JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Chemometric Characterization of Fruit Juices from Spanish Cultivars According to Their Phenolic Compound Contents: I. Citrus Fruits

Beatriz Abad-García,[†] Luis A. Berrueta,*[,†] Sergio Garmón-Lobato,[†] Arantza Urkaregi,[‡] Blanca Gallo,[†] and Francisca Vicente[†]

[†]Departamento de Química Analítica, Facultad de Ciencia y Tecnología, Universidad del País Vasco/Euskal Herriko Unibertsitatea, P.O. Box 644, 48080 Bilbao, Spain

[‡]Departamento de Matemática Aplicada, Estadística e Investigación Operativa, Universidad del País Vasco/Euskal Herriko Unibertsitatea, P.O. Box 644, 48080 Bilbao, Spain

Supporting Information

ABSTRACT: The data set composed by phenolic compound profiles of 83 *Citrus* juices (determined by HPLC-DAD-MS/MS) was evaluated by chemometrics to differentiate them according to *Citrus* species (sweet orange, tangerine, lemon, and grapefruit). Cluster analysis (CA) and principal component analysis (PCA) showed natural sample grouping among *Citrus* species and even the *Citrus* subclass. Most of the information contained in the full data set can be captured if only 15 phenolic compounds (concentration ≥10 mg/L), which can be quantified with fast and accurate methods in real samples, are introduced in the models; a good classification which allows the confirmation of the authenticity of juices is achieved by linear discriminant analysis. Using this reduced data set, fast and routine methods have been developed for predicting the percentage of grapefruit in adulterated sweet orange juices using principal component regression (PCR) and partial least-squares regression (PLS). The PLS model has provided suitable estimation errors.

KEYWORDS: *phenolic compounds, polyphenols, flavonoids, citrus, orange, tangerine, lemon, grapefruit, juice, chemometrics, pattern recognition, regression, PCA, PCR, PLS*

INTRODUCTION

Phenolic compounds, widely distributed in fruits, are very suitable as chemotaxonomic markers. Some of them are characteristic of some species or varieties, whereas quantitative differences may occur depending on fruit variety, stages of maturity, environmental conditions during growth, storage conditions,[1,2] postharvest treatments,[3] the presence of the peel in fruit-based products[4,5] and the extraction system.[6] For certain fruits, characteristic phenolic compounds have been successfully used for the determination of adulteration of fruit juices,[7,8] nectars,[9,10] and jams[11,12] with cheaper fruits.

Moreover, phenolic compounds have great importance in the nutritional, organoleptic, and commercial properties of fruits and their derivated products through their contributions to sensory attributes of fruits (color, sweet taste, bitterness, and astringency).[13,14] Other important aspects of phenolic compounds are the positive health benefits to humans.[15,16] Epidemiological studies have shown an inverse relationship between the intake of fruits, vegetables, and their products, rich in phenolic compounds, and these chemoprotective effects.[17,18]

Thus, knowledge of the precise composition of *Citrus* fruit cultivars and their products such as fruit juices, whose consumption has increased significantly in the last few years, may contribute to a better understanding of their influence on the quality and biological properties of these products/foods.[19] For this reason, characterization studies based on the phenolic profiles have been carried out with *Citrus* fruit and *Citrus* juice.[20−25] The phenolic composition of *Citrus* juice comprises flavanones (major group), flavones, and flavonols.[21,22,25−27] These flavonoids, found in different parts of *Citrus* fruits, usually occur as glycosides. Polymethoxylated flavones have been also found in large amounts in the peel of some *Citrus*.[28]

Several publications have detailed improvements in *Citrus* phenolic compound determination, especially using HPLC in conjunction with diode array detection for their identification and characterization.[8,21] Liquid chromatography coupled to electrospray ionization and tandem mass spectrometry (LC-ESI-MS/MS) is one of the most successful techniques applied to qualitative and quantitative determination of phenolic compounds in fruits. Its superior sensitivity, high selectivity, and resolution power allow direct screening of natural products, avoiding the previous need for laborious isolation of phenolics.[28,29] Several studies have reported phenolic compound identification in *Citrus* fruit,[25,30,31] apples,[32] and tomato[33] etc. by applying this technique.

These modern analytical instruments allow the production of great amounts of information (variables or features) for a large number of samples (objects) that are analyzed in a relatively short time. This leads to multivariate data matrices that require the use of mathematical and statistical procedures in order to efficiently extract the maximum useful information from data.[34]

In this sense, the need for guaranteeing food authenticity, demanded by food producers, consumers, and regulatory

Received: April 6, 2011
Revised: March 16, 2012
Accepted: March 19, 2012
Published: March 19, 2012

bodies, requires methods not only based on chemical analysis but also on sophisticated data analysis procedures as a relevant quality criterion. Statistical analyses are commonly used in order to develop systems for the determination of geographical origin or quality brand of foodstuffs and fraud detection.[35] Twomey et al. showed the potential of NIR spectroscopy and statistical analysis for the detection of the adulteration of orange juice.[36]

In this study, phenolic profiles of *Citrus* juices were quantified by LC-DAD. These profiles are representative of the Spanish *Citrus* fruit juice production and were analyzed by chemometric techniques with the aim of differentiating juices according to species of *Citrus* used for their elaboration: sweet orange, tangerine, lemon, and grapefruit. A classification system was developed in order to confirm the authenticity of juices. Principal component regression (PCR) and partial least-squares (PLS) were used with a high number of samples to obtain a prediction model. The ultimate objective was to predict the adulteration percentage in sweet orange juices to which grapefruit was added, which is a cheaper fruit than sweet orange.

MATERIALS AND METHODS

Reagents and Standards. Methanol and dimethyl sulfoxide (Romil, Chemical Ltd., Heidelberg, Germany) were of HPLC grade. Water was purified on a Milli-Q system (Millipore, Bedford, MA, USA). Glacial acetic acid, ascorbic acid, and sodium fluoride provided by Merck (Darmstadt, Germany) were of analytical quality. All solvents used were previously filtered through 0.45 μm nylon membranes (Lida, Kenosha, WI, USA).

Phenolics standards were supplied as follows: eriodictyol-7-*O*-rutinoside, eriodictyol-7-*O*-neohesperidoside, naringenin-7-*O*-rutinoside, hesperetin-7-*O*-rutinoside, hesperetin-7-*O*-neohesperidoside, isosakuranetin-7-*O*-rutinoside, hesperetin, homoeriodictyol, ferulic acid, sinapic acid, quercetin-3-*O*-galactoside, quercetin-3-*O*-glucofuranoside, quercetin-3-*O*-glucopyranoside, quercetin-3-*O*-rhamnoside, kaempferol-3-*O*-glucoside, kaempferol-3-*O*-rutinoside, kaempferol-7-*O*-neohesperidoside, kaempferol-3-*O*-robinoside-7-*O*-rhamnoside, isorhamnetin-3-*O*-glucoside, isorhamnetin-3-*O*-rutinoside, isorhamnetin, tamarixetin, myricetin, scopoletin, luteolin-7-*O*-glucoside, luteolin-6-*C*-glucoside, luteolin-8-*C*-glucoside, luteolin-3′,7-di-*O*-glucoside, luteolin-4′-*O*-glucoside, diosmetin-7-*O*-rutinoside, apigenin-7-*O*-glucoside, apigenin-6-*C*-glucoside, apigenin-8-*C*-glucoside, apigenin-7-*O*-neohesperidoside, apigenin-7-*O*-rutenoside, diosmetin, chrysoeriol, and sinensetin were from Extrasynthèse (Genay, France); while naringenin, 5′-caffeoylquinic acid, caffeic acid, *p*-coumaric acid, and quercetin-3-*O*-rutinoside were provided by Sigma-Aldrich Chemie (Steinheim, Germany); apigenin-8-*C*-glucoside-4′-*O*-rhamnoside, kaempferol-3-*O*-(*p*-coumaroyl)glucoside, tangeretin, and nobiletin were by Chromadex (Santa Ana, CA, USA); and naringenin-7-*O*-neohesperidoside, dehydrated quercetin, and apigenin were by Fluka Chemie (Steinheim, Germany).

All stock standard solutions (in concentrations ranging from 250 to 2500 μg/mL, depending on each phenolic compound) were prepared in methanol, except for hesperetin-7-*O*-rutinoside, hesperetin, homoeriodictyol, chrysoeriol, and isorhamnetin which were dissolved with water−dimethyl sulfoxide (80:20, v/v), and all were stored at 4 °C in darkness.

