison to the decaffeinate one.

Noteworthy differences were then observed in the obtained fingerprints among the three groups of analyzed samples and since these features were reproducible among the samples belonging to the same group, HPLC-UV and HPLC-FLD fingerprintings were then evaluated to classify the samples through multivariate chemometric approaches.

6.2. EXPLORATORY PCA

In order to achieve sample classification, as a first approach the chromatographic fingerprints obtained for all the analyzed samples were submitted to PCA chemometric method in order to detect possible irregularities during analysis time due to instrumentation instability.

The PCA scores plots of PC1 vs PC2 obtained when using HPLC-UV and HPLC-FLD fingerprints as chemical descriptors are shown in Figure 6.2a and 6.2b respectively. As can be seen, when using the HPLC-UV fingerprinting method (Figure 6.2a), QCS formed a compact cluster, which ensures the reproducibility of the proposed method in function of time. However, with the HPLC-FLD fingerprinting method (Figure 6.2b), QC's appeared in a straight-line following the QC injection order in the sample sequence employed indicating a clear tendency of the signal decline during the analysis time (b). To sort out that instrumental problem, a mathematical correction was employed in order to readjust the analysis results. The correction consisted of relativize sample intensities to the nearest QC by employing its intensities as a common intensity reference. This could be achieved dividing every sample intensity into the nearest QC intensity measurement, while QC signal was divided by itself (resulting in fingerprints variables defined by the value 1). The PCA results obtained after applying this correction can be seen in the Figure 6.2c.

As both chromatographic fingerprints were obtained simultaneously for each analyzed sample (HPLC-UV-FLD method), the differentiation on the repeatability of the obtained results cannot be attributed to reproducibility problems during injection and chromatographic separation but to the decay of the fluorescence intensity throughout the sequence of analysis time (probably because of the drift in the Xe lamp intensity), in comparison to the UV-vis detector.

After this point, this correction will be applied to all the chemometrics methods when employing HPLC-FLD fingerprints.

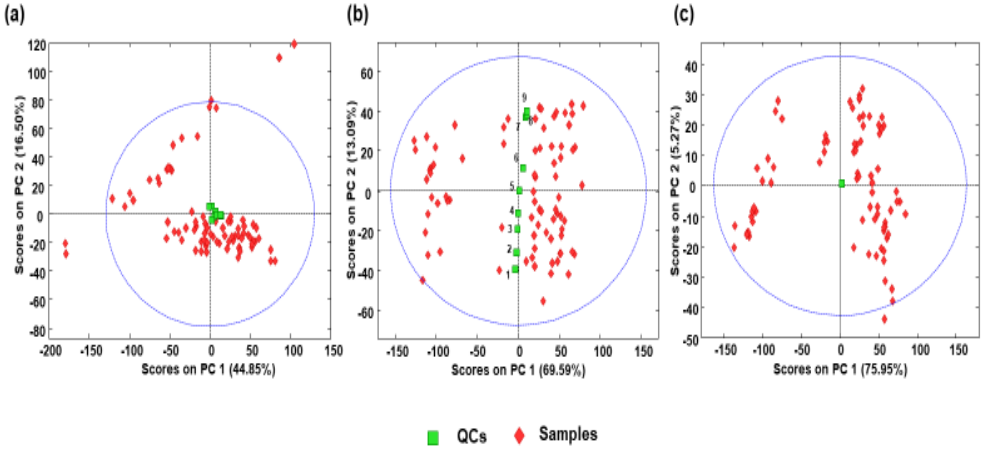

Figure 6.2. PCA scores plot map PC1 vs PCA2 of (a) UV method, (b) FLD method with no QC correction and (c) FLD method with QC correction

6.3. SAMPLE CLASSIFICATION BY PLS-DA

Once it is ensured that the elution order do not affect the reproducibility of the methods, it is possible to proceed to characterize and classify samples using PLS-DA.

The PLS-DA model was intended to generate scores scatter map plots in order to check if the chemical descriptors achieved with both proposed methods can discriminate the different samples into 3 clusters: coffee, decaffeinate coffee and chicory. As can be seen in Figure 6.3, in both methods, the PLS-DA score plots defined by LV1 vs LV2 allowed to discriminate chicory samples (which occupy the left section of the scores plots forming a diagonal cluster) from regular and decaffeinated coffee samples. So, as a first approach, it can be expected that both fingerprinting methods would be able to authenticate coffees adulterated with chicory.

