JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

J. Agric. Food Chem. **2008,** *56,* 9835–9842 **9835**

Improving the Screening Process for the Selection of Potato Breeding Lines with Enhanced Polyphenolics Content

CECILIA SHIROMA-KIAN,[†] DAVID TAY,[‡] IVÁN MANRIQUE,[‡] M. MONICA GIUSTI,[†] AND LUIS E. RODRIGUEZ-SAONA*,[†]

Department of Food Science and Technology, Ohio State University, 2015 Fyffe Court, Columbus, Ohio 43210, and International Potato Center, Av. La Molina 1895, La Molina, Lima 12, Perú

Efficient selection of potato varieties with enhanced nutritional quality requires simple, rapid, accurate, and cost-effective assays to obtain tuber chemical composition information. Our objective was to develop simple protocols to determine phenolics, anthocyanins, and antioxidant capacity in polyphenolic extracts of potatoes using Fourier transform infrared spectroscopy combined with multivariate techniques. Lyophilized potato samples (23) were analyzed. Polyphenolic compounds were extracted from potatoes and applied directly applied onto a three-bounce ZnSe crystal for attenuated total reflectance measurements in the infrared region of 4000 to 700 cm^{-1}. Robust models were generated ($r \geq 0.99$) with standard error of cross-validation values of 4.17 mg gallic acid equivalent/100 g (total phenolics), 0.87 mg pelargonidin-3-glucoside/100 g (monomeric anthocyanins), and 130.8 μmol Trolox equivalent/100 g (antioxidant capacity) potato powder. In addition, classification models discriminated potato samples at the species and variety level. Application of a simple infrared spectroscopic protocol allowed simultaneous rapid quantification of specific nutritional components in potatoes and efficient selection of value-added potato varieties.

KEYWORDS: FTIR; phenolics content; anthocyanins content; antioxidant capacity; potatoes

INTRODUCTION

Potato is the world's fourth major food crop and is considered a key product in alleviating poverty and improving food security (*1*). The Food and Agriculture Organization (FAO) of the United Nations reported that in 2005 two-thirds of the 320 million tons of the cultivated potato was consumed as food (*1*). Potato is an appealing crop due to its ability to grow fast, adaptability to different environments, production in high yield, and response to low inputs (*2*). In addition, its characteristics (up to 85% of the plant is edible as compared to ~50% in cereals) and ability to produce high yields under harsher climates and using less land than any other major crop make potato a suitable staple food in many countries (*1*). Furthermore, potato is an important source of carbohydrates, proteins (due to the highest protein content among tuber crops), and vitamin C, and some varieties of potato are a rich source of polyphenolic compounds (*3, 4*).

Although most of cultivated potatoes belong to the botanical species *Solanum tuberosum*, there are many other species and varieties that differ in size, shape, color, texture, cooking characteristics, and taste (*1, 3*−*6*). Skin- and/or flesh-colored potatoes are gaining popularity not only for their appealing colors and culinary uses but also for their higher content of polyphenolic compounds and reported health benefits (*3, 5, 7, 8*). Naturally occurring polyphenolics include anthocyanins, phenolic acids, hydroxycinnamic acid derivatives, and lignans (*11*). Many of these compounds have been shown to be effective scavengers of free radicals with recognized antioxidant capacities (*4, 5, 10*). It is believed that polyphenolic compounds are associated with certain health benefits such as inhibition of cholesterol accumulation in blood, reduction in the risk of coronary heart disease, prevention of some types of cancer, and retardation of macular degeneration, among others (*4, 8*). Depending on the potato variety, a wide range of phenolic acids (chlorogenic acid, caffeic acid, ferulic acid, coumaric acid, and quinic acid) and anthocyanins [3-(*p*-coumaroyl-rhamnosylglucoside)-5-glucoside derivatives of pelargonidin, peonidin, petunidin, and malvidin] have been reported (*4, 8, 11*). Generally, pigmented potatoes have shown higher (2−3 times) antioxidant capacities than white-fleshed potatoes as shown by the oxygen radical absorbance capacity (ORAC) of plasma assays (*7*).