Fruit Samples. *Citrus* fruit of different Spanish cultivars from 2003 to 2004 and 2004−2005 harvests used in Spain for making juices were analyzed (Table 1). These *Citrus* fruits were purchased from a local market at maturity and were as follows: sweet orange *Citrus sinensis*, cv. Navel-Late (NVLA), cv. Navelina (NVL), cv. Navel (NV), cv. Salustiana (SA), and cv. Valencia Late (VL); tangerine *Citrus reticulate* and *Citrus unshiu*, cv. Hernandina (CLH), cv. Marisol (CLM), cv. Clemenule (CLN), cv. Clementina (CL), cv. Satsuma (SAT), cv. Fortuna (FOR), and cv. Clemenvilla (CLV); lemon *Citrus limon*, cv. Verna (V) and cv. Primafiori (VP); grapefruit *Citrus paradisi*, cv. Star Ruby (SR), cv. Red Ruby (RR), and cv. Blanco (BL).

Table 1. *Citrus* Cultivars Used for Elaboration Juices

				harvest	
fruit	subclass	cultivar	abbrev.	03/04	04/05
sweet orange	Navel	Navel-Late	NVLA	x	x
		Navelina	NVL	x	x
		Navel	NV	x	
	Blanca	Valencia Late	VL	x	x
		Salustiana	SA	x	x
tangerine	Clementina	Clementina Hernandina	CLH	x	x
		Clementina	CL	x	x
		Clemenule	CLN	x	x
		Clementina Marisol	CLM		x
	Satsuma	Satsuma	SAT	x	x
	Hybrids	Fortuna	FOR	x	
		Clemenvilla	CLV		x
lemon	Verna	Verna	V	x	x
		Primafiori	VP	x	x
grapefruit	Pigmented	Star Ruby	SR	x	x
		Red Ruby	RR	x	x
	Blanca	Blanco	BL		x

***Citrus* Juice Preparation.** Three batches of fruit (1 kg) were constituted for each fruit cultivar and harvest. Each batch was peeled separating the flavedo and the albedo from the pulp and squeezed using a home juicer. Despite the fact that this extraction procedure is not used in an industrial scale by fruit juice manufacturers, it is widely used by small manufacturers and allows a suitable control of the conditions and fruits from which the juice is extracted. The collected juice, after measuring its volume, was mixed with 50 mL of an aqueous solution containing 0.2 g/mL ascorbic acid and 0.2 g/mL sodium fluoride, in order to inactive polyphenoloxidases and prevent phenolic degradation,[37] and centrifuged at 6000 rpm for 15 min at 4 °C. Aliquots of 1 mL were sampled, stored at −20 °C, and lyophilized later. The freeze-dried material was stored at room temperature in a desiccator in darkness until analysis.

Analytical Procedure. *Solvent Extraction of Freeze-Dried Samples.* A 1 mL aliquot of this juice was freeze-dried for preservation and extracted at the time of analysis with 2 mL of a mixture of methanol−water−acetic acid 30:69:1 (v/v/v) using ascorbic acid as the preservative (2 g/L). Mixing was carried out by vortexing, and the extraction was performed in an ultrasonic bath for 15 min at room temperature. The extract was centrifuged at 4000 rpm for 4 min and passed through a 0.45um PTFE filter (Waters, Milford, USA) prior to its injection into the chromatographic system. This solvent extraction procedure of freeze-dried aliquots of fruit juices followed by the analysis of phenolic compounds by reversed-phase high-performance liquid chromatography with photodiode array detection was optimized and validated in a previous work.[38]

Reversed-Phase HPLC Analysis. Chromatographic analysis was performed on a Shimadzu (Kyoto, Japan) liquid chromatograph, equipped with a vacuum degasser DGU-14A, a quaternary pump LC-10DVP, a thermostatted autosampler SIL-10ADVP, a thermostatted column compartment, and a DAD detector SPD-M10AVP, and controlled by CLASS-VP software. A reversed-phase Phenomenex (Torrance, USA) Luna C18(2) column (150 × 4.6 mm i.d. and 3 μm particle size) with a Waters Nova-Pack C18 guard column (10 × 3.9 mm i.d, 4 μm) was used. A gradient program for general phenolic compound analysis[38] was employed: the eluents were acetic acid−water (0.5:99.5, v/v) (phase A) and methanol (phase B); initially, 0% B for 2 min, a linear gradient to 15% B at 6 min, held isocratic until 12 min, linear gradient to 20% B at 15 min, 20% B constant until 35 min,

Table 2. Concentrations (mg/L) of Flavanones Present in *Citrus* Juices

	concentrations (mg/L) (mean ± SD) (min-max)			
flavanones	orange	tangerine	lemon	grapefruit
eri-7-*O*-rut-4′-*O*-glc	nd	nd	19 ± 8 (9−29)	nd
nar-7-*O*-rut-4′-*O*-glc	39 ± 13 (19−74)	11 ± 11 (0.9−30)	nd	14 ± 2 (9−17)
nar-7-*O*-nhes-4′-*O*-glc	nd	nd	nd	11 ± 2 (7−14)
nar-*O*-hexhex	nd	nd	nd	5.3 ± 0.9 (4−7)
eri-7-*O*-rut	8 ± 3 (2−13)	5 ± 2 (3−9)	341 ± 140 (163−538)	nd
nar-7-*O*-rut	139 ± 47 (66−237)	90 ± 72 (13−203)	9 ± 7 (3-21)	273 ± 64 (164−381)
nar-7-*O*-nhes	nd	nd	nd	1064 ± 300 (652−1472)
hes-7-*O*-rut	642 ± 271 (203−1074)	549−289 (144−984)	536 ± 234 (266−900)	33 ± 3 (24−37)
hes-7-*O*-nhes	nd	nd	nd	62 ± 12 (42−80)
isk-7-*O*-rut	46 ± 19 (17−75)	42 ± 58 (1−174)	4 ± 4 (0.006−9)	6 ± 8 (0.006−21)
nar-*O*-rhamalonylhex	nd	nd	nd	17 ± 7 (4−26)
isk-7-*O*-nhes	nd	nd	nd	95−24 (65−151)

Table 3. Concentrations (mg/L) of Flavones Present in *Citrus* Juices

	concentrations (mg/L) (mean ± SD) (min−max)			
flavones	orange	tangerine	lemon	grapefruit
lut-6,8-di-*C*-glc	1.7 ± 0.4 (1.1−2.6)	nd	3 ± 2 (0.6−6)	nd
api-6,8-di-*C*-glc	42 ± 14 (25−80)	14 ± 23 (1−66)	11 ± 6 (3−17)	36 ± 7 (27−47)
chry-6,8-di-*C*-glc	nd	nd	2.2 ± 0.8 (1.4−3.3)	nd
api-7-*O*-rut-4′-*O*-glc	nd	nd	1.4 ± 0.6 (0.6−2.6)	nd
dio-6,8-di-*C*-glc	nd	1.5 ± 0.9 (0.01−3.4)	28 ± 12 (13−46)	nd
chry-6,8-di-*C*-acylhexhex	nd	nd	4 ± 3 (0.4−8)	nd
dio-6,8-di-*C*-acylhexhex	nd	nd	5 ± 2 (3−8)	nd
lut-7-*O*-nhes-4′-*O*-glc	nd	nd	nd	0.3 ± 0.3 (0.01−0.8)
api-6-*C*-hex-*O*-hex	nd	nd	nd	5.5 ± 0.7 (4.0−6.8)
api-8-*C*-glc-*O*-pent	3 ± 1 (2−7)	nd	nd	nd
api-6-*C*-glc-*O*-pent	8 ± 4 (2−14)	nd	3 ± 1 (0.9−5)	9 ± 2 (6−12)
api-8-*C*-hex-*O*-acylgly	2 ± 1 (0.02−4)	nd	nd	nd
dio-8-*C*-glc	nd	0.2 ± 0.6 (0.01−2.3)	7 ± 2 (4−10)	nd
lut-7-*O*-rut	nd	2 ± 3 (0.01−8)	17 ± 7 (9−29)	nd
dio-6-*C*-glc	nd	0.1 ± 0.4 (0.01−1.5)	12 ± 9 (3−26)	nd
api-7-*O*-rut	nd	9 ± 15 (0.02−44)	7 ± 3 (3−12)	nd
api-7-*O*-nhes	nd	nd	nd	11 ± 4 (5−17)
chry-7-*O*-rut	nd	1 ± 2 (0.01−5)	4 ± 2 (2−7)	nd
dio-7-*O*-rut	nd	0.5 ± 0.8 (0.01−2.5)	10 ± 12 (2−39)	nd

linear up to 35% B at 90 min, and 35% B constant until 136 min, and finally, washing and reconditioning of the column was done. The flow rate was 0.8 mL/min, and injection volume was 50 μL. The column was operated at 30 °C, and sample vials on the injector were preserved at 4 °C. Flavanones were monitored and quantified at 280 nm, hydroxycinnamic acids at 320 nm, and flavonols, flavones and coumarins at 370 nm.

