

Article

Determination of HPLC-UV Fingerprints of Spanish Paprika (*Capsicum annuum* L.) for Its Classification by Linear Discriminant Analysis

Xavier Cetó ^{1,*}, Núria Serrano ^{1,*}, Miriam Aragón ¹, Alejandro Gámez ¹, Miquel Esteban ¹, José Manuel Díaz-Cruz ¹ and Oscar Núñez ^{1,2,3}

¹ Department of Chemical Engineering and Analytical Chemistry, University of Barcelona, Martí i Franquès 1-11, E08028 Barcelona, Spain; xavier.ceto@ub.edu (X.C.); miriara2@hotmail.com (M.A.); alejandrogamezcr@gmail.com (A.G.); miquelstebanc@ub.edu (M.E.); josemanuel.diaz@ub.edu (J.M.D.-C.); oscar.nunez@ub.edu (O.N.)

² Research Institute in Food Nutrition and Food Safety, University of Barcelona, Av. Prat de la Riba 171, Edifici Recerca (Gaudí), E-08901 Santa Coloma de Gramanet, Barcelona, Spain

³ Serra Hunter Fellow, Generalitat de Catalunya, Spain

* Correspondence: nuria.serrano@ub.edu; Tel.: +34-93-403-3706

Received: 29 November 2018; Accepted: 16 December 2018; Published: 18 December 2018



Abstract: The development of a simple HPLC-UV method towards the evaluation of Spanish paprika's phenolic profile and their discrimination based on the former is reported herein. The approach is based on C₁₈ reversed-phase chromatography to generate characteristic fingerprints, in combination with linear discriminant analysis (LDA) to achieve their classification. To this aim, chromatographic conditions were optimized so as to achieve the separation of major phenolic compounds already identified in paprika. Paprika samples were subjected to a sample extraction stage by sonication and centrifugation; extracting procedure and conditions were optimized to maximize the generation of enough discriminant fingerprints. Finally, chromatograms were baseline corrected, compressed employing fast Fourier transform (FFT), and then analyzed by means of principal component analysis (PCA) and LDA to carry out the classification of paprika samples. Under the developed procedure, a total of 96 paprika samples were analyzed, achieving a classification rate of 100% for the test subset (n = 25).

Keywords: HPLC-UV; Spanish paprika; polyphenols; Protected designation of origin; linear discriminant analysis; food authentication

1. Introduction

The quality of natural food products is an issue of great interest in our society, especially when dealing with products with protected designation of origin (PDO). Several health-promoting properties of those have been partly attributed to the presence of bioactive compounds, which also play an important role in food sensorial and functional properties. Paprika, or chili pepper, which is a red seasoning powder with a characteristic flavor obtained from the drying and grinding of certain varieties of red peppers (*Capsicum annuum* L.), is one of the most widely used food colorants in both culinary and industrial applications. The intense and characteristic red color of paprika is due to its high carotenoid pigment content produced during fruit ripening, that apart from conferring color to food provides important health benefits [1–4]. Along with carotenoids, chlorophylls, tocopherols, capsaicinoids, and phenolic compounds are also important bioactive compounds present in paprika. The latter act as antioxidants, and are important in plant defense responses, having an impact in the resulting fruit quality [5].

Spain is one of the main paprika producing countries in Europe, where it is mainly cultivated in two regions: Murcia, with 21% of the country's production in 2015 (1010 tones distributed over 381 ha) and Extremadura (district of La Vera, Cáceres) with 71% of total production (3362 tones distributed over 1241 ha) [6]. Both regions have been awarded a designation of origin: "Pimentón de La Vera" and "Pimentón de Murcia", respectively. Murcia paprika is mainly produced from red peppers of the *Bola* variety, whereas La Vera paprika is produced from *Jaranda*, *Jariza*, *Jeromín*, and *Bola* red pepper varieties. Depending on the red pepper variety employed, three different paprika types are commercialized: sweet (*Bola* and *Jaranda* varieties), bittersweet (*Jariza* and *Jaranda* varieties) and spicy (*Jeromín*, *Jaranda*, and *Jariza* varieties).

Phenolic compounds are aromatic secondary metabolites ubiquitously spread through the plant kingdom comprising more than 8000 substances with highly diverse structures. Over the last few years, this family of compounds has caught a lot of attention due to the recognition of their antioxidant properties and their probable role in the prevention of several diseases [7–10], as well as their contribution to the organoleptic and functional properties in food products. Furthermore, the impact of these compounds in characterization, classification, and authentication studies cannot be underestimated. Polyphenolic distribution and content in natural food products seems to be related to food features such as the plant/fruit/seed variety, the geographical climate conditions of their production area, and the cultivation and manufacturing practices, between others. Consequently, phenolic profiling can be exploited as a source of analytical data to establish product classifications, as well as for the evaluation of food quality and the detection of adulterations [11,12].

Liquid chromatography (LC) in reversed-phase mode with UV detection or coupled to mass spectrometry (LC-MS) are among the most common techniques for the determination of polyphenols in a great variety of natural food products [13–20]. The use of polyphenolic profiles has also been described to achieve the characterization and classification of red paprika by means of chemometrics. For example, polyphenols and carbohydrates were employed as indicators of botanical and geographical origin of Serbian autochthonous clones of red spice paprika [16]. In relation with Spanish red paprika, a few works can be found dealing with the classification and authentication of this kind of samples. For example, geographical characterization of Spanish PDO paprika by multivariate analysis of multi-elemental content has been proposed [21]. Detection of adulterations in La Vera PDO paprika by free zone capillary electrophoresis (FZCE) [22,23] or by means of DNA typing methods [24] has also been described. However, to the best of our knowledge, phenolic profiling and HPLC-UV fingerprinting have not been yet described for that purpose.