Breeders are working on the production of new potato varieties with enhanced nutritional traits and appropriate growing conditions (*7*). The development of varieties with enhanced characteristics and nutritive value is possible through direct selection in segregating populations and through molecular-assisted strategies. Biofortification, the process of enriching the nutritional quality of crops through traditional practices of

* To whom correspondence should be addressed. Tel: 614-292-3339. Fax: 614-292-0218. E-mail: rodriguez-saona.1@osu.edu.

[†] Ohio State University.

[‡] International Potato Center.

10.1021/jf801716b CCC: $40.75
Published on Web 10/03/2008

Table 1. Description Related to Samples' CIP Codification, Species Name, Total Phenolics, Monomeric Anthocyanin Content, and Antioxidant Capacity as Determined by Reference Methods

sample	ID[a]	species name (ploidi level)	total phenolics (mg gallic acid equiv/100 g DW)	monomeric anthocyanin (mg pelargonidin-3-glucoside/100 g DW)	antioxidant capacity (μmol Trolox equiv/100 g DW)
1	700234	*S. tuberosum* ssp. *andigena* (4×)	85.98	13.75	2796.08
2	700304	*S. stenotomum* ssp.*goniocalyx* (2×)	10.79	2.41	206.77
3	701165	*S. stenotomum* ssp. *stenotomum* (2×)	44.50	1.17	1444.00
4	701273	*S. tuberosum* ssp. *andigena* (4×)	102.81	15.96	3220.27
5	701515	*S. tuberosum* ssp. *andigena* (4×)	86.35	17.81	2757.79
6	701675	*S. chaucha* (3×)	65.67	7.91	2576.73
7	701997	*S. tuberosum* ssp. *andigena* (4×)	108.69	20.67	4137.17
8	702395	*S. tuberosum* ssp. *andigena* (4×)	91.57	14.17	3061.48
9	702464	*S. tuberosum Hyb*	88.74	24.70	3325.15
10	702736	*S. stenotomum* ssp. *stenotomum* (2×)	31.55	3.04	1419.91
11	702815	*S. stenotomum* ssp. *stenotomum* (2×)	82.18	17.55	3050.44
12	703168	*S. stenotomum* ssp.*goniocalyx* (2×)	23.59	2.22	743.30
13	703265	*S. tuberosum* ssp. *andigena* (4×)	49.58	3.56	1363.92
14	703274	*S. phureja* (2×)	109.16	16.04	4370.94
15	703287	*S. stenotomum* ssp. *stenotomum* (2×)	76.19	21.64	3006.65
16	703312	*S. stenotomum* ssp. *stenotomum* (2×)	45.40	5.04	2216.46
17	703317	*S. stenotomum* ssp. *stenotomum* (2×)	36.14	13.32	1490.91
18	703421	*S. stenotomum* ssp. *stenotomum* (2×)	40.08	3.38	1701.86
19	704022	*S. stenotomum* ssp. *stenotomum* (2×)	8.52	0.66	595.79
20	704058	*S. tuberosum* ssp. *andigena* (4×)	153.70	102.71	4394.15
21	704327	*S. chaucha* (3×)	23.86	7.37	1625.61
22	705543	*S. chaucha* (3×)	41.96	7.63	2105.34
23	707135	*S. chaucha* (3×)	58.89	6.90	1979.86

[a] Identification number assigned by CIP.