Identification and Quantitation of Phenolic Compounds. The identification of phenolic compounds will be reported in another work from this research group. For compounds whose standards were available, this identification was carried out by comparison of their retention times, their UV−visible spectra, and ESI-MS/MS spectra (recorded in MS^1 full scan and MS^2 product ion mode using as precursor ion the protonated molecule $[M+H]^+$ and the protonated aglycone $[Y_0]^+$) with those obtained by injecting standards in the same conditions. Other compounds were identified by comparison of UV−visible and MS spectra with those of standards of the same phenolic family and following the general strategy for the characterization of phenolic compounds in fruit juices by HPLC with diode array detection coupled to ESI triple quadrupole mass spectrometry published in a previous work.[29]

Quantitation was performed using integration areas in the calibration regression of the standards most similar to each phenolic compound quantified. Thus, flavanones were quantified as naringenin-7-*O*-rutinoside; apigenin glycosides as apigenin-7-*O*-glucoside; luteolin, diosmetin, and chrysoeriol glycosides as luteolin-7-*O*-glucoside; quercetin, isorhamnetin and kaempferol glycosides as quercetin-3-*O*-rutinoside, isorhamnetin-3-*O*-rutinoside, and kaempferol-3-*O*-rutinoside, respectively; ferulic and sinapic acid derivates as 5′-caffeoylquinic and sinapic acid, respectively; and scopoletin glycosides as scopoletin. These concentrations were corrected with the recovery factors previously published.[38]

Data Analysis and Chemometric Procedures. The data set consisted of a 83 × 49 matrix. Rows represented the *Citrus* fruit juices analyzed (83 objects) belonging to four categories and columns, the concentrations of 49 individual phenolics determined by HPLC-DAD. The four categories of analyzed *Citrus* fruit consisted of (1) 26 samples of sweet orange juices from five different cultivars and two subclasses, Navel and 'Blancas'; (2) 30 samples of tangerine juices from seven different cultivars and three subclasses, Clementinas, Satsuma, and hybrids; (3) 12 samples of lemon juices from two different cultivars belonging to the Verna subclass; and (4) 15 samples of grapefruit juices from three different cultivars belonging to pigmented and 'Blanca' subclasses (Table 1).

Statistical and chemometric data analyses were performed using the statistical software packages: Statgraphics plus 5.0 (Statistical Graphics Corp., 1994−2000) for PCA; Parvus (Forina M., Lanteri, S., and Armanino, C., University of Genova, 2004) for CA; SPSS for Windows

Table 4. Concentrations (mg/L) of Flavonols, Hydroxycinnamic Acids and Coumarins Present in *Citrus* Juices

polyphenols	concentrations (mg/L) (mean ± SD) (min−max)			
	orange	tangerine	lemon	grapefruit
Flavonols				
que-3-*O*-rhahex-7-*O*-hex	4 ± 1 (3−7)	3 ± 6 (0.01−16)	4 ± 2 (2−8)	nd
kaem-3-*O*-rhahex-7-*O*-hex	0.9 ± 0.5 (0.3−1.9)	1 ± 3 (0.01−8)	nd	nd
iso-3-*O*-hex-7-*O*-rhahex	1.6 ± 0.6 (0.8−3.0)	0.2 ± 0.3 (0.02−1.1)	nd	nd
iso-3-*O*-rhahex-7-*O*-hex	4 ± 1 (3−7)	1 ± 3 (0.02−7)	3 ± 2 (1−9)	nd
tam-3-*O*-rhahex-7-*O*-hex	nd	1 ± 1 (0.02−4)	nd	nd
que-3-*O*-rha-7-*O*-rhahex	1.3 ± 0.6 (0.01−2.3)	0.6 ± 0.8 (0.01−2.6)	nd	nd
que-7-*O*-rut	1.7 ± 0.4 (0.9−2.4)	2 ± 2 (0.01−6)	4 ± 4 (0.01−9)	2.2 ± 0.5 (1.6−3.0)
kaem-3-*O*-rhahex-7-*O*-rha	0.7 ± 0.8 (0.01−2.5)	2 ± 4 (0.01−10)	nd	nd
iso-3-*O*-rha-7-*O*-rhahex	1.6 ± 0.4 (0.9−2.3)	0.4 ± 0.9 (0.02−2.5)	nd	nd
que-3-*O*-rut	3 ± 1 (1−5)	5 ± 5 (2−18)	24 ± 11 (7−35)	nd
iso-7-*O*-rut	2 ± 1 (0.4−4)	nd	2.4 ± 0.7 (1.5−3.9)	nd
tam-7-*O*-rut	nd	1 ± 1 (0.02−4)	nd	nd
kaem-3-*O*-rut	0.4 ± 0.5 (0.01−1.6)	2 ± 4 (0.01−11)	nd	nd
iso-3-*O*-rut	2.3 ± 0.7 (1.3−4.1)	3 ± 3 (0.3−10)	6 ± 3 (3−10)	nd
Hydroxycinnamic Acids				
fer-*O*-hex	17 ± 7 (8−33)	16 ± 7 (5−29)	9 ± 3 (6−13)	45 ± 8 (36−58)
Snp-*O*-hex	5 ± 1 (3−7)	2 ± 1 (0.006−5)	3 ± 2 (2−8)	nd
Coumarins				
Sco-*O*-hex	nd	nd	nd	3 ± 1 (2−5)
Sco-*O*-rhahex	nd	nd	nd	1 ± 1 (0.009−3)

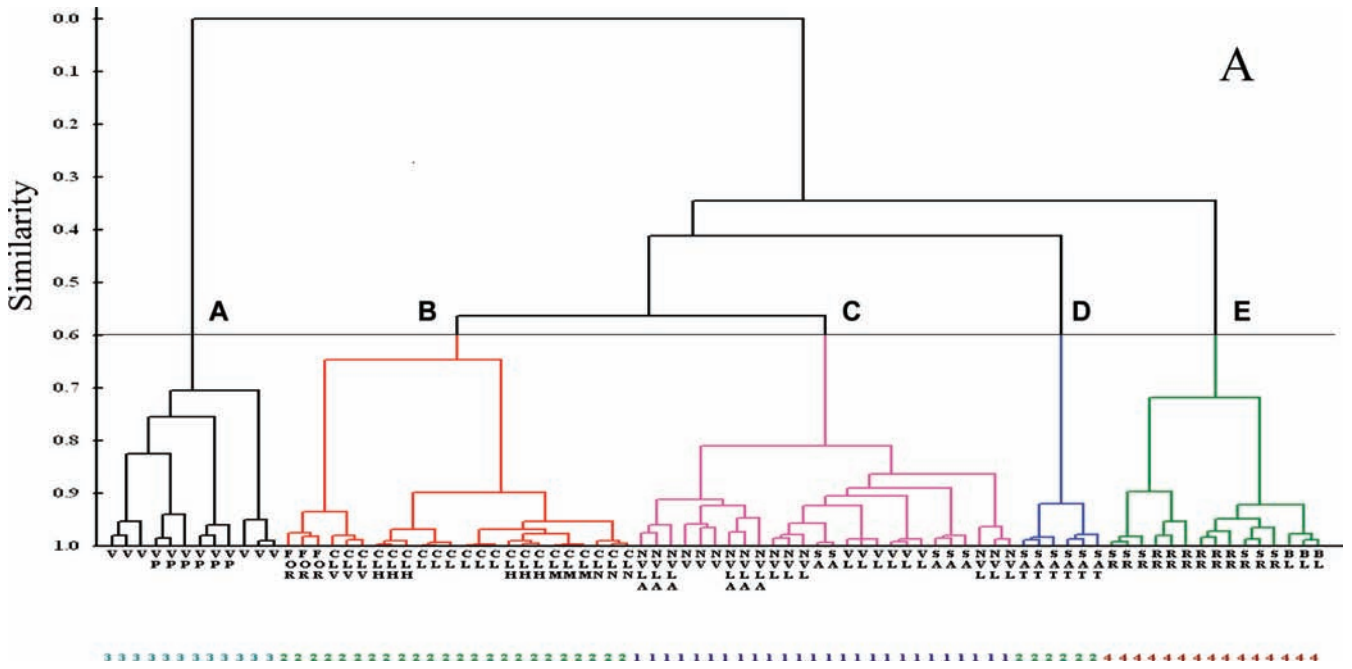

Figure 1. Dendrogram of cluster analysis. Samples codes, considering all variables: 1, sweet orange juice category; 2, tangerine juice category; 3, lemon juice category; and 4, grapefruit category.