In this work, a simple and feasible HPLC-UV method for the extraction of fingerprinting profiles based on phenolic compounds to achieve the classification of Spanish paprika samples was developed. For that purpose, a total of 17 phenolic compounds belonging to different families were selected and their separation by reversed-phase chromatography using a C₁₈ column optimized. A simple and cheap sample extraction procedure based on sonication and centrifugation was employed. Extracting solution was selected according to the obtention of discriminant HPLC-UV fingerprints among the analyzed paprika samples. The hypotheses is that the data corresponding to the phenolic composition as well as the HPLC-UV chromatographic fingerprints can be considered as a source of potential descriptors to be exploited for the classification and characterization of Spanish red paprika samples by exploratory principal component analysis (PCA) and linear discriminant analysis (LDA) methods.

2. Materials and Methods

2.1. Chemicals and Standard Solutions

All reagents employed, unless otherwise stated, were of analytical grade. Standards of arbutine, resveratrol, ethyl gallate, syringaldehyde, polydatin, tyrosol, umbelliferon, 4-hydroxybenzoic acid, caffeic acid, chlorogenic acid, gallic acid, homogentisic acid, *p*-coumaric acid, vanillic acid, ferulic acid, syringic acid, and *trans*-cinnamic acid were obtained from Sigma-Aldrich (Steinheim, Germany).

The structures and CAS numbers of all studied phenolic compounds are summarized in Table 1. Stock standard solutions of all phenolics (*ca.* 1000 mg/L) were prepared in methanol in amber glass vials. Intermediate working solutions were then prepared by appropriate dilution with Milli-Q water. All stock solutions were stored at 4 °C.

Methanol and acetonitrile (both UHPLC-gradient grade), absolute ethanol and acetone were purchased from Panreac (Barcelona, Spain). Water was purified using an Elix 3 coupled to a Milli-Q system (Millipore, Bedford, MA, USA) and filtered through a 0.22 µm nylon membrane integrated into the Milli-Q system.

2.2. Instrumentation

An Agilent 1200 Series HPLC instrument equipped with a G1311A quaternary pump, a G1322A degasser, a G1329A autosampler, and a G1314B ultraviolet-visible detector from Agilent Technologies (Palo Alto, CA, USA) was employed. Instrument was controlled with the Agilent ChemStation software. Chromatographic separation was performed in reversed-phase mode by using a Kinetex C₁₈ (100 × 4.6 mm i.d., 2.6 µm particle size) column from Phenomenex (Torrance, CA, USA) at room temperature. Gradient elution employing 0.1% (*v/v*) formic acid aqueous solution (solvent A) and methanol (solvent B) as mobile phase components was applied following the next elution program: 0–2 min, isocratic step at 5% B; 2–4 min linear gradient from 5 to 25% B; 4–12 min, at 25% B; 12–14 min, from 25 to 45% B; 14–16 min, at 45% B; 16–18 min, from 45 to 95% B; 18–20 min, at 95% B; 20–21 min, back to initial conditions at 5% B; and from 21–30 min, at 5% B for column re-equilibration. The mobile phase flow rate was 1 mL/min and the injection volume was 20 µL. Direct UV-absorption detection at 280 nm was employed.

2.3. Samples and Sample Treatment

Two different scenarios were considered when purchasing the different paprika samples, *viz.* its region of production and its type. All paprika samples considered in this study were produced in Spain, either in the La Vera or Murcia regions, which are the two main producing regions in Spain. Furthermore, for each of the regions, the main types of paprika were considered; namely spicy and sweet, plus the bittersweet produced in La Vera. Overall, 96 Spanish paprika samples from different brands, purchased from different local shops and producers, were analyzed. Among them, 72 paprika samples were from La Vera PDO and belonging to three different varieties: sweet, bittersweet, and spicy, with 26, 23, and 23 samples, respectively, and 24 were from Murcia PDO and belonging to two varieties: sweet and spicy, with 12 samples each.

Sample treatment was performed as follows: 0.3 g of paprika sample was dispersed in 3 mL of a water:acetonitrile (20:80 *v/v*) solution in a 15 mL PTFE tube. Next, samples were stirred in a Vortex for 1 min (Stuart, Stone, United Kingdom) and sonicated for 15 min (2510 Branson ultrasonic bath, Hampton, NH, USA). Finally, samples were centrifuged for 30 min at 4500 rpm (Rotana 460 HR centrifuge, Hettich, Germany), and the supernatant extract filtered through 0.45 µm nylon filters (Whatman, Clifton, NJ, USA) and stored at −18 °C in 2 mL glass injection vials until analysis.

2.4. Data Analysis

Chemometric analysis of the chromatographic data was done in Matlab 7.1 (MathWorks, Natick, MA, USA) employing basic scripts written by the authors. Chromatograms were first baseline corrected and then compressed employing fast Fourier transform (FFT) to reduce its dimensionality while ensuring the preservation of the chromatographic profiles [25]. Calculated FFT coefficients were then analyzed by means of principal component analysis (PCA) and linear discriminant analysis (LDA) to assess the feasibility of the approach for the classification of paprika samples. The former allowed to reveal trends in the chromatographic responses, whereas the latter allowed to actually build a classification model. See further details in the Supplementary information file.

In this way, data matrices to be treated by PCA/LDA consisted on the whole HPLC-UV chromatographic fingerprint obtained at 280 nm.

3. Results and Discussion

3.1. HPLC-UV Method

The main objective of the present work was to develop a simple and non-expensive HPLC-UV method for the classification, characterization and authentication of Spanish paprika samples that can be feasible and easily accessible for any food control laboratory. To this aim, the first step was to achieve the separation of major phenolic compounds present in paprika, so that afterwards, when processing the whole chromatogram, differences in the chromatographic profiles (i.e., peaks and areas) of the different compounds are identified.