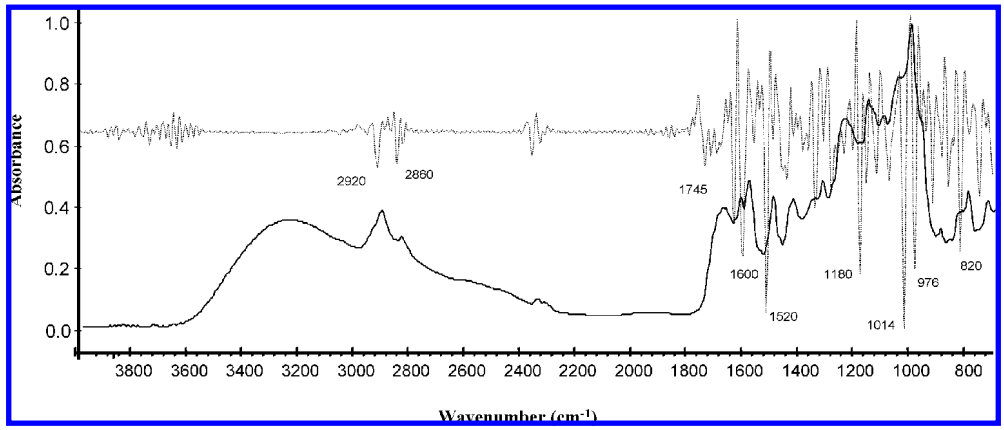

Figure 1. Fourier transform IR absorption and respective second derivative (five-point window) for polyphenolic extracts of potato collected in the 4000 to 700 cm^{-1} region and measured on a three-bounce ATR-IR zinc selenide (ZnSe) crystal.

breeding and biotechnology, has been suggested as an alternative approach to increase the micronutrients content in potatoes (*12*). Through biofortification practices, it will be possible to provide farmers with improved crop varieties (higher content of nutrient and higher productivity) in a cost-effective manner (*12*). Efficient selection of improved varieties requires simple, rapid, accurate, and inexpensive assays to obtain tuber composition information. However, current methods to analyze potato composition are time-consuming, expensive, and use hazardous organic solvents. Fourier transform infrared spectroscopy (FTIR) is an attractive technology for the rapid, sensitive, and high-throughput analysis of food components. It can provide real-time measurements, reduce the use of hazardous organic solvents, simplify data acquisition, and enable immediate predictions without the need for special skills from the users (*13*−*17*). The mid-IR spectra (4000 to 700 cm^{-1}) allow for the chemically based discrimination of organic constituents and produce distinct and reproducible biochemical fingerprints. Recent advances in FTIR spectroscopic instrumentation and multivariate techniques have shown the potential for the analysis of complex multispectroscopic information for the discrimination, classification, and identification of various components in foods.

The objectives of this study were to develop simple and rapid IR protocols for simultaneous determination of total phenolics, anthocyanins, and antioxidant capacity in polyphenolic extracts from potato. In addition, we evaluated the feasibility of using IR spectroscopy to classify potatoes at species and variety levels to support the breeder's ability in screening potatoes with enhanced nutritional traits.

MATERIALS AND METHODS

Potato Samples. Lyophilized potato samples with varying contents of phenolics, anthocyanins, and antioxidant capacity were supplied by the International Center of Potato (CIP, Lima, Perú). A total of 23 cultivated samples from six different species (**Table 1**) were analyzed. Samples were kept in a desiccator containing drierite and in the dark until analysis.

Polyphenolics Extraction and Purification. Extractions were performed using the procedure described by Giusti and Wrolstad (*18*) with some modifications. Lyophilized potato material (≈3 g) was

Table 2. Multivariate Analysis PLSR Model Parameters for Potato Extracts with FTIR Technique

analysis	model parameters					
	composition range[a]	no. of factors	SECV	rVAL	SEC	rCAL
total phenolics (mg gallic acid equiv/100 g DW)	8.52−153.20	20	4.17	0.99	3.18	0.99
monomeric anthocyanin (mg pelargonidin-3-glucoside/100 g DW)	0.66−102.71	20	0.87	0.99	0.57	0.99
antioxidant capacity (μmol Trolox equiv/100 g DW)	206.77−4394.15	20	130.8	0.99	97.99	0.99

[a] From **Table 1**. rVAL, correlation coefficient of cross-validation; SEC, standard error of calibration; and rCAL, correlation coefficient of calibration.