However, to discern between coffee and decaffeinated coffee, both methods show a partial overlaid cluster between these two sample groups, although samples tend to be grouped according to coffee type.

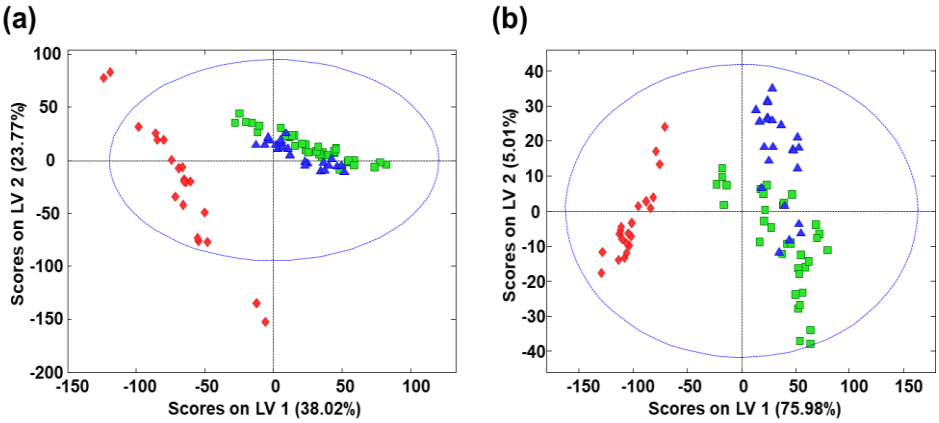

Figure 6.3. PLS-DA LV1 vs LV2 with (a) UV method and (b) FLD method

To evaluate the capability of both HPLC-UV and HPLC-FLD fingerprinting methods, the classification models were studied in pairs to perform a PLS-DA validation. The studied pair cases were coffee vs decaffeinated coffee, coffee vs chicory and decaffeinated coffee vs chicory. PLS-DA validation was performed by employing only 70% of the samples of each class (randomly selected) to build the model and the remaining 30% of the samples of each class as predictors. Assignation plots of sample vs Y predicted 1 (where the horizontal red line delimites between classes) obtained are shown in Figure 6.4. As can be seen, 100% classification rates were obtained with both fingerprinting methodologies when chicory was discriminated from coffee and decaffeinate coffee, in agreement with the information extracted from the previous PLS-DA scores plot where a completely independent chicory cluster was observed (Figure 6.3). In contrast, as it was stated in the Figure 6.3 too, a most complex situation was observed when coffee is tried to be discriminated from decaffeinated coffee. In this case, sample classification rates for the predictions were 94.4% and 83.3% with HPLC-UV and HPLC-FLD fingerprinting methods respectively.