For that purpose, a total of 17 phenolic compounds belonging to different families (Table 1) were selected as target analytes, based on major phenolic compounds already identified in paprika, [16, 26]. Their chromatographic separation was evaluated by reversed-phase chromatography using a porous-shell Kinetex C₁₈ column. As it is well described in the literature, reversed-phase separation of phenolic compounds can be addressed with acidified aqueous solutions and methanol or acetonitrile as organic components in the mobile phase [20]. Thus, as a first approach, the separation of the 17 studied phenolic compounds was attempted employing 0.1% formic acid aqueous solution and methanol as mobile phase components, and using a universal gradient elution profile (from 5 to 95% methanol in 20 min, and then back to initial conditions for column re-equilibration). Under these conditions, peak signals for all the studied compounds were detected although full baseline separation for all of them was not achieved, and numerous partial co-elutions were observed. Hence, chromatographic separation was optimized by applying different gradient elution profiles combining methanol isocratic elution and linear gradient elution steps. Optimal gradient conditions (see Instrumentation section) were selected as a compromise between the separation of the targeted phenolic compounds and total analysis time. Figure 1 shows the HPLC-UV chromatogram obtained at 280 nm when a standard solution of targeted phenolic compounds (each at 15 mg/L) was analyzed under the optimized gradient program. As can be seen, the most critical phenolic pairs to be separated were chlorogenic acid/vanillic acid (peaks 6 and 7, respectively), and syringaldehyde/ethyl gallate (peaks 10 and 11, respectively). Achieving baseline separation of all compounds implied to considerably increase the total analysis time. For this reason, 20 min were considered as an acceptable limit to address the separation of the targeted phenolic compounds, taking also into consideration that the total analysis time will be higher as the column has to re-equilibrate at initial gradient conditions. Under these conditions, an acceptable separation for all compounds except for the pair syringaldehyde and ethyl gallate was obtained within 19 min, with a total chromatographic elution profile of 30 min.

Table 1. Chemical structures and classification of the studied phenolic compounds.

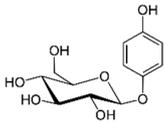
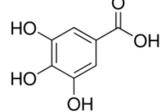
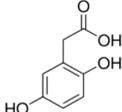
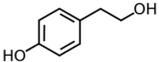
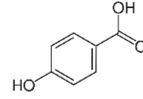
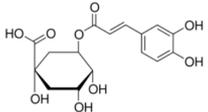
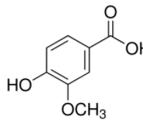
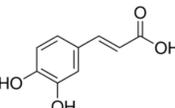
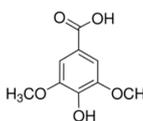
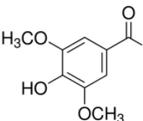
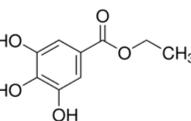
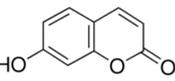
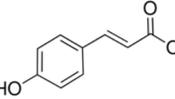
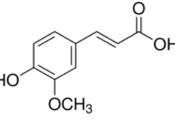
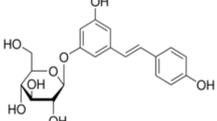
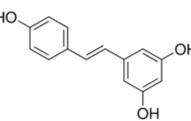
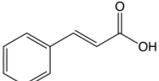
Peak	Phenolic Compound	Family	Structure	CAS Number
1	Arbutine	Phenolic glucoside		497-76-7
2	Gallic acid	Phenolic acid		149-91-7
3	Homogentisic acid	Phenolic acid		451-13-8

Table 1. Cont.

Peak	Phenolic Compound	Family	Structure	CAS Number
4	Tyrosol	Other phenolics		501-94-0
5	4-Hydroxybenzoic acid	Phenolic acid		99-96-7
6	Chlorogenic acid	Phenolic acid		327-97-9
7	Vanillic acid	Phenolic acid		121-34-6
8	Caffeic acid	Phenolic acid		331-39-5
9	Syringic acid	Phenolic acid		530-57-4
10	Syringaldehyde	Phenolic aldehyde		134-96-3
11	Ethyl gallate	Phenolic acid		831-61-8
12	Umbelliferon	Coumarin		93-35-6
13	<i>p</i> -Coumaric acid	Phenolic acid		501-98-4
14	Ferulic acid	Phenolic acid		537-98-4
15	Polydatin	Stilben		65914-17-2
16	Resveratrol	Stilben		501-36-0
17	<i>trans</i> -Cinnamic acid	Cinnamic acid		140-10-3

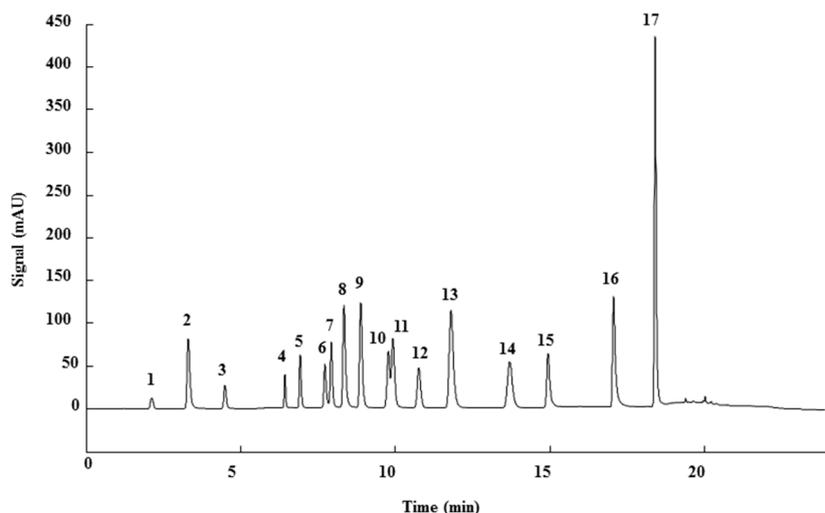


Figure 1. HPLC-UV chromatogram of a 15 mg/L standard solution of the 17 targeted phenolic compounds under optimal gradient elution conditions at 280 nm. Peak identification as in Table 1.

Instrumental quality parameters of the proposed HPLC-UV method were calculated for the 17 phenolic compounds studied, and the figures of merit are summarized in Table 2. Instrumental limits of detection (ILODs), based on a signal-to-noise ratio of 3:1, were calculated using standard solutions prepared in Milli-Q water at low concentration levels, and values between 12.5 µg/L (vanillic acid) and 38.0 µg/L (arbutine) were obtained. Instrumental limits of quantitation (ILOQs), based on a signal-to-noise ratio of 10:1, in the range of 44.0 to 126.7 µg/L were achieved. In general, these values are slightly lower than the ones achieved with photo-diode array (PDA) detectors for these kind of compounds [27–29], which it is expected due to the characteristic improvement in sensitivity when employing monochromatic-UV detectors.