blended with 45 mL of aqueous acetone (30:70 v/v) for 10 min in a model RKDYNAL Rotamix (Dynal Biotech Inc., Lake Success, NY) and centrifuged, and the supernatant was collected. The residual cake was re-extracted with 25 mL of aqueous acetone to ensure that a clear solution was obtained. Filtrates were combined and gently mixed with chloroform (1:1 chloroform:acetone), and phases were allowed to separate. The aqueous extract (top layer) was taken to a known volume (25 mL) with distilled water and passed through a C-18 mini-column to remove sugars and acids that could affect the signal of the spectra (*19*). The C18 mini-column had a 10 mL capacity (Waters, Milford, MA) and was previously activated with methanol followed by 0.01% aqueous HCl (*18*). Phenolics and anthocyanins were adsorbed onto the mini-column and recovered with methanol containing 0.01% HCl (v/v). Polyphenolic extracts were split in halves concentrated under nitrogen gas. One half was resuspended in 0.5 mL of acidified methanol and utilized for analysis with attenuated total reflectance IR (ATR-IR) spectroscopy, and the other half was redissolved in deionized water (1 mL) containing 0.01% HCl for further analysis. Acidified conditions were used to ensure the stability of anthocyanins as well as reproducible FTIR spectra (*19, 20*). Three extractions per sample were performed.

Monomeric Anthocyanin Content. The pH-differential method was used to quantify the monomeric anthocyanin content following the procedure reported by Giusti and Wrolstad (*18*). A model UV-2450 UV/vis spectrophotometer (Shimadzu Corp., Addison, IL) and 1 cm path length cells were used for spectroscopic measurements at 510 (maximum absorbance) and 700 nm (to correct for haze). The pigment content was calculated as pelargonidin-3-glucoside, using an extinction coefficient of 31600 L/cm/mol and molecular weight of 433.2 g/mol (*8*). The analysis was performed in triplicate.

Total Phenolic Contents. The total phenolic content was determined using the microscale Folin−Ciocalteu (FC) colorimetric method described by Waterhouse (*21*). The reaction mixture containing 20 μL of potato polyphenolic extract, 1.55 mL of water, and 100 μL of FC reagent was mixed thoroughly and incubated for up to 8 min. Sodium carbonate solution (300 μL) was added, mixed, and incubated for 2 h at room temperature. The absorbance at 765 nm was measured using a Shimadzu UV−visible spectrophotometer (UV-2450, Shimadzu Corp.) and 1 cm path length cells. A gallic acid calibration standard curve was developed and used to calculate the phenolic content in samples, expressed as mg/L of gallic acid equiv. Quantification of total phenolics was done in triplicate. FC reagent and gallic acid standard were obtained from Sigma Aldrich (St. Louis, MO) and Fisher Scientific (Pittsburgh, PA), respectively.

Antioxidant Capacity ORAC Assay. The antioxidant capacity ORAC assay was carried out according to the method reported by Cao et al. (*24*) with slight modifications. In the assay mixture, fluorescein was used as a target of free radical attack, and 2,2′-azobis (2-amidinopropane) dihydrochloride (AAPH) was used as a peroxyl radical generator. Trolox, a water-soluble analogue of vitamin E, was used to construct a standard curve in the ranges between 6.25 and 125 μM. A Cary Eclipse fluoremeter (Varian Inc., Palo Alto, CA) was set to record fluorescence for 90 min with an excitation and emission wavelength of 490 and 520 nm, respectively. The antioxidant capacity was calculated using the differences in areas under the fluorescein decay curves between the blank and a sample and expressed as μmol of Trolox equiv per 100 g of dry weight (DW). Measurements were done by triplicate. Reagents and standards were obtained from Fisher Scientific.

FTIR. A Varian 3100 FTIR Excalibur series spectrometer (Varian Inc.) with a potassium bromide beam splitter and deuterated triglycine sulfate (DTGS) detector was used for all readings, operating at 8 cm^{-1} resolution. Polyphenolic extracts in methanol (10 μL) were dispensed onto a three-bounce ATR-IR ZnSe crystal (Pike Technologies, Madison, WI) and vacuum-dried to form a thin film. Spectra were collected over the wavenumber region from 4000 to 700 cm^{-1} and were displayed in terms of absorbance. Interferograms of 64 scans were coadded. For each polyphenolic extract, four independent measurements were collected. The absorbance spectrum was obtained by ratioing the sample single beam spectrum against that of a blank optical path (reference spectrum). The instrument was continuously purged with CO_2-free dry air from a CO_2 RP140 dryer (Dominick Hunter, Charlotte, NC).