(SPSS Inc., 1989−1999), and Matlab 2009 (MathWorks, Natick, MA) for LDA, and Unscrambler program v 9.7 (Camo, ASA, 2007) for the application of the PCR and PLS methods.

RESULTS AND DISCUSSION

Analytical Data. Tables 2−4 summarize analytical data of sweet orange, tangerine, lemon, and grapefruit phenolic compounds in juices for each cultivar obtained by chromatographic determinations.

Chemometric Study. *Multivariate Data Analysis.* The aim of this study is to propose a methodology based on multivariate analysis to determinate the possible adulteration in *Citrus* juices. This study was carried out (i) considering all variables (phenolic compounds) and (ii) variables whose concentrations were close to or higher than 10 mg/L at least in one of the four types of studied *Citrus* juices. If this last data set is used, the few variables used in the statistical treatment can be easily quantified; this would allow simpler and feasible methods (fast and accurate enough for screening purposes) for detecting *Citrus* juices adulterated in real samples.

The variables of this reduced data set were the flavanones, eriodictyol-7-*O*-rutinoside-4′-*O*-glucoside, naringenin-7-*O*-ruti-

noside-4′-*O*-glucoside, eriodictyol-7-*O*-rutinoside, naringenin-7-*O*-rutinoside, naringenin-7-*O*-neohesperidoside, hesperetin-7-*O*-rutinoside, hesperetin-7-*O*-neohesperidoside, isosakuranetin-7-*O*-rutinoside, isosakuranetin-7-*O*-neohesperidoside; the flavones, apigenin-6,8-di-*C*-glucoside, chrysoeriol-6,8-di-*C*-glucoside, diosmetin-6-*C*-glucoside, and luteolin-7-*O*-rutinoside; the flavonol, quercetin-3-*O*-rutinoside; and, the hydroxycinnamic acid, *O*-hexoside of ferulic acid.

Unsupervised Pattern Recognition Techniques. A preliminary evaluation of the information content in the data matrices was carried out using the unsupervised pattern recognition techniques: CA and PCA. CA highlights the existence of natural groupings among samples inside the data set,[39,40] and PCA is a factor analysis which reduces the number of variables retaining the maximum amount of variability present in the data in order to provide a better visualization of data structure in a reduced dimension.[34]

In order to carry out CA, data were autoscaled, sample similarities were calculated on the basis of squared Euclidean distance, and the Ward hierarchical agglomerative method was used to establish clusters. PCA was performed on the autoscaled data, and components with eigenvalues ≥1 were selected.

The results achieved by these methods considering all variables (49 phenolic compounds) are presented as a dendrogram in Figure 1 and as a tridimensional plot of sample scores in the space defined by the first three PCs in Figure 2.

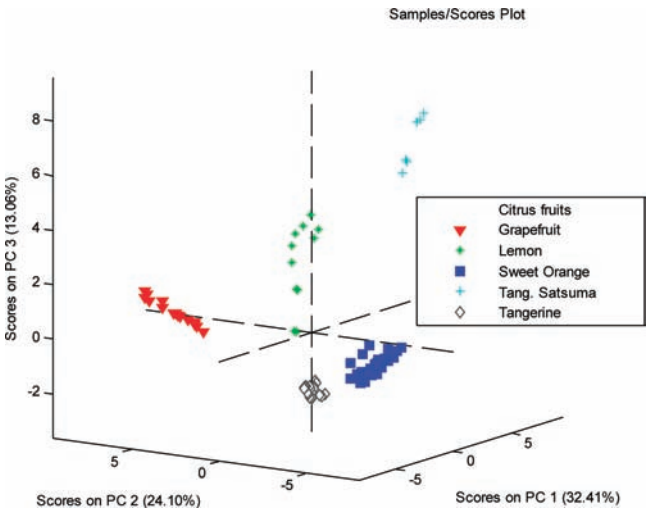

Figure 2. Principal component score projection of *Citrus* fruit juices: 1, sweet orange juices; 2, tangerine juices; 3, tangerine juices of Satsuma cultivar; 4, lemon juices; and 5, grapefruit juices.

The dendrogram (Figure 1) shows five clusters at a similarity level of 0.6: cluster A contains only lemon juices; clusters B and D are made up with tangerine juices (D consists of Satsuma tangerine cultivar with a special flavonol composition, and B is composed of the rest of tangerine juices); cluster C contains only sweet orange juices; and finally, cluster E contains grapefruit juices.

In the tridimensional plot (Figure 2), five separated groups were observed: 1, sweet orange juices, 2, lemon juices, 3, grapefruit juices, 4, tangerine juices of Clementina and the hybrid subclass, and 5, tangerine juices of the Satsuma subclass.

The three main PCs accounted for 70% of the total system variability. The variables that contributed more to the PC1 (see loading components in Supporting Information, File 1) were flavanones, naringenin-7-*O*-neohesperidoside-4′-*O*-glucoside, naringenin-*O*-hexosylhexoside, naringenin-7-*O*-neohesperidoside, hesperetin-7-*O*-neohesperidoside, and isosakuranetin-7-*O*-neohesperidoside; the flavone, apigenin-6-*C*-hexoside-*O*-hexoside; the flavonol, quercetin-3-*O*-rutinoside; and the coumarin, scopoletin-*O*-hexoside. Dominant features in the second principal component (PC2) were the flavanone naringenin-7-*O*-rutinoside-4′-*O*-glucoside; and the flavonols, isorhamnetin-3-*O*-hexoside-7-*O*-rhamnosylhexoside, quercetin-3-*O*-rhamnoside-7-*O*-rhamnosylhexoside, and isorhamnetin-3-*O*-rhamnoside-7-*O*-rhamnosylhexoside. Finally, the major contribution to PC3 was due to flavonols: kaempferol-3-*O*-rutinoside-7-*O*-glucoside, tamarixetin-3-*O*-rutinoside-7-*O*-glucoside, kaempferol-3-*O*-rhamnosylhexoside-7-*O*-rhamnoside, tamarixetin-7-*O*-rutinoside, and kaempferol-3-*O*-rutinoside.

These plots revealed that sweet orange juices are significantly more homogeneous than tangerine juices because the latter included juices from different tangerine cultivars that exhibit unique phenolic profiles such as those of the Satsuma cultivar. The Satsuma variety is remarkably different from Clementinas and hybrids and also from sweet orange juices; this fact is a consequence of Satsuma's higher content of flavonols (mainly explained by the third principal component). These results are in accordance with those from CA results: one cluster (cluster D) formed by tangerine juice made up with the Satsuma cultivar is separated from the rest of the tangerine juices and sweet orange juices.

These analyses were repeated considering only the variables which have a concentration close to or higher than 10 mg/L at least in one of the four species of studied *Citrus* juice. In this reduced data set, four clusters were observed at a similarity level of 0.70 (Figure 3A). These clusters were identified as follows: cluster A, made up of lemon juices; cluster D, consisting of grapefruit juices; B, containing mainly sweet orange juices and tangerine juices of the Satsumas and hybrid (Fortuna and Clemenvilla) cultivars; and D, due to tangerine juices of the Clementina subclass and sweet orange juices of the Salustiana cultivar belonging to the 'Blanca' subclass. It is interesting to highlight the behavior of sweet orange and tangerine juices: in spite of being different *Citrus* species, they are mixed in clusters B and C. This observation suggests that sweet orange and tangerine juices present similar phenolic profiles.

In PCA analysis, the bidimensional plot of the sample scores in the space defined by the two first principal components shows a natural separation of *Citrus* juice according to species (Figure 4A). Two groups, lemon and grapefruit juices, are clearly separated between them and from the rest of the groups (sweet orange and tangerine juices). Three principal components accounted for 87% of the total variability, and the loadings of the variables are showed in Supporting Information, File 2. The extreme variations in the phenolic profile between the samples of grapefruit and lemon juices and between these and sweet orange and tangerine juices may distort the CA and PCA. For this reason, samples of lemon and grapefruit juices were excluded, and the CA and PCA were repeated for only sweet orange and tangerine juices (data set consisted of a 56 × 11 matrix in which the variables eriodictyol-7-*O*-rutinoside-4′-*O*-glucoside, naringenin-7-*O*-neohesperidoside, hesperetin-7-*O*-neohesperidoside, and isosakuranetin-7-*O*-neohesperidoside, not detected in these juices, were removed).