External calibration curves based on peak area within the concentration range from ILOQ to 50 mg/L were established for each phenolic compound, and very good linearities, with correlation coefficients (r^2) values higher than 0.9996, were obtained. Sensitivity values, expressed as the slope of the calibration curves, between 6.29 (arbutine) and 138.66 (*trans*-cinnamic acid) were observed.

In order to evaluate the performance of the proposed HPLC-UV method both run-to-run (intra-day) and day-to-day (inter-day) precisions based on compound quantitation at several concentration levels (15, 2.5, 0.75, and 0.25 mg/L) were calculated and the results are also depicted in Table 2. With few variations depending on the concentration level evaluated, run-to-run precision values (calculated as percentage of relative standard deviation, RSD) were always lower than 0.7% (except for *trans*-cinnamic acid at 250 µg/L showing a RSD value of 2.7%). The dispersion increased for the day-to-day precision, as expected, with RSD values between 0.4 and 19.3%, and obtaining higher dispersion as the concentration level decreases.

Table 2. Instrumental quality parameters of the proposed HPLC-UV method.

Peak	Polyphenol	ILOD (µg/L)	ILOQ (µg/L)	Linearity (r ²)	Sensitivity	run-to-run precision (%RSD, n = 5)				day-to-day precision (%RSD, n = 3 × 5)				Trueness (% Error)			
						Level 1	Level 2	Level 3	Level 4	Level 1	Level 2	Level 3	Level 4	Level 1	Level 2	Level 3	Level 4
1	Arbutine	38.0	126.7	0.9996	6.29	0.1	0.1	0.1	0.4	0.4	3.9	1.4	19.3	0.5	0.7	3.1	7.9
2	Gallic acid	33.6	112.0	0.9997	36.70	0.1	0.05	0.4	0.3	0.4	2.7	2.1	10.7	1.0	0.7	1.6	2.5
3	Homogentisic acid	31.7	105.7	0.9996	11.50	0.1	0.2	0.4	0.4	1.8	4.5	9.1	17.2	1.0	0.3	0.4	5.1
4	Tyrosol	31.3	104.3	0.9996	9.07	0.1	0.1	0.1	0.7	2.1	3.3	4.7	9.1	1.0	0.2	5.3	13.3
5	4-Hydroxybenzoic acid	13.4	44.7	0.9996	18.05	0.04	0.1	0.1	0.2	1.8	3.3	3.7	10.0	1.0	1.3	4.7	4.3
6	Chlorogenic acid	33.3	111.0	0.9996	19.24	0.03	0.1	0.1	0.3	1.1	4.4	8.1	8.6	1.1	0.6	0.1	7.2
7	Vanillic acid	12.5	41.7	0.9999	26.39	0.1	0.1	0.2	0.3	0.8	3.2	3.3	13.5	0.04	1.0	4.6	1.0
8	Caffeic acid	13.2	44.0	0.9996	48.28	0.05	0.03	0.1	0.2	1.1	3.1	5.5	14.9	0.3	0.9	2.1	8.6
9	Syringic acid	12.8	42.7	0.9996	53.45	0.1	0.1	0.1	0.2	2.1	3.4	4.2	10.3	1.1	1.1	5.3	5.7
10	Syringaldehyde	31.0	103.3	0.9999	28.06	0.1	0.1	0.1	0.1	1.6	3.2	2.3	8.1	0.6	0.7	4.0	25.7
11	Ethyl gallate	13.0	43.3	0.9996	41.57	0.1	0.1	0.3	0.2	2.0	5.7	10.8	14.3	1.0	3.4	0.5	1.3
12	Umbelliferon	31.2	104.0	0.9996	25.50	0.1	0.03	0.1	0.5	1.1	3.8	5.9	17.7	0.3	0.2	2.4	8.9
13	<i>p</i> -Coumaric acid	14.3	47.7	0.9999	75.10	0.1	0.2	0.1	0.1	1.2	4.2	6.0	13.6	0.8	1.1	0.1	4.4
14	Ferulic acid	31.1	103.7	0.9999	46.79	0.1	0.04	0.1	0.2	1.4	3.9	6.3	14.9	0.9	0.9	1.6	5.2
15	Polydatin	33.0	110.7	0.9999	37.26	0.1	0.2	0.5	0.4	1.4	2.5	7.9	13.7	0.5	2.0	0.6	8.7
16	Resveratrol	31.1	103.7	0.9999	63.62	0.04	0.2	0.3	0.4	1.4	4.5	8.2	18.3	1.2	0.6	4.6	3.7
17	<i>trans</i> -Cinnamic acid	12.6	42.0	0.9999	138.66	0.1	0.1	0.4	2.7	3.3	2.7	5.6	8.1	1.4	0.5	8.3	29.2

Level 1 = 15 mg/L; Level 2 = 2.5 mg/L; Level 3 = 750 µg/L; Level 4 = 250 µg/L.

Because of the lack of reference materials containing the 17 targeted phenolic compounds, intra-day (within the same day) method trueness was evaluated by comparing spiked phenolic compounds concentrations with the calculated ones using external calibration at the four concentration levels previously described, and the results expressed as the relative errors (%) are shown in Table 2. With some variations depending on the concentration level, the method trueness was, in general, very good with relative errors in the range from 0.04 to 8.9%, except for three compounds at the lowest concentration level (tyrosol, syringaldehyde, and *trans*-cinnamic acid) showing higher errors (13.3–29.2%).

The results obtained showed that the proposed HPLC-UV method is, in general, very satisfactory in terms of sensitivity, precision and trueness for the determination of phenolic compounds.