Chemometric Analyses. Data analysis was performed using a multivariate statistics program called Pirouette (version 3.11, Infometrix, Bothell, WA). Spectra were imported as GRAMS (.spc) files into Pirouette and normalized to the most intense band in the spectrum. IR spectra were transformed to its second-derivative using a five-point polynomial-fit Savitzky−Golay function. The transformed spectra were analyzed by partial least-squares regression (PLSR) that was cross-validated (leave-one-out) to generate calibration models. Models were evaluated in terms of loading vectors, standard error of cross-validation (SECV), correlation coefficient (r), and outlier diagnostics. Classification of potatoes by species/variety was done by using soft independent modeling of class analogy (SIMCA). SIMCA is a supervised pattern recognition classification technique based on principal component analysis (PCA) that requires previous category membership knowledge of samples. Scores-plot projections of the original data onto principal component axes were used to visualize sample clustering (patterns, groupings, or outliers). Class borders, which are 95% probability clouds, are defined around the samples of the class. Between-class distances were calculated using between-class residuals, and variable importance was determined by comparing average residual variance of each class to all classes and residual variance of all classes to themselves (*23*). Variable importance, also known as discriminating power, were used to identify the variables (wavenumbers) that have a predominant effect on sample classification while minimizing the difference between samples within a cluster and maximizing differences between samples from different clusters (*24*).

Performance of Calibration Models. Validation of calibration models included a set of potato polyphenolic extracts from four different varieties. A total of 12 independent spectra for each sample were collected and used to test the predictive ability of the calibration model for phenolics, monomeric anthocyanin, and antioxidant capacity content in potatoes. These values were compared to those obtained by the reference methods. Quantitative predictions of the unknowns were done by PLSR using Pirouette software. A SIMCA calibration model was also used to evaluate the ability to predict the species/variety class to which each sample belongs.

RESULTS AND DISCUSSION

Screening of Potato Samples. Potato samples were supplied by the International Potato Center (CIP), and they were sent in the lyophilized powder form to minimize undesirable degradative reactions of the plant material during shipment (*19, 25*). The diverse potato samples provided by CIP had differences in tuber shape, primary and secondary color in skin and flesh, and color distribution within the skin and flesh of the tuber as reported by CIP collaborators. Predominant colors ranged from white to deep purple in the flesh and from white-cream to purple black in the skin. These color differences were evident not only