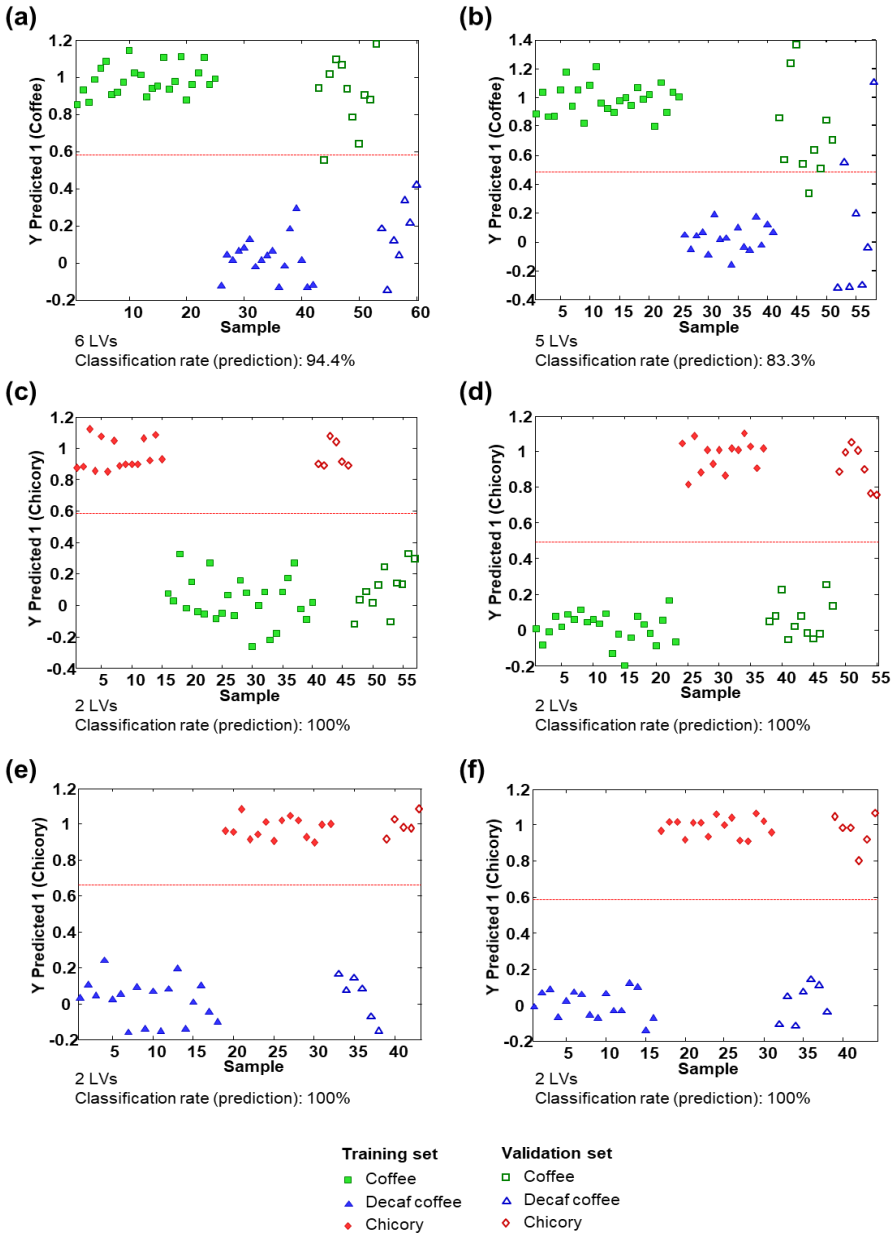

Figure 6.4. PLS-DA pair validation of (a) Coffee vs Decaffeinated with HPLC-UV, (b) Coffee vs Decaffeinated with HPLC-FLD, (c) Coffee vs Chicory with HPLC-UV, (d) Coffee vs Chicory with HPLC-FLD, (e) Decaffeinated vs Chicory with HPLC-UV and (f) Decaffeinated coffee vs Chicory with HPLC-FLD.

In conclusion, the information we can extract from the PLS-DA methods performed is that the evaluated fingerprinting chemicals descriptors are able to achieve the characterization and classification of the samples in both methods achieving an excellent chicory discrimination. For discrimination among coffees the best results are achieved when employing UV detection. This experimental results can be explained due to the lack of absorbance in fluoresence of caffeine being a key chemical descriptor to discern between regular and decaffeinated coffee.

The principal result that can be highlighted up to this point is that 100% of classification rate were achieved when chicory is involved with both proposed. Therefore, adulteration cases of coffee and decaffeinated coffee with chicory will be studied by PLS regression chemometric method to quantify the chicory adulterant percentage in the coffee samples.

6.4. PREDICTION OF ADULTERATION PERCENTAGES BY PLS

In order to detect and quantify adulterations in instant coffee PLS regression was employed. For that purpose, a regular coffee sample and a decaffeinated one were chosen randomly and were adulterated with chicory, at different adulterant levels as described in Table 6.1, using similar adulteration rations as described in a previous work[8]. Then, coffees were brewed following the same procedure described in the experimental section. PLS calibration and validation sets of adulterated samples were prepared as indicated in Table 6.1. Furthermore, an adulterated sample containing 50% chicory was employed as a QC standard. Each adulteration level was prepared by quintuplicate.

Table 6.1. Samples used in the PLS adulteration studies as calibration and validation sets.