3.2. Sample Extraction Optimization

The applicability of the proposed HPLC-UV method to obtain discriminant UV fingerprinting profiles for the classification and characterization of Spanish paprika samples was evaluated. For that purpose, a simple sample treatment procedure consisting in the extraction of 0.3 g of paprika sample with 3 mL of an extracting solvent by stirring, sonication and centrifugation was proposed. Several extraction solutions including 100% water, methanol, acetonitrile, ethanol and acetone, as well as the mixtures of water with the other organic solvents at 50:50, 20:80, and 80:20 *v/v* ratios, were employed. As a preliminary study, the proposed sample treatment using the different extracting solutions was applied to three paprika samples from La Vera PDO belonging to the same producer, but of different variety (sweet, bittersweet, and spicy). The obtained HPLC-UV chromatograms at 280 nm were compared in terms of signal abundance, signal profile, number of peaks detected, etc., in order to achieve discriminant fingerprints. The best results were obtained when the samples were extracted with a water:acetonitrile 20:80 *v/v* solution (see HPLC-UV fingerprints in Figure 2). As can be seen, noticeable differences in signal profiles and intensities were obtained in the time segments from 5 to 10 min and from 15 to 23 min, as well as a characteristic peak eluting close to 21 min whose intensity is clearly different depending on the paprika variety. A priori, we considered that these extracting conditions were appropriate to obtain sufficient discriminant HPLC-UV fingerprints for the characterization and classification of paprika samples. Hence, this extracting solution was selected, the proposed method applied to all the Spanish paprika samples, and the obtained HPLC-UV fingerprints subjected to chemometric analysis by PCA and LDA.

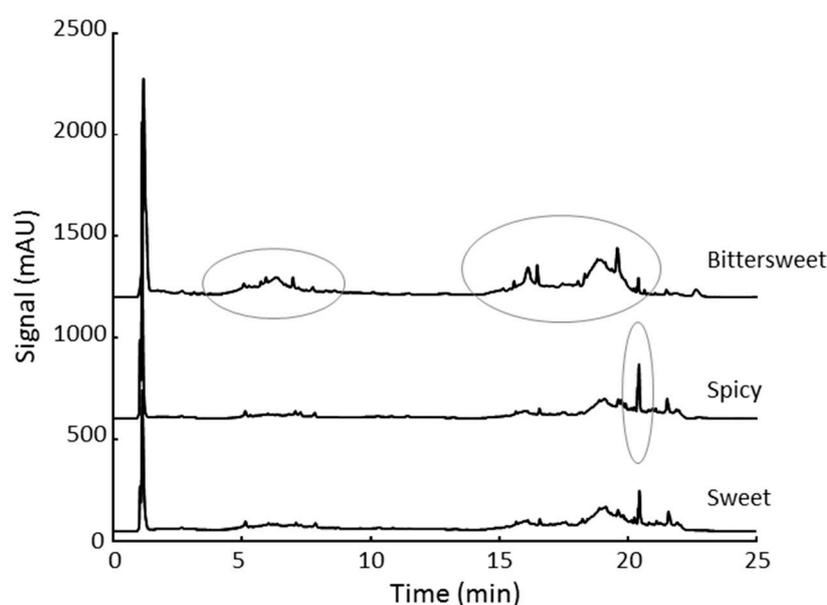


Figure 2. Representative chromatograms obtained for certain arbitrary paprika samples extracts.

3.3. Qualitative Analysis

As already stated, our aim was to assess whether or not we can identify a pattern in the chromatographic profile that allows the classification of paprika samples based on its region and type. However, rather than trying to match each of the peaks observed on the chromatogram to a specific compound and analyze their areas, we aim to process the whole chromatogram (after its compression with FFT) as a unique profile and analyze it with the aid of pattern recognition methods such as PCA and LDA. Such a holistic approach based on the analysis of the patterns of several compounds has higher potential to address new challenges in food authentication than traditional quality control strategies focused on the analysis of a few compounds as there is not a unique marker that can be related to e.g., food origin.

Thus, the first step was the compression of the signals by means of FFT (see Supplementary information for further details). In this way, each chromatogram was reduced from 12,325 points down to 512 coefficients, without any loss of significant information (Figure S1). This was assessed by calculating both the correlation of determination (R^2) and a comparison factor f_c between the original and the reconstructed signal after doing the inverse of the FFT [25], with values higher than 0.99 for both of them.

Next, PCA was used to assess initial patterns in the data as by plotting the PCs it provides a better representation of samples (dis)similarities, without taking into account prior expected similarities, but based only on their variance (Figure S2). With only the first two PCs, the cumulative variance was ca. 46.0%; meaning that almost half of the initial variance in the original data is now summarized on the plot shown in Figure 3.

Although not completely separated clusters were obtained for all the classes that we initially considered, we can still see some significant trends, e.g., the cluster of samples belonging to the Murcia spicy group. In order to further confirm these patterns, 95% confidence ellipses for each of the classes were also calculated and plotted [30], showing quite promising results taking into account that PCA is an unsupervised method that does not seek samples classification.

The next step was to build an actual model using a supervised method that allows the classification of samples into the different classes, to which aim LDA was the chosen tool (Figure S2). As before, chromatograms were baseline corrected and compressed with FFT, but in this case, calculated coefficients were then pruned employing a stepwise inclusion method in order to remove the less-significant coefficients [31]. The scores plot is shown in Figure 4, where clear clusters for each of the classes were obtained.

To numerically confirm the goodness of the classification model, the set of samples was randomly split into two subsets: training (71 samples) and testing (25 samples, 5 of each class). The former was used to build the classification model, providing the system with responses and target matrices to adjust the model parameters; whereas the latter was used to assess its predictive ability, providing the system only with responses matrix as input data to be interpolated in the built model and calculate the predicted targets matrix. This allows to obtain more unbiased metrics since the testing subset is not used at all during the modelling stage, and can therefore be considered as a blind set of samples, providing a measure of the actual performance of the classification model. Predicted and actual classes for each of the samples were compared, the confusion matrix was built (Table S1) and performance metrics were calculated, achieving 100% for classification rate, sensitivity and specificity [32]. Values that confirm the potential of HPLC-UV fingerprinting for the authentication of Spanish paprika samples, confirming the hypotheses established in the introduction section.