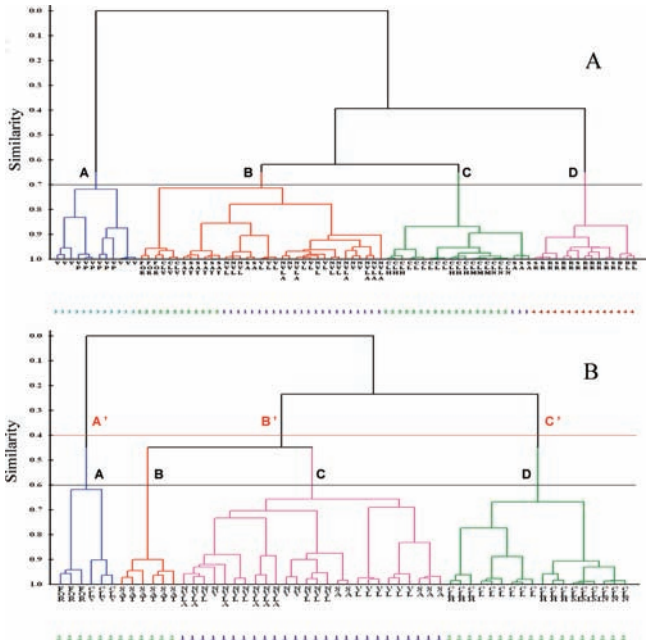

Figure 3. Dendrogram of cluster analysis. Samples codes: 1, sweet orange juice category; 2, tangerine juice category; 3, lemon juice category; and 4, grapefruit category (A); considering only the variables whose concentrations were close to or higher than 10 mg/L at least in one of the four species of the studied *Citrus*; and (B) dendrogram of CA considering only the variables whose concentrations were close to or higher than 10 mg/L at least in one of two the species of the studied *Citrus* juices. Sample codes: 1, orange juice category; and 2, tangerine juice category.

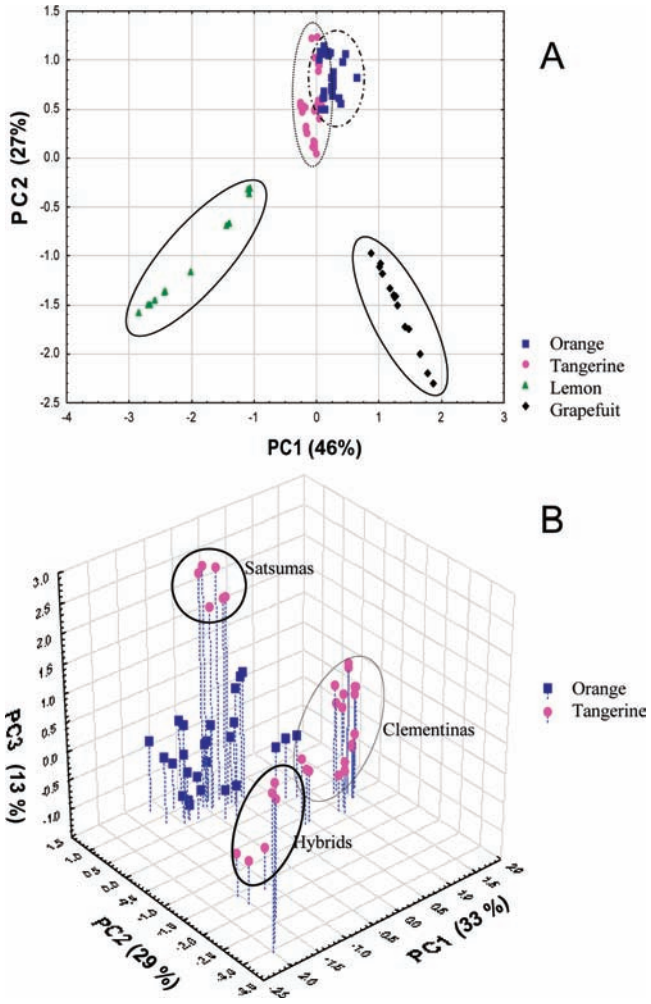

Figure 4. Principal component score plot of *Citrus* fruit juices considering only the variables whose concentrations were close to or higher than 10 mg/L at least in one of the four species of the studied *Citrus* juices. (A) PCA model of all *Citrus* fruits: 1, sweet orange juices; 2, tangerine juices; 3, lemon juices; and 4, grapefruit juices. (B) PCA model comprising only sweet oranges and tangerines; 1, sweet orange juices; and 2, tangerine juices.

In CA analysis, it is interesting to note the behavior of Satsuma tangerine juices (cluster B): despite being tangerine juices, they were included in the sweet orange juice cluster (cluster C) at a similarity level of 0.4 to generate cluster B′ as can observed in the dendrogram of the Figure 3B. As previously mentioned, this cultivar of tangerine presents a phenolic profile clearly different from the rest of the tangerine juices and more similar to the sweet orange juices. Increasing the level of similarity from 0.4 to 0.6, four clusters are distinguished: cluster C, due to sweet orange juices, and clusters A, B, and D consisting of tangerine juices grouped according to subclass type (A contains the juices of tangerine hybrid varieties (Fortuna and Clemenvilla); B includes juices of the group of Satsumas; and D, tangerine juices of the Clementina subclass).

The tridimensional plot of the sample scores in the space defined by the three principal components in PCA analysis shows a natural separation in four well separated groups (Figure 4B). A group contains sweet orange juices and the other three groups are made up of tangerine juices: one consisted of juices of the Satsuma tangerine cultivar and the other two were juices of hybrid cultivars and Clementina cultivars. These results correlate well with those obtained by CA since the samples from two categories are grouped in four well-defined areas.

Four PCs retain eigenvalues greater than 1, accounting for 84% of the total system variability. Observing the loadings of the variables (Supporting Information, File 3), the ones that contributed most to the PC1 (accounting for 33% of total variability) were two flavanones, naringenin-7-*O*-rutinoside and isosakuranetin-7-*O*-rutinoside; and the flavone apigenin-6,8-di-*C*-glucoside. Dominant features in the PC2 (accounting for 29% of total variability) were flavones diosmetin-6,8-di-*C*-glucoside and luteolin-7-*O*-rutinoside; the major contribution to PC3 (accounting for 13% of total variability) was due to flavonol quercetin-3-*O*-rutinoside and the hydroxycinnamic acid, ferulic acid *O*-hexose; and PC4 (accounting for 9% of total variability) was mainly due to the flavanone hesperetin-7-*O*-rutinoside.

Therefore, natural separation of samples depending on species and even *Citrus* subclass can be achieved with a smaller number of variables. This set of variables (15) captures almost all the information contained in the full set of variables (49 initially identified phenolic compounds) and it offers several advantages, among others, achieving a good classification of samples excluding many of the noisiest variables of the data set and the possibility of developing fast methods to guarantee *Citrus* fruit juice authenticity. These methods can save time and money in routine chemical analysis.

Supervised Pattern Recognition Technique. A supervised pattern recognition technique, linear discriminant analysis (LDA) was applied in order to attain classification rules for predicting composition of juices according to their phenolic

dx.doi.org/10.1021/jf300022u | *J. Agric. Food Chem.* 2012, 60, 3635−3644

profiles[34,41] using variables whose concentrations were close to or higher than 10 mg/L at least in one of the four species of studied *Citrus* juices.

Autoscaled data matrix of the phenolic profiles of *Citrus* fruit juices was the input data set of LDA. The LDA classification method was validated by cross-validation using random subsets containing 10% of the samples (700 iterations). The reliability of these classification models was studied in terms of recognition ability (correctly classified percentage of the members of the training set) and prediction ability (correctly classified percentage of the members of the validation test set, using the rules developed in the training step).

LDA produced classification rules that were 100% successful for tangerine, lemon, and grapefruit juices when assigning samples to classes, however, were 96% successful in sweet orange juices. Checking the model using cross-validation, the percentage of success was 100% for tangerine and grapefruit, whereas the prediction percentages were lower for sweet orange (95.7%) and lemon juices (89.9%). This fact indicated that the model established by this technique was very selective for tangerine and grapefruit juices. However, worse prediction percentages for sweet orange and lemon juices were due to probabilities of the classification of sweet orange and lemon juices as tangerine juices (4.3 and 10.1%, respectively) (Table 5).