	COFFEE OR DECAF COFFEE [%]	CHICORY AS ADULTERANT [%]
CALIBRATION SET	100	0
	80	20
	60	40
	40	60
	20	80
	0	100
VALIDATION SET	85	15
	75	25
	50	50
	25	75
	15	85

A total of 55 sample extracts were then prepared for each adulteration case, and they were randomly analyzed to obtain the chromatographic fingerprints. For each adulteration case studied, first, PLS-DA was evaluated, in order to study if the distribution of the samples in the scores plot followed a coherent trend according to its adulteration percentage, and the resulting scores plots of LV1 vs LV2 for the coffee sample adulterated with chicory are depicted in Figure 6.5.

As it can be seen in the PLS-DA scores plot, for both fingerprinting methods a coherent trend according to the adulterant percentage is observed, with 0% of chicory adulteration pure original coffee) at the left of the plots, and 100% of chicory at the right side. In between, samples are distributed according to their adulterant percentage, with the increment of the adulterant directly related to the LV1. Similar and comparable results were also obtained with the decaffeinate coffee adulterated with chicory for both HPLC-UV and HPLC-FLD data.

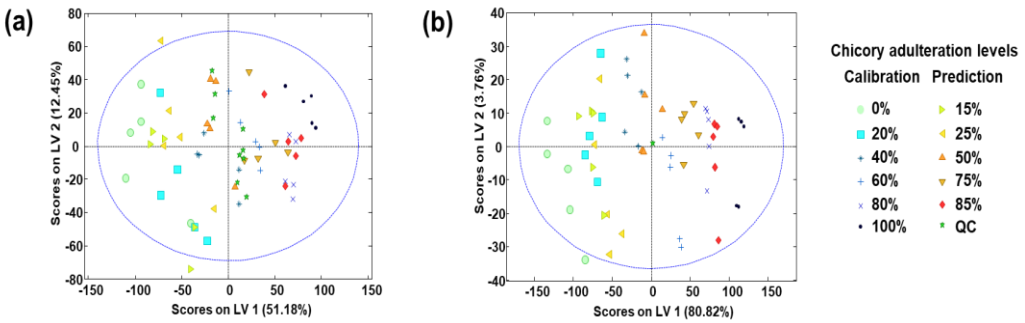

Figure 6.5. PLS-DA scores plot of LV1 vs LV2 for the coffee adulterated with chicory when using (a) HPLC-UV and (b) HPLC FLD fingerprints.

Then, PLS regression was employed to quantify the chicory adulterant levels in the adulteration cases under study. PLS predictions are depicted in Figure 6.6, representing the scatter plot of actual (measured) vs. predicted values for the coffee adulterated with chicory case using both fingerprinting methods. The following statistic parameters were used to evaluate the method performance: R^2 that indicates the degree of correlation of the regression, and the RMSE (Root mean square errors) of the calibration (RMSEC, calibration error) and of the prediction (RMSEP, prediction error). Results for both adulteration studies are summarized in Table 6.2.

As can be seen in Figure 6.6, a pretty lineal tendency is followed forming two close tendency correlation lines with slopes close to 1 and interccepts close to 0, and where predicted Y values practically matches to the measured Y values.

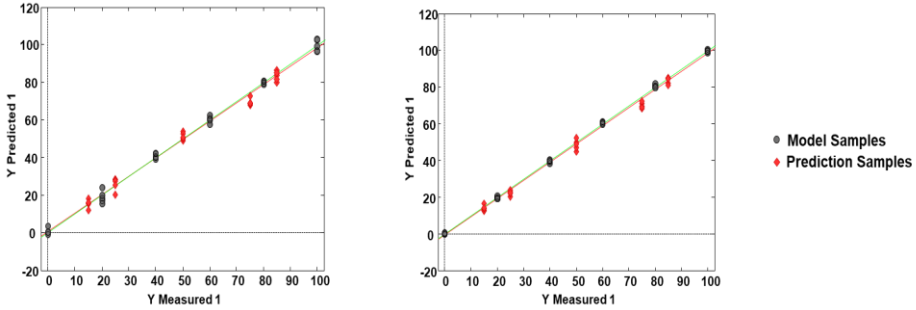

Figure 6.6. PLS adulteration models for Coffee adulterated with Chicory. Y measured vs Y predicted plots with (a) HPLC-UV and (b) HPLC-FLD fingerprints