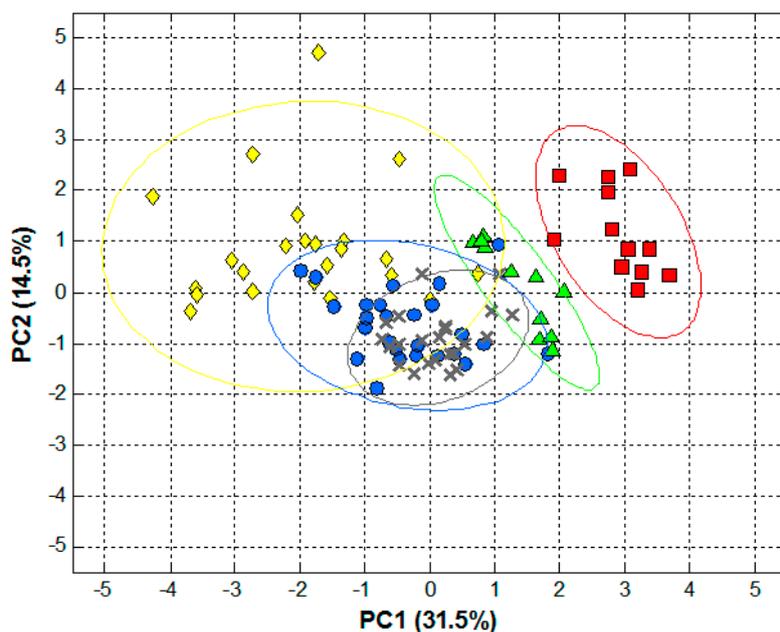


Figure 3. Score plot obtained after PCA analysis of the paprika samples chromatograms: (green \blacktriangle) Murcia sweet, (red \blacksquare) Murcia spicy, (blue \bullet) La Vera sweet, (yellow \blacklozenge) La Vera spicy, and (grey \times) La Vera bittersweet. Ellipses plotted correspond to 95% confidence limits for each of the groups.

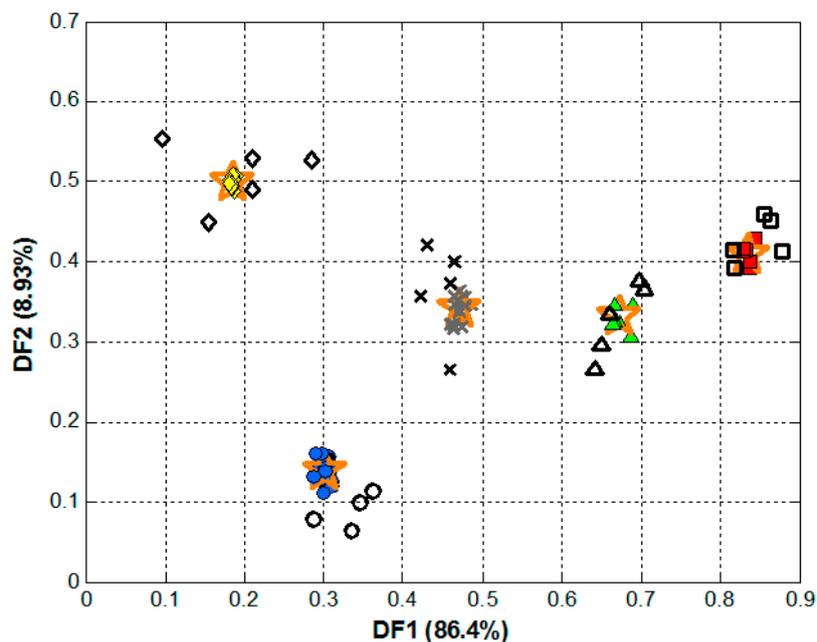


Figure 4. Score plot obtained after LDA analysis of the paprika samples chromatograms: (green \blacktriangle) Murcia sweet, (red \blacksquare) Murcia spicy, (blue \bullet) La Vera sweet, (yellow \blacklozenge) La Vera spicy, and (grey \times) La Vera bittersweet. Additionally, the centroid for each of the classes is also plotted (\star). Colored, filled symbols correspond to the samples of the training subset, black empty ones to the testing subset.

4. Conclusions

The combination of liquid chromatography with chemometric methods has been proven to be a suitable approach for the authentication of paprika samples, providing an analytical tool able to overcome the lack of knowledge of unique characteristic compounds that can be used to achieve such discrimination.

Extraction of phenolic compounds present in paprika was achieved by sonication/centrifugation, and its determination was carried out in reversed-phase chromatography using a C₁₈ column. The developed method showed very good linearities, with slightly lower LOD and LOQ values than the ones reported with PDA detection. Run-to-run (intra-day) precisions, calculated as percentage of RSD were lower than 0.5% at 750 µg/L, whereas day-to-day precisions were around 2–10%. For the authentication of paprika samples, the registered chromatograms were compressed employing FFT and qualitative models built employing LDA, achieving the correct classification of all samples.

Overall, the main advantages of our approach are its simplicity, versatility, low-cost, and short-time analysis, all of them characteristics required to achieve the engagement of the stakeholders. Moreover, the approach developed herein can also be adapted to the discrimination of other solid food samples given the great relevance of phenolic compounds in plant products (e.g., fruits, vegetables, cereals, tea, or wine). In such cases, same extracting procedure and chromatographic conditions could be used as starting point given same target analytes would be considered. However, optimization of the chromatographic conditions, plus the analysis of a proper set of samples to build the chemometric model would still be required.