Table 5. Classification Results for Supervised Pattern Recognition Technique LDA Applied to *Citrus* Juice Samples

Correct Classification (%)					
		predicted group membership			
	class	orange	tangerine	lemon	grapefruit
recognition ability (%)	sweet orange	96.2	3.8	0	0
	tangerine	0	100	0	0
	lemon	0	0	100	0
	grapefruit	0	0	0	100
prediction ability (%)	sweet orange	95.7	4.3	0	0
	tangerine	0	100	0	0
	lemon	0	10.1	89.9	0
	grapefruit	0	0	0	100

Multivariate Linear Regression Techniques. Principal component regression (PCR) and partial least-squares regression (PLS) are methods widely used to fit the observed data and to create models that can be used for prediction in many research fields such as food analysis.[42−44]

PCR combines linear regression and PCA. PCR establishes a relationship between the response variable (y) and the selected PCs of the input variables (x_i), by fitting a linear equation to the observed data. The dependent variable (y) is given by

$$y = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i + \in$$

where x_i (i = 1, ..., k) are the explanatory independent variables, $\hat{\beta}_i$ (i = 0, ..., k) are the regression coefficients, and $\in$ is the error associated with the regression and assumed to be normally distributed with both expectation value zero and constant variance.[45]

The predicted value given the regression model ($\hat{y}$) is calculated by

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i$$

PLS is a technique that is closely related to PCR. However, in PLS, the decomposition is performed in a slightly different fashion. Instead of first decomposing the matrix into a set of eigenvector and scores, and regressing them against the dependent variable as a separate step, PLS actually uses the dependent variable information during the decomposition process.[46]

In this work, these methods were applied for establishing predictive models for the estimation of the adulteration percentage in sweet orange juices in which grapefruit (fruit with a lower price than sweet orange) was added. For this purpose, these different multivariate statistical techniques (PLS and PCR) have been compared.

A new data set was composed using the data of phenolic concentrations in pure sweet orange and grapefruit juices and estimating phenolic concentrations in sweet orange juices adulterated with grapefruit in the following percentages: 5, 10, 20, 30, 50, and 70%. This data set consists of a 275 × 11 matrix, of which 26 samples were pure sweet orange juices, and 15 samples were pure grapefruit juices (taking each batch separately); 39 samples of the juice mixture for each percentage of adulteration with grapefruit (39 × 6 = 234 adulterated juices). These concentration values (phenolic compounds) were calculated using the average of three analyzed replicates for each variety and harvest. For adulterated juices, their composition values were theoretically calculated from pure juice values.

To select the number of components used to build PCR and PLS models and in order to model the system without overfitting, a segmented cross-validation procedure was used. Twenty groups for cancellation were used in the internal cross-validation of the regression models. Internal cross-validation consisted of randomly removing a group of training samples (5%) in a turn, performing the calibration with the rest of the samples, and using the model for predicting the excluded samples. These steps are repeated iteratively for each group of samples considered.[47] The root-mean-square error of calibration (RMSEC) depends on the number of components used for the calibration. Its value was chosen as an optimizing criterion to select the optimal number of components. The maximum number of components used to calculate the optimum RMSEC was selected as 4 for PLS and 3 for PCR. RMSEC is an indicator of the average error in the analysis for each component and how well the model fits to the data. RMSEC is defined by the following formula:

$$\mathrm{RMSEC} = \sqrt{\sum_{i=1}^{n} \frac{(y_{ci} - y_i)^2}{n}}$$

where y_{ci} is the predicted percentage of adulteration in each sweet orange juice in calibration sample i, y_i is the real percentage in calibration sample i, and n is the number of calibration samples. The RMSEC values were 6.0272 for PLS and 7.6891 for PCR.

The square correlation coefficient (R^2), which indicates the fraction of the total variance explained, is higher for the PLS

(0.9541) model than for the PCR model (0.9254). Adequate robustness of PCR and PLS models were checked estimating R^2 of the internal cross-validation (0.9508 for PLS and 0.9240 for PCR). The regression coefficients of the two models are shown in Table 6.

Table 6. Regression Coefficients Estimated in the PLS and PCR Models

variables	PLS	PCR
naringenin-7-*O*-rutinoside-4′-*O*-glucoside	−0.017	−0.187
eriodictyol-7-*O*-rutinoside	0.510	−0.785
naringenin-7-*O*-rutinoside	0.023	0.068
naringenin-7-*O*-neohesperidoside	0.006	0.014
hesperetin-7-*O*-rutinoside	−0.046	−0.007
hesperetin-7-*O*-neohesperidoside	0.462	0.242
isosakuranetin-7-*O*-rutinoside	0.239	−0.091
isosakuranetin-7-*O*-neohesperidoside	0.210	0.153
apigenin-6,8-di-*C*-glucoside	−0.186	−0.040
quercetin-3-*O*-rutinoside	−7.357	−3.199
ferulic acid-*O*-hexoside	0.027	0.378
intercept	42.043	20.039

The first factor for PLS and PCR is the most important to explain the *X* and *Y* variance as is shown in Supporting Information, File 4. This factor explains very well (>93%) the variance of the phenolic compounds: naringenin-7-*O*-neohesperidoside, hesperetin-7-*O*-neohesperidoside, and isosakuranetin-7-*O*-neohesperidoside; these three polyphenols are exclusively from grapefruit. Therefore, this latent variable can be considered as an indicator of grapefruit percentage in sweet orange juices. Considering the loadings of variables, it is clear that variables most important in grapefruit have positive values (naringenin-7-*O*-neohesperidoside, hesperetin-7-*O*-neohesperidoside, and isosakuranetin-7-*O*-neohesperidoside) and that they contribute the most to the first factor. However, the negative values of loadings were due to variables that are in higher concentration in sweet orange than grapefruit.

It is well known that the real predictive ability of any calibration model should not be judged solely by using internal validation. It is advisible to validate it on the basis of predictions for samples not included in the calibration test.[48] In order to check the quality of the proposed models, the external validation set, not included in the calibration step of the model, was used for checking the predicted percentages of grapefruit in the sweet orange juices by the proposed models (see Table 7). In this case, root-mean-square error of prediction (RMSEP) was chosen as a reference criterion to evaluate the built calibrations, which is given by the following formula:

$$\mathrm{RMSEP} = \sqrt{\sum_{i=1}^{m} \frac{(y_{pi} - y_i)^2}{m}}$$

where y_{pi} is based on the previously developed calibration models, y_i is the real percentage in calibration samples *i*, and *m* is the real number of evaluation samples. The RMSEP value showed us that both statistical techniques are suitable for the prediction of the percentage of grapefruit in sweet orange juices (Supporting Information, File 4). However, the best results were found for the PLS model with a RMSEP value of 4.744.

Thus, a reduced data set of the 15 phenolic compounds present at high concentration in *Citrus* fruit juices is capable of retaining most of the significant information of the full data set from the whole phenolic profile (49 compounds). The same groupings are described for both data sets, and the same conclusions are achieved. The proposal of this data set for *Citrus* juice monitoring allows the use of faster and simpler methods of analysis. These methods would enable the high throughput demanded by food producers to guarantee food authenticity.

Table 7. Prediction of Results for Adulteration Percentages of Grapefruit in Sweet Orange Juices by PLS and PCR

	adulteration (%)			
	PLS		PCR	
sample	reference	predicted	reference	predicted
OrG4-1	10.0	10.8	10.0	2.0
OrG4-2	10.0	7.9	10.0	14.6
OrG4-3	10.0	13.3	10.0	10.5
OrG4-4	10.0	6.8	10.0	13.9
OrG4-5	10.0	13.6	10.0	10.1
OrG4-6	10.0	13.7	10.0	5.7
OrG5-1	20.0	17.0	20.0	12.3
OrG5-2	20.0	28.5	20.0	31.6
OrG5-3	20.0	22.9	20.0	21.7
OrG5-4	20.0	26.4	20.0	30.1
OrG5-5	20.0	23.5	20.0	21.1
OrG5-6	20.0	22.9	20.0	19.7
OrG6-1	30.0	23.0	30.0	30.2
OrG6-2	30.0	24.7	30.0	26.8
OrG6-3	30.0	28.0	30.0	27.9
OrG6-4	30.0	23.7	30.0	24.1
OrG6-5	30.0	28.9	30.0	26.9
OrG6-6	30.0	31.8	30.0	41.3
OrG7-1	50.0	44.3	50.0	39.6
OrG7-2	50.0	50.2	50.0	44.0
OrG7-3	50.0	46.5	50.0	52.3
OrG7-4	50.0	48.6	50.0	39.6
OrG7-5	50.0	48.0	50.0	50.6
OrG7-6	50.0	59.0	50.0	58.2
OrG8-1	70.0	60.5	70.0	55.7
OrG8-2	70.0	68.5	70.0	64.9
OrG8-3	70.0	69.9	70.0	73.9
OrG8-4	70.0	66.3	70.0	58.6
OrG8-5	70.0	72.0	70.0	71.6
OrG8-6	70.0	81.1	70.0	81.7
RMSEP		4.744		6.853

All *Citrus* species are easily differentiated by chemometric tools except for tangerines and oranges. Cv. Satsuma exhibits a phenolic profile more similar to sweet oranges and is rather different from the rest of the tangerine cultivars. These phenolic profiles have been studied in fresh laboratory made juices; thus, some differences with commercial juices are expected. However, these conditions represent the desirable optimal conditions of obtaining a quality fruit juice in a traditional way.