Table 6.2. Results for the evaluation of the adulteration of instant coffee and instant decaffeinated coffee with instant chicory using HPLC-UV and HPLC-FLD as chemical descriptors for PLS

	HPLC-UV fingerprints				HPLC-FLD fingerprints			
	LVs	Linearity [R^2]	Calibration error [%]	Prediction error [%]	LVs	Linearity [R^2]	Calibration error [%]	Prediction error [%]
Instant coffee	3	0.996	2.1	2.4	5	0.999	0.8	2.9
Decaf coffee	5	0.996	2.0	3.5	5	1.000	0.7	3.2

The obtained PLS calibration models were very good, as indicated by the low calibration errors (≤2.1%), bias values tending towards zero, and good linearities (R^2≥0.996) (Table 6.2), independently of the type of fingerprints employed. Excellent prediction performance was also accomplished, with prediction errors lower than 3.5% in all cases. Although very good results were achieved with both HPLC-UV and HPLC-FLD fingerprints, overall HPLC-FLD seemed to provide better calibration (≤0.8%), and similar to better prediction errors than HPLC-UV, probably due to the superior selectivity of fluorescence detection.

7. CONCLUSIONS

In this work, two simultaneous non-targeted HPLC-UV and HPLC-FLD fingerprinting methods have been developed to deal with the characterization, classification and authentication of instant coffee and chicory beverages. After evaluating all the results obtained in this study, several conclusions can be extracted:

- Both fingerprinting methods showed to be excellent to achieve the instant coffee authentication and to detect and quantify frauds with chicory adulteration.

- After analyzing the PLS-DA score plots obtained for both methods it can be concluded that the methods allow to classify the analyzed samples, providing a good clustering of coffee and decaffeinate coffee perfectly discriminated from chicory samples.

- Satisfactory results were obtained for PLS-DA pair validation achieving a 100% classification rate for the classification of coffee or decaffeinate coffee against chicory samples for both proposed methods. In the case of coffee versus decaf coffee, the classification obtained for HPLC-UV was 94.4%, providing better classification results than HPLC-FLD where it classification rate dropped to 83.3%

- Two adulteration cases were evaluated by PLS regression, being able to predict the percentage of adulteration of the samples by obtaining calibration and prediction errors lower than 2.1% and 3.5%, respectively, for both proposed methods. The HPLC-FLD method provided better results that the HPLC-UV one, in term of better linearities and lower calibration errors.