Supplementary Materials: This supplementary provides further description of the fingerprinting approach, namely about the signal compression method used (fast Fourier transform, FFT) and the chemometric analysis (both principal component analysis (PCA) and linear discriminant analysis (LDA)). The following are available online at <http://www.mdpi.com/1424-8220/18/12/4479/s1>, Figure S1: Selection of the optimal number of Fourier coefficients for signal compression. Changes in (A) chromatogram reconstruction and (B) R² and *f*_c coefficients depending on the number of coefficients considered; Figure S2. Schematic comparison between PCA and LDA. Briefly, PCA seeks for the direction of maximum variance between the variables, while LDA seeks for the direction that provides maximum separation between the classes. Note also that in PCA as many new PCs as variables in the original data are obtained, whereas in LDA only as many #groups-1 DFs are obtained. Adapted from Cetó, X., et al. *Electroanalysis*, 25 (2013), 1635; Table S1. Confusion matrix built according to the classes assigned by the LDA model to the samples of the testing subset.

Author Contributions: N.S. and O.N. conceived and designed the experiments. M.A. and A.G. performed the experiments. X.C. processed the data. N.S., O.N., X.C., and J.M.D.-C. discussed the first results and suggested additional experiments. The first draft of the manuscript was prepared by O.N., X.C. and N.S. and it was revised and substantially improved by J.M.D.-C. and M.E.

Acknowledgments: The authors gratefully acknowledge the financial support received from the Spanish Ministry of Economy and Competitiveness under the project CTQ2015-63968-C2-1-P and from the Agency for Administration of University and Research Grants (Generalitat de Catalunya, Spain) under the projects 2017SGR-311 and 2017SGR-310. X. Cetó thanks the support of the Beatriu de Pinós postdoctoral program of the Government of Catalonia's Secretariat for Universities and Research of the Ministry of Economy and Knowledge.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Govindarajan, V.S.; Sathyanarayana, N.M. Capsicum—Production, technology, chemistry, and quality. Part V. Impact on physiology, pharmacology, nutrition, and metabolism; structure, pungency, pain, and desensitization sequences. *Crit. Rev. Food Sci. Nutr.* **1991**, *29*, 435–474. [[CrossRef](#)] [[PubMed](#)]
2. Surh, Y.J.; Lee, S.S. Capsaicin in hot chili pepper: Carcinogen, co-carcinogen or anticarcinogen? *Food Chem. Toxicol.* **1996**, *34*, 313–316. [[CrossRef](#)]
3. Márkus, F.; Daood, H.G.; Kapitány, J.; Biacs, P.A. Change in the carotenoid and antioxidant content of spice red pepper (Paprika) as a function of ripening and some technological factors. *J. Agric. Food Chem.* **1999**, *47*, 100–107. [[CrossRef](#)] [[PubMed](#)]
4. Hornero-Méndez, D.; Gómez-Ladrón de Guevara, R.; Mínguez-Mosquera, M.I. Carotenoid Biosynthesis Changes in Five Red Pepper (*Capsicum annuum* L.) Cultivars during Ripening. Cultivar Selection for Breeding. *J. Agric. Food Chem.* **2000**, *48*, 3857–3864. [[CrossRef](#)] [[PubMed](#)]
5. Pol, L.C.; Siddiq, M.; Shahzad, T. Chili, Peppers, and Paprika. In *Handbook of Vegetables and Vegetable Processing*; Siddiq, M., Uebersax, M.A., Eds.; John Wiley & Sons: Hoboken NJ, USA, 2018; pp. 633–660.