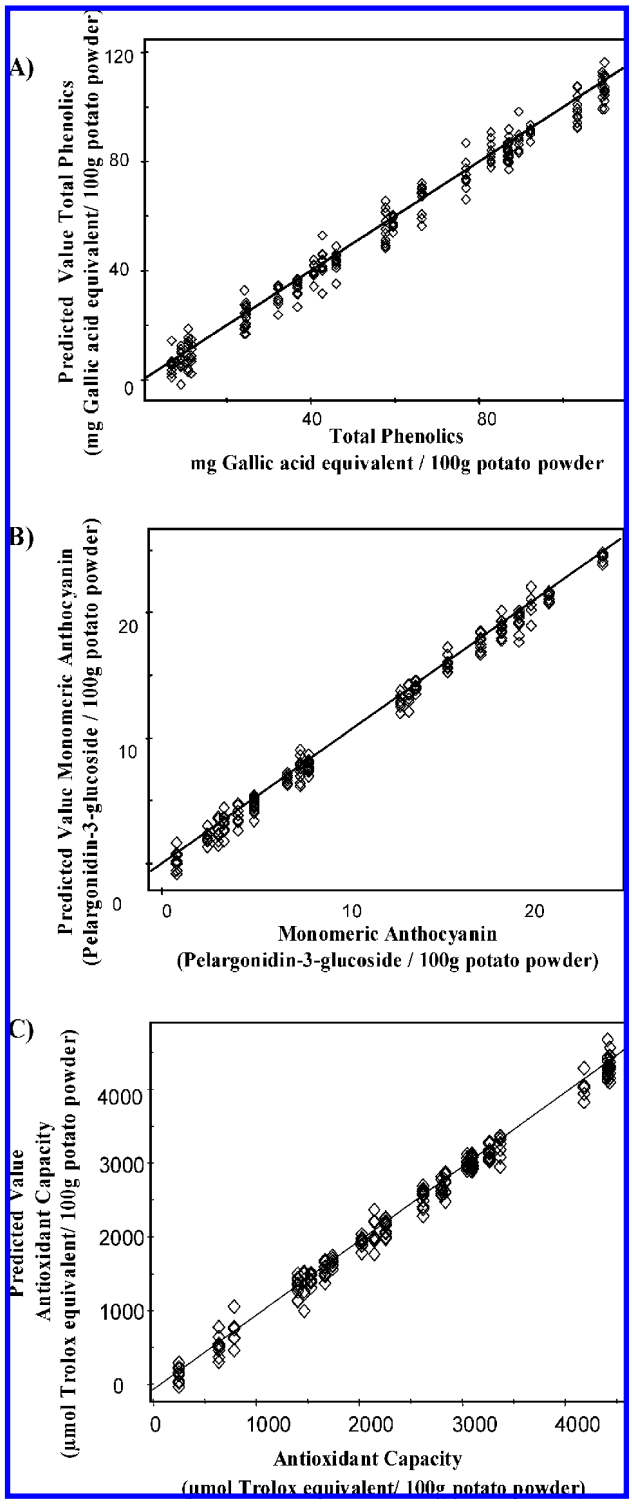

Figure 2. PLSR calibration and validation plots for total phenolic, monomeric anthocyanin, and antioxidant capacity with FTIR spectroscopy techniques. (**A**) PLSR model of total phenolic content, (**B**) PLSR model of monomeric anthocyanins, and (**C**) PLSR model of antioxidant capacity. The Savitzky−Golay second derivative transformation was used for the analysis of FTIR spectra.

between different species but also among varieties within the same species. The large diversity of native potatoes (more than 4500 distinct native varieties at CIP), ranging from diploids to pentaploids, has been reported to be responsible for such variability in tuber shape, flesh and skin color, flavor, and cooking quality (*3*).

Table 3. Comparison between Expected Reference Values and Predicted Values Obtained with FTIR PLSR Models for Total Phenolics, Monomeric Anthocyanin, and Antioxidant Capacity

			predicted value with PLSR	
analysis	sample[a]	expected value (reference method)	IR	deviation[b] (%)
total phenolics (mg gallic acid equiv/100 g DW)	7	108.40	108.81	0.38
	8	93.43	93.89	0.49
	9	88.72	88.44	0.32
	16	45.40	44.41	2.17
monomeric anthocyanin (mg pelargonidin-3-glucoside/100 g DW)	7	20.50	20.47	0.13
	8	14.46	14.59	0.90
	9	24.01	24.11	0.46
	16	5.04	5.10	1.26
antioxidant capacity (umol Trolox equiv/ 100 g DW)	7	4057.89	4068.77	0.27
	8	3109.09	3083.96	0.81
	9	3335.17	3329.21	0.18
	16	2216.46	2206.23	0.46

[a] See **Table 1** for sample identification. [b] Calculated as the percent ratio of the difference between expected and predicted value to the expected value.

A characteristic FTIR spectrum of phenolic extracts from potatoes with bands representative of specific functional groups is shown in **Figure 1**. A Savitzky−Golay second derivative algorithm was applied to spectra to enhance spectroscopic features by removing of baseline shifts, resolving overlapping peaks, and reducing variability between replicates (*26*). Bands were mainly associated to C−H asymmetric and symmetric stretching of CH_3 and CH_2 (2930 to 2850 cm^{-1}), C═C vibrations of aromatic systems, and C═O stretching of conjugated ketone and quinones (1640 to 1600 cm^{-1}), O−H deformation vibrations of an aromatic compound (≈ 1520 cm^{-1}), and C−O stretching of ester groups as well as O−H vibrations of phenolic groups (1160−1180 cm^{-1}). Bands with the strongest absorption were associated with C−C and C═C−C stretching of benzene ring modes (1520 cm^{-1}) and C−O stretching vibrations of primary alcohols (1014 cm^{-1}) possibly from sugar substitutions of flavonoids (*27−29, 33, 34*).