8. REFERENCES

1. Kennedy, G.; Nantel, G.; Shetty, P. *Globalization of food systems in developing countries: impact on food security and nutrition.*; **2004**; Vol. 83; ISBN 925105228X.
2. Esteki, M.; Shahsavari, Z.; Simal-Gandara, J. Food identification by high performance liquid chromatography fingerprinting and mathematical processing. *Food Res. Int.* **2019**, *122*, 303–317, doi:10.1016/j.foodres.2019.04.025.
3. Tarannum, N.; Khatoon, S.; Dzantiev, B.B. Perspective and application of molecular imprinting approach for antibiotic detection in food and environmental samples: A critical review. *Food Control* **2020**, *118*, 107381, doi:10.1016/j.foodcont.2020.107381.
4. Navaneethan, Y.; Effarizah, M.E. Prevalence, toxigenic profiles, multidrug resistance, and biofilm formation of Bacillus cereus isolated from ready-to eat cooked rice in Penang, Malaysia. *Food Control* **2020**, *121*, 107553, doi:10.1016/j.foodcont.2020.107553.
5. Cuadros-Rodríguez, L.; Ruiz-Samblás, C.; Valverde-Som, L.; Pérez-Castaño, E.; González-Casado, A. Chromatographic fingerprinting: An innovative approach for food "identitation" and food authentication - A tutorial. *Anal. Chim. Acta* **2016**, *909*, 9–23, doi:10.1016/j.aca.2015.12.042.
6. Angeloni, S.; Nzekoue, F.K.; Navarini, L.; Sagratini, G.; Torregiani, E.; Vittori, S.; Caprioli, G. An analytical method for the simultaneous quantification of 30 bioactive compounds in spent coffee ground by HPLC-MS/MS. *J. Mass Spectrom.* **2020**, *55*, doi:10.1002/jms.4519.
7. Puerta, G. Composición química de una taza de café. *Av. en Química* **2011**, 1–12.
8. Núñez, N.; Collado, X.; Martínez, C.; Saurina, J.; Núñez, O. Authentication of the origin, variety and roasting degree of coffee samples by non-targeted HPLC-UV fingerprinting and chemometrics. Application to the detection and quantitation of adulterated coffee samples. *Foods* **2020**, *9*, doi:10.3390/foods9030378.
9. Schmit, S.L.; Nwogu, O.; Matejcic, M.; DeRenzis, A.; Lipworth, L.; Blot, W.J.; Raskin, L. Coffee consumption and cancer risk in African Americans from the Southern Community Cohort Study. *Sci. Rep.* **2020**, *10*, 1–8, doi:10.1038/s41598-020-72993-6.
10. Romualdo, G.R.; Prata, G.B.; da Silva, T.C.; Evangelista, A.F.; Reis, R.M.; Vinken, M.; Moreno, F.S.; Cogliati, B.; Barbisan, L.F. The combination of coffee compounds attenuates early fibrosis-associated hepatocarcinogenesis in mice: involvement of miRNA profile modulation. *J. Nutr. Biochem.* **2020**, *85*, 108479, doi:10.1016/j.jnutbio.2020.108479.
11. Zeng, A.; Liang, X.; Zhu, S.; Liu, C.; Wang, S.; Zhang, Q.; Zhao, J.; Song, L.

Chlorogenic acid induces apoptosis, inhibits metastasis and improves antitumor immunity in breast cancer via the NF-κB signaling pathway. *Oncol. Rep.* **2020**, *45*, 717–727, doi:10.3892/or.2020.7891.

12. Montenegro, J.; dos Santos, L.S.; de Souza, R.G.G.; Lima, L.G.B.; Mattos, D.S.; Viana, B.P.P.B.; da Fonseca Bastos, A.C.S.; Muzzi, L.; Conte-Júnior, C.A.; Gimba, E.R.P.; et al. Bioactive compounds, antioxidant activity and antiproliferative effects in prostate cancer cells of green and roasted coffee extracts obtained by microwave-assisted extraction (MAE). *Food Res. Int.* **2020**, *140*, doi:10.1016/j.foodres.2020.110014.
13. Bartrina, J.; Blay Cortés, G.; Echevarría Guitiérrez, F.J.; Inmaculada;, G.C.; Hernández Cabria, M.; Iglesias Barcia, J.R.; López Díaz-Ufano, M.L. *Atención primaria de calidad: Guía de buena práctica clinica en Alimentos funcionales*; 2011; ISBN 9788469497999.
14. Cheah, W.L.; Fang, M. HPLC-Based Chemometric Analysis for Coffee Adulteration. *Foods* **2020**.
15. Song, H.Y.; Jang, H.W. Analytical method to detect adulteration of ground roasted coffee. *Int. J. Food Sci. Technol.* **2019**, *54*, 256–262, doi:10.1111/ijfs.13942.
16. Indzere, Z.; Khabdullina, Z.; Khabdullin, A.; Blumberga, D. The benchmarking of chicory coffee's production. *Energy Procedia* **2018**, *147*, 631–635, doi:10.1016/j.egypro.2018.07.081.
17. Wu, T.; Cadwallader, K.R. Identification of Characterizing Aroma Components of Roasted Chicory "coffee" Brews. *J. Agric. Food Chem.* **2019**, *67*, 13848–13859, doi:10.1021/acs.jafc.9b00776.
18. Perović, J.; Tumbas Šaponjac, V.; Kojić, J.; Krulj, J.; Moreno, D.A.; García-Viguera, C.; Bodroža-Solarov, M.; Ilić, N. Chicory (Cichorium intybus L.) as a food ingredient – Nutritional composition, bioactivity, safety, and health claims: A review. *Food Chem.* **2020**, *336*, 127676, doi:10.1016/j.foodchem.2020.127676.
19. Cadot, P.; Kochuyt, A.M.; Van Ree, R.; Ceuppens, J.L. Oral allergy syndrome to chicory associated with birch pollen allergy. *Int. Arch. Allergy Immunol.* **2003**, *131*, 19–24, doi:10.1159/000070430.
20. Ballin, N.Z.; Laursen, K.H. To target or not to target? Definitions and nomenclature for targeted versus non-targeted analytical food authentication. *Trends Food Sci. Technol.* **2019**, *86*, 537–543, doi:10.1016/j.tifs.2018.09.025.
21. Pauli, E.D.; Barbieri, F.; Garcia, P.S.; Madeira, T.B.; Acquaro, V.R.; Scarminio, I.S.; da Camara, C.A.P.; Nixdorf, S.L. Detection of ground roasted coffee adulteration with roasted soybean and wheat. *Food Res. Int.* **2014**, *61*, 112–119, doi:10.1016/j.foodres.2014.02.032.
22. Kucharska-Ambrożej, K.; Karpinska, J. The application of spectroscopic techniques in combination with chemometrics for detection adulteration of some herbs and spices. *Microchem. J.* **2020**, *153*, 104278, doi:10.1016/j.microc.2019.104278.
23. Núñez, N.; Martínez, C.; Saurina, J.; Núñez, O. High-performance liquid chromatography with fluorescence detection fingerprints as chemical descriptors to