6. Anuario de Estadística 2016. Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente, Gobierno de España, Madrid 2017. Available online: <http://www.mapama.gob.es/estadistica/pags/anuario/2016/AE16.pdf> (accessed on 17 February 2017).
7. Manach, C.; Scalbert, A.; Morand, C.; Rémésy, C.; Jiménez, L. Polyphenols—Food Sources and Bioavailability. *Am. J. Clin. Nutr.* **2004**, *79*, 727–747. [[CrossRef](#)] [[PubMed](#)]
8. Scalbert, A.; Manach, C.; Morand, C.; Rémésy, C.; Jiménez, L. Dietary polyphenols and the prevention of diseases. *Crit. Rev. Food Sci. Nutr.* **2005**, *45*, 287–306. [[CrossRef](#)] [[PubMed](#)]
9. Choi, D.Y.; Lee, Y.J.; Hong, J.T.; Lee, H.J. Antioxidant properties of natural polyphenols and their therapeutic potentials for Alzheimer’s disease. *Brain Res. Bull.* **2012**, *87*, 144–153. [[CrossRef](#)] [[PubMed](#)]
10. Tomé-Carneiro, J.; Visioli, F. Polyphenol-based nutraceuticals for the prevention and treatment of cardiovascular disease: Review of human evidence. *Phytomedicine* **2016**, *23*, 1145–1174. [[CrossRef](#)] [[PubMed](#)]
11. Hidalgo-Serrano, M.; Núñez, O.; Saurina, J. Chemometrics in the polyphenolic characterization of natural products by liquid chromatography and capillary electrophoresis. In *Chemometrics. Methods, Applications and New Research*; Luna, A.S., Ed.; Nova Science: New York, NY, USA, 2017.
12. Saurina, J.; Sentellas, S. Determination of Phenolic Compounds in Food Matrices: Applications to Characterization and Authentication. In *Fast Liquid Chromatography-Mass Spectrometry Methods in Food and Environmental Analysis*; Núñez, O., Gallart-Ayala, H., Eds.; Imperial College Press: London, UK, 2015.
13. Nagy, Z.; Daoud, H.; Ambrózy, Z.; Helyes, L. Determination of Polyphenols, Capsaicinoids, and Vitamin C in New Hybrids of Chili Peppers. *J. Anal. Methods Chem.* **2015**, *2015*, 1–10. [[CrossRef](#)] [[PubMed](#)]
14. Silva, C.L.; Haesen, N.; Câmara, J.S. A new and improved strategy combining a dispersive-solid phase extraction-based multiclass method with ultra high pressure liquid chromatography for analysis of low molecular weight polyphenols in vegetables. *J. Chromatogr. A* **2012**, *1260*, 154–163. [[CrossRef](#)] [[PubMed](#)]
15. Navarro, M.; Núñez, O.; Saurina, J.; Hernández-Cassou, S.; Puignou, L. Characterization of fruit products by capillary zone electrophoresis and liquid chromatography using the compositional profiles of polyphenols: Application to authentication of natural extracts. *J. Agric. Food Chem.* **2014**, *62*, 1038–1046. [[CrossRef](#)] [[PubMed](#)]
16. Mudric, S.Z.; Gašić, U.M.; Dramicanin, A.M.; Ćirić, I.Z.; Milojković-Opsenica, D.M.; Popović-Dorđević, J.B.; Momirović, N.M.; Tešić, Z.L. The polyphenolics and carbohydrates as indicators of botanical and geographical origin of Serbian autochthonous clones of red spice paprika. *Food Chem.* **2017**, *217*, 705–715. [[CrossRef](#)] [[PubMed](#)]
17. Martí, R.; Valcárcel, M.; Herrero-Martínez, J.M.; Cebolla-Cornejo, J.; Roselló, S. Fast simultaneous determination of prominent polyphenols in vegetables and fruits by reversed phase liquid chromatography using a fused-core column. *Food Chem.* **2015**, *169*, 169–179. [[CrossRef](#)] [[PubMed](#)]
18. Puigventós, L.; Navarro, M.; Alechaga, É.; Núñez, O.; Saurina, J.; Hernández-Cassou, S.; Puignou, L. Determination of polyphenolic profiles by liquid chromatography-electrospray-tandem mass spectrometry for the authentication of fruit extracts. *Anal. Bioanal. Chem.* **2015**, *407*, 597–608. [[CrossRef](#)] [[PubMed](#)]
19. Morales-Soto, A.; Gómez-Caravaca, A.M.; García-Salas, P.; Segura-Carretero, A.; Fernández-Gutiérrez, A. High-performance liquid chromatography coupled to diode array and electrospray time-of-flight mass spectrometry detectors for a comprehensive characterization of phenolic and other polar compounds in three pepper (*Capsicum annuum* L.) samples. *Food Res. Int.* **2013**, *51*, 977–984. [[CrossRef](#)]
20. Lucci, P.; Saurina, J.; Núñez, O. Trends in LC-MS and LC-HRMS analysis and characterization of polyphenols in food. *TrAC Trends Anal. Chem.* **2017**, *88*, 1–24. [[CrossRef](#)]
21. Palacios-Morillo, A.; Jurado, J.M.; Alcázar, Á.; De Pablos, F. Geographical characterization of Spanish PDO paprika by multivariate analysis of multielemental content. *Talanta* **2014**, *128*, 15–22. [[CrossRef](#)] [[PubMed](#)]
22. Hernández, A.; Martín, A.; Aranda, E.; Bartolomé, T.; Córdoba, M.G. Detection of Smoked Paprika “Pimentón de La Vera” adulteration by Free Zone Capillary Electrophoresis (FZCE). *J. Agric. Food Chem.* **2006**, *54*, 4141–4147. [[CrossRef](#)] [[PubMed](#)]
23. Hernández, A.; Martín, A.; Aranda, E.; Bartolomé, T.; de Córdoba, M.G. Application of temperature-induced phase partition of proteins for the detection of smoked paprika adulteration by free zone capillary electrophoresis (FZCE). *Food Chem.* **2007**, *105*, 1219–1227. [[CrossRef](#)]
24. Hernández, A.; Aranda, E.; Martín, A.; Benito, M.J.; Bartolomé, T.; Córdoba, M.D.G. Efficiency of DNA typing methods for detection of smoked paprika “pimenton de la Vera” adulteration used in the elaboration of dry-cured iberian pork sausages. *J. Agric. Food Chem.* **2010**, *58*, 11688–11694. [[CrossRef](#)] [[PubMed](#)]

25. Cetó, X.; Céspedes, F.; del Valle, M. Comparison of methods for the processing of voltammetric electronic tongues data. *Microchim. Acta* **2013**, *180*, 319–330. [[CrossRef](#)]
26. Ciulu-Costinescu, F.; Chifiriuc, M.C.; Popa, M.; Bleotu, C.; Neamtu, J.; Averis, L.M.E.; Bubulica, M.V.; Simionescu, A.; Aldea, I.M.; Belu, I. Screening of Polyphenol Content and In vitro Studies of Antioxidant, Antibacterial and Cytotoxic Activities of Capsicum Annuum Extracts. *Rev. Chim.* **2015**, *66*, 1261–1266.
27. Pardo-Mates, N.; Vera, A.; Barbosa, S.; Hidalgo-Serrano, M.; Núñez, O.; Saurina, J.; Hernández-Cassou, S.; Puignou, L. Characterization, classification and authentication of fruit-based extracts by means of HPLC-UV chromatographic fingerprints, polyphenolic profiles and chemometric methods. *Food Chem.* **2017**, *221*, 29–38. [[CrossRef](#)] [[PubMed](#)]
28. Larrauri, A.; Núñez, O.; Hernández-Cassou, S.; Saurina, J. Determination of polyphenols in white wines by liquid chromatography: Application to the characterization of alella (Catalonia, Spain) wines using chemometric methods. *J. AOAC Int.* **2017**, *100*, 323–329. [[CrossRef](#)] [[PubMed](#)]
29. Farrés-Cebrián, M.; Seró, R.; Saurina, J.; Núñez, O. HPLC-UV Polyphenolic Profiles in the Classification of Olive Oils and Other Vegetable Oils via Principal Component Analysis. *Separations* **2016**, *3*, 33. [[CrossRef](#)]
30. Wagner, M.S.; Castner, D.G. Characterization of adsorbed protein films by time-of-flight secondary ion mass spectrometry with principal component analysis. *Langmuir* **2001**, *17*, 4649–4660. [[CrossRef](#)]
31. Johnson, R.A.; Wichain, D.W. *Applied Multivariate Statistical Analysis*; Pearson International Edition: Upper Saddle River, NJ, USA, 2007.
32. Cetó, X.; Voelcker, N.H.; Prieto-Simón, B. Bioelectronic tongues: New trends and applications in water and food analysis. *Biosens. Bioelectron.* **2016**, *79*, 608–626. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).