Development of Calibration Models. PLSR models generated from FTIR spectra allowed the development of calibration models for quantification of total phenolics, monomeric anthocyanins content, and antioxidant capacity. The diverse genetic pool of the potato varieties provided a wide range of levels for the specific nutritional traits evaluated in this study. Total phenolics ranged from 8.52 to 153.20 mg gallic acid equiv/100 g DW, as determined by the FC reference method. Monomeric anthocyanins content ranged from 0 to 102.71 mg anthocyanin/100 g DW, as analyzed by the pH-differential method, and antioxidant capacity ranged from 206.77 to 4394.15 μmol Trolox equiv/100 g DW of potato extracts, according to the ORAC reference method (**Table 1**). For most potato varieties, the levels obtained in this study are close to those reported for total phenolic content (3−180 mg gallic acid equiv/100 g DW; *3, 10, 11, 30*), monomeric anthocyanin content (0−200 mg anthocyanin/100 g FW; *8−11*), and ORAC values (500−1700 μmol Trolox equiv/100 g FW; *3, 10, 30*). Genotype and location (environmental conditions) have been observed as the sources of major variation in polyphenolic content among potatoes (*9−11, 31*). Positive correlations were observed among the specific potato characteristics evaluated. The Pearson correlation between total phenolics and antioxidant capacity ($r = 0.95$) was higher than between total phenolics and monomeric anthocyanins ($r = 0.75$)

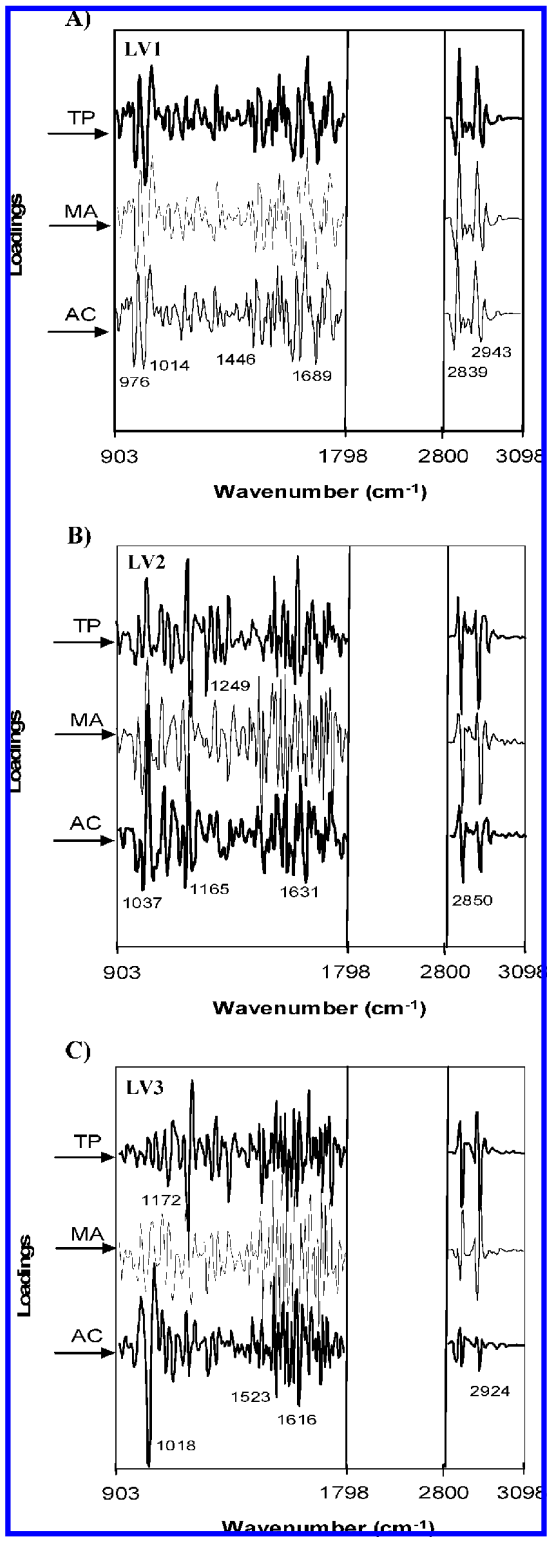

Figure 3. Nonlinear iterative partial least-squares algorithm (NIPAL) loadings for the generated PLSR calibration models of total phenolics