Although more studies and a comprehensive external validation with adulterated samples are required, the regression methods presented here seem to be promising for detecting adulterations. The proposed PLS calibration model allowed successful detection of adulteration at the 10%−50% level with a suitable confidence interval (RMSEP = 4.7%) for screening purposes.

■ ASSOCIATED CONTENT

Supporting Information

Loadings (correlation coefficients between original variables and principal components) of the three PCs (PC1, PC2, and PC3) for each of the *Citrus* juice samples; loadings of the three PCs for each *Citrus* juice sample, considering only the variables whose concentrations were close to or higher than 10 mg/L at least in one of the four species of the studied *Citrus* juices; loadings of the four PCs (PC1, PC2, PC3, and PC4) for sweet orange and tangerine juice samples, considering only the variables whose concentrations were close to or higher than 10 mg/L at least in one of the four species of studied *Citrus* juices; PC1 loadings and total X and Y variance for each factor and explained variance of the polyphenols in each factor. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +34946015505. Fax: +34946013500. E-mail: luisangel.berrueta@ehu.es.

Funding

This research was supported by Gobierno del País Vasco/Eusko Jaurlaritza (Dpto. de Educación, Universidades e Investigación, poject number IT413-10) and Ministerio de Ciencia e Innovación, project number CTQ2009-08390). B.A.G. and S.G.L. thank Universidad del País Vasco/Euskal Herriko Unibertsitatea and Gobierno Vasco/Eusko Jaurlaritza, respectively, for their Ph.D. grants.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Technical and staff support provided by SGIker (UPV/EHU, MICINN, GV/EJ, and ESF) is gratefully acknowledged.

■ ABBREVATIONS USED

HPLC, high performance liquid chromatography; DAD, diode array detector; ESI, electrospray ionization; MS, mass spectrometry; nd, not detected; SD, standard deviation; CA, cluster analysis; PCA, principal component analysis; LDA, linear discriminant analysis; PCR, principal component regression; PLS, partial least-squares regression; PCs, principal components; RMSEC, root-mean-square error of calibration; RMSEP, root-mean-square error of prediction; Or, sweet orange; Ta, tangerine; Le, lemon; Gr, grapefruit; NVLA, Navel-Late; NVL, Navelina; NV, Navel; SA, Salustiana; VL, Valencia Late; CLH, Clementina Hernandina; CLM, Clementina Marisol; CLN, Clemenule; CL, Clementina; SAT, Satsuma; FOR, Fortuna; CLV, Clemenvilla; V, Verna; VP, Primafiori; SR, Star Ruby; RR, Red Ruby; BL, Blanco; Nar, Naringenin; Eri, eriodictyol; Isk, isosakuranetin; Hes, hesperetin; Lut, luteolin; Dio, diosmetin; Chry, chrysoeriol; Api, apigenin; Kaem, kaempferol; Que, quercetin; Iso, isorhamnetin; Tam, tamarixetin; Fer, ferulic acid; Snp, sinapic acid; Sco, scopoletin; rha, rhamnoside; hex, hexoside; pent, pentoside; glc, glucoside; rut, rutinoside; nhes, neohesperidoside; gly, glycoside

■ REFERENCES

(1) Spanos, G. A.; Wrolstad, R. E.; Heatherbell, D. A. Influence of processing and storage on the phenolic composition of apple juice. *J. Agric. Food Chem.* **1990**, *38*, 1572−1579.

(2) Spanos, G. A.; Wrolstad, R. E. Influence of processing and storage on the phenolic composition of Thompson seedless grape juice. *J. Agric. Food Chem.* **1990**, *38*, 1565−1571.

(3) Djoukeng, J. D.; Arbona, V.; Argamasilla, R.; Gomez-Cadenas, A. Flavonoid profiling in leaves of citrus genotypes under different environmental situations. *J. Agric. Food Chem.* **2008**, *56*, 11087−11097.

(4) Garciaviguera, C.; Bridle, P.; Ferreres, F.; Tomasbarberan, F. A. Influence of variety, maturity and processing on phenolic-compounds of apricot juices and jams. *Z. Lebensm. Unters. Forsch.* **1994**, *199*, 433−436.

(5) Bengoechea, M. L.; Sancho, A. I.; Bartolome, B.; Estrella, I.; GomezCordoves, C.; Hernandez, M. T. Phenolic composition of industrially manufactured purees and concentrates from peach and apple fruits. *J. Agric. Food Chem.* **1997**, *45*, 4071−4075.

(6) Marin, F. R.; Martinez, M.; Uribesalgo, T.; Castillo, S.; Frutos, M. J. Changes in nutraceutical composition of lemon juices according to different industrial extraction systems. *Food Chem.* **2002**, *78*, 319−324.

(7) Versari, A.; Biesenbruch, S.; Barbanti, D.; Farnell, P. J. Adulteration of fruit juices: Dihydrochalcones as quality markers for apple juice identification. *Food Sci. Technol.* **1997**, *30*, 585−589.

(8) Mouly, P.; Gaydou, E. M.; Auffray, A. Simultaneous separation of flavanone glycosides and polymethoxylated flavones in citrus juices using liquid chromatography. *J. Chromatogr., A* **1998**, *800*, 171−179.

(9) Dragovic-Uzelac, V.; Pospisil, J.; Levaj, B.; Delonga, K. The study of phenolic profiles of raw apricots and apples and their purees by HPLC for the evaluation of apricot nectars and jams authenticity. *Food Chem.* **2005**, *91*, 373−383.

(10) Pan, G. G.; Kilmartin, P. A.; Smith, B. G.; Melton, L. D. Detection of orange juice adulteration by tangelo juice using multivariate analysis of polymethoxylated flavones and carotenoids. *J. Sci. Food Agric.* **2002**, *82*, 421−427.

(11) Silva, B. M.; Andrade, P. B.; Mendes, G. C.; Valentao, P.; Seabra, R. M.; Ferreira, M. A. Analysis of phenolic compounds in the evaluation of commercial quince jam authenticity. *J. Agric. Food Chem.* **2000**, *48*, 2853−2857.

(12) Tomaslorente, F.; Garciaviguera, C.; Ferreres, F.; Tomasbarberan, F. A. Phenolic-compounds analysis in the determination of fruit jam genuineness. *J. Agric. Food Chem.* **1992**, *40*, 1800−1804.

(13) Lesschaeve, I.; Noble, A. C. Polyphenols: factors influencing their sensory properties and their effects on food and beverage preferences. *Am. J. Clin. Nutr.* **2005**, *81*, 330S−335S.

(14) Bartolome, B.; Bengoechea, M. L.; Sancho, A. I.; Estrella, I.; Hernandez, M. T.; Gomez-Cordoves, C. Differentiation of intermediate products (concentrates and purees) from the huh industry by means of phenolic content. *Z. Lebensm. Unters. Forsch., A* **1998**, *206*, 355−359.

(15) Del Caro, A.; Piga, A.; Vacca, V.; Agabbio, M. Changes of flavonoids, vitamin C and antioxidant capacity in minimally processed citrus segments and juices during storage. *Food Chem.* **2004**, *84*, 99−105.

(16) Prouillet, C.; Maziere, J. C.; Maziere, C.; Wattel, A.; Brazier, M.; Kamel, S. Stimulatory effect of naturally occurring flavonols quercetin and kaempferol on alkaline phosphatase activity in MG-63 human osteoblasts through ERK and estrogen receptor pathway. *Biochem. Pharmacol.* **2004**, *67*, 1307−1313.