authenticate the origin, variety and roasting degree of coffee by multivariate chemometric methods. *J. Sci. Food Agric.* **2020**, *101*, 65–73, doi:10.1002/jsfa.10615.

24. Chirfa, G.; Merdassa, Y.; Gure, A. Salting-out Assisted Liquid–Liquid Extraction for Analysis of Caffeine and Nicotinic Acid in Coffee by HPLC–UV/Vis Detector. *J. Anal. Test.* **2020**, *4*, 298–306, doi:10.1007/s41664-020-00148-7.
25. Abreu, M.B.; Marcheafave, G.G.; Bruns, R.E.; Scarminio, I.S.; Zeraik, M.L. Spectroscopic and Chromatographic Fingerprints for Discrimination of Specialty and Traditional Coffees by Integrated Chemometric Methods. *Food Anal. Methods* **2020**, *13*, 2204–2212, doi:10.1007/s12161-020-01832-1.
26. Ongo, E.A.; Montevecchi, G.; Antonelli, A.; Sberveglieri, V.; Sevilla, F. Metabolomics fingerprint of Philippine coffee by SPME-GC-MS for geographical and varietal classification. *Food Res. Int.* **2020**, *134*, doi:10.1016/j.foodres.2020.109227.
27. Marek, G.; Dobrzański, B.; Oniszczuk, T.; Combrzyński, M.; Ćwikła, D.; Rusinek, R. Detection and differentiation of volatile compound profiles in roasted coffee arabica beans from different countries using an electronic nose and GC-MS. *Sensors (Switzerland)* **2020**, *20*, doi:10.3390/s20072124.
28. Pérez-Míguez, R.; Sánchez-López, E.; Plaza, M.; Marina, M.L.; Castro-Puyana, M. Capillary electrophoresis-mass spectrometry metabolic fingerprinting of green and roasted coffee. *J. Chromatogr. A* **2019**, *1605*, doi:10.1016/j.chroma.2019.07.007.
29. Olivieri, A.C. *Introduction to Multivariate Calibration*; **2018**; ISBN 9783319970967.

9. ACRONYMS

A.U: Arbitrary units

CE: Capillary electrophoresis

FL: Fluorescence

FLD: Fluorescence detection

GC: Gas chromatography

GCP: Good clinical practices

HPLC: High-performance liquid chromatography

ICS: International chemometric society

ICO: International coffee organization

I.D: Internal diameter

LV: Latent variable

MS: Mass spectrometry

PC: Principal component

PCA: Principal component analysis

PLS: Partial least squares

PLS-DA: Partial least squares - discriminant analysis

QC: Quality control

R.F.U: Relative fluorescence units

RMSE: Root mean square error

RMSEC: Root mean square error of calibration

RMSEP: Root mean square error of prediction

SOLO: Stand alone chemometrics software

UV: Ultraviolet