(17) Hollman, P. C. H. Evidence for health benefits of plant phenols: local or systemic effects? *J. Sci. Food Agric.* **2001**, *81*, 842−852.

(18) Rimm, E. B.; Katan, M. B.; Ascherio, A.; Stampfer, M. J.; Willett, W. C. Relation between intake of flavonoids and risk for coronary heart disease in male health professionals. *Ann. Intern. Med.* **1996**, *125*, 384−389.

(19) Ranganna, S; Govindarajan, V. S.; Ramana, K. V. R.; Kefford, J. F. Citrus fruits: Varieties, chemistry, technology, and quality

evaluation. Part II. Chemistry, technology, and quality evaluation. A. Chemistry. *Crit. Rev. Food Sci. Nutr.* **1983**, *18*, 313−386.
(20) Gattuso, G.; Barreca, D.; Gargiulli, C.; Leuzzi, U.; Caristi, C. Flavonoid composition of citrus juices. *Molecules* **2007**, *12*, 1641−1673.
(21) Robards, K.; Li, X.; Antolovich, M.; Boyd, S. Characterisation of citrus by chromatographic analysis of flavonoids. *J. Sci. Food Agric.* **1997**, *75*, 87−101.
(22) Dugo, P.; Lo Presti, M.; Ohman, M.; Fazio, A.; Dugo, G.; Mondello, L. Determination of flavonoids in citrus juices by micro-HPLC-ESI/MS. *J. Sep. Sci.* **2005**, *28*, 1149−1156.
(23) Caristi, C.; Bellocco, E.; Gargiulli, C.; Toscano, G.; Leuzzi, U. Flavone-di-C-glycosides in citrus juices from southern Italy. *Food Chem.* **2006**, *95*, 431−437.
(24) Peterson, J. J.; Dwyer, J. T.; Beecher, G. R.; Bhagwat, S. A.; Gebhardt, S. E.; Haytowitz, D. B.; Holden, J. M. Flavanones in oranges, tangerines (mandarins), tangors, and tangelos: a compilation and review of the data from the analytical literature. *J. Food Comp. Anal.* **2006**, *19*, S66−S73.
(25) Caristi, C.; Bellocco, E.; Panzera, V.; Toscano, G.; Vadala, R.; Leuzzi, U. Flavonoids detection by HPLC-DAD-MS-MS in lemon juices from sicilian cultivars. *J. Agric. Food Chem.* **2003**, *51*, 3528−3534.
(26) Horowitz, R. H.; Gentili, B. Flavonoid Constituents of Citrus. In *Citrus Science and Technology*; Nagy, S. S. P. E. V. M. K., Ed.; Avi Publishers: Westport, CT, 1977; pp 397−426.
(27) Rouseff, R. L.; Martin, S. F.; Youtsey, C. O. Quantitative survey of narirutin, naringin, hesperidin, and neohesperidin in citrus. *J. Agric. Food Chem.* **1987**, *35*, 1027−1030.
(28) Mouly, P.; Gaydou, E. M.; Estienne, J. Column liquid-chromatographic determination of flavanone glycosides in citrus: Application to grapefruit and sour orange juice adulterations. *J. Chromatogr.* **1993**, *634*, 129−134.
(29) Abad-Garcia, B.; Berrueta, L. A.; Garmon-Lobato, S.; Gallo, B.; Vicente, F. A general analytical strategy for the characterization of phenolic compounds in fruit juices by high-performance liquid chromatography with diode array detection coupled to electrospray ionization and triple quadrupole mass spectrometry. *J. Chromatogr., A* **2009**, *1216*, 5398−5415.
(30) Barreca, D.; Bellocco, E.; Caristi, C.; Leuzzi, U.; Gattuso, G. Distribution of C- and O-glycosyl flavonoids, (3-hydroxy-3-methylglutaryl)glycosyl flavanones and furocoumarins in *Citrus aurantium* L. juice. *Food Chem.* **2011**, *124*, 576−582.
(31) Gattuso, G.; Caristi, C.; Gargiulli, C.; Bellocco, E.; Toscano, G.; Leuzzi, U. Flavonoid glycosides in bergamot juice (*Citrus bergamia risso*). *J. Agric. Food Chem.* **2006**, *54*, 3929−3935.
(32) Cao, X. L.; Wang, C.; Pei, H. R.; Sun, B. G. Separation and identification of polyphenols in apple pomace by high-speed counter-current chromatography and high-performance liquid chromatography coupled with mass spectrometry. *J. Chromatogr., A* **2009**, *1216*, 4268−4274.
(33) Gomez-Romero, M.; Segura-Carretero, A.; Fernandez-Gutierrez, A. Metabolite profiling and quantification of phenolic compounds in methanol extracts of tomato fruit. *Phytochemistry* **2010**, *71*, 1848−1864.
(34) Berrueta, L. A.; Alonso-Salces, R. M.; Heberger, K. Supervised pattern recognition in food analysis. *J. Chromatogr., A* **2007**, *1158*, 196−214.
(35) Ashurts, P. R.; Dennis, M. J. *Food Authentication*; Blackie Academic & Proffesional: London, 1996.
(36) Twomey, M.; Downey, G.; McNulty, P. B. The potential of NIR spectroscopy for the detection of the adulteration of orange juice. *J. Sci. Food Agric.* **1995**, *67*, 77−84.
(37) Tomas-Barberan, F. A.; Gil, M. I.; Cremin, P.; Waterhouse, A. L.; Hess-Pierce, B.; Kader, A. A. HPLC-DAD-ESIMS analysis of phenolic compounds in nectarines, peaches, and plums. *J. Agric. Food Chem.* **2001**, *49*, 4748−4760.
(38) Abad-Garcia, B.; Berrueta, L. A.; Lopez-Marquez, D. M.; Crespo-Ferrer, I.; Gallo, B.; Vicente, F. Optimization and validation of a methodology based on solvent extraction and liquid chromatography for the simultaneous determination of several polyphenolic families in fruit juices. *J. Chromatogr., A* **2007**, *1154*, 87−96.
(39) Lavine, B. K. Pattern recognition. *Crit. Rev. Anal. Chem.* **2006**, *36*, 153−161.
(40) Massart, L.; Kaufman, L. *The Interpretation of Analytical Chemical Data by the Use of Clusters Analysis*; Wiley Interscience: New York, 1983.
(41) Massart, D. L.; Vandeginste, B. G. M.; Uydens, L. M. C.; Jong, S.; Ewi, P. J.; Meyers-Verbeke, J. *Handbook of Chemometrics and Qualimetrics, Part A*; Elsevier: Amsterdam, 1997.
(42) Rodriguez-Nogales, J. M. Approach to the quantification of milk mixtures by partial least-squares, principal component and multiple linear regression techniques. *Food Chem.* **2006**, *98*, 782−789.
(43) Casal, V.; MartinAlvarez, P. J.; Herraiz, T. Comparative prediction of the retention behaviour of small peptides in several reversed-phase high-performance liquid chromatography columns by using partial least squares and multiple linear regression. *Anal. Chim. Acta* **1996**, *326*, 77−84.
(44) Perez-Coello, M. S.; Martin-Alvarez, P. J.; Cabezudo, M. D. Prediction of the storage time in bottles of Spanish white wines using multivariate statistical analysis. *Z. Lebensm. Unters. Forsch., A* **1999**, *208*, 408−412.
(45) Agirre-Basurko, E.; Ibarra-Berastegi, G.; Madariaga, I. Regression and multilayer perceptron-based models to forecast hourly O-3 and NO_2 levels in the Bilbao area. *Environ. Modell. Softw.* **2006**, *21*, 430−446.
(46) Vandeginsten, B. G. M.; Massart, D. L.; Buydens, L. M. C.; De Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J. *Handbook of Chemometrics and Qualimetrics, Part B*; Elsevier: The Netherlands, 1998.
(47) Khanchi, A. R.; Mahani, M. K.; Hajihosseini, M.; Maragheh, M. G.; Chaloosi, M.; Bani, F. Simultaneous spectrophotometric determination of caffeine and theobromine in Iranian tea by artificial neural networks and its comparison with PLS. *Food Chem.* **2007**, *103*, 1062−1068.
(48) Boeris, M. S.; Luco, J. M.; Olsina, R. A. Simultaneous spectrophotometric determination of phenobarbital, phenytoin and methylphenobarbital in pharmaceutical preparations by using partial least-squares and principal component regression multivariate calibration. *J. Pharm. Biomed. Anal.* **2000**, *24*, 259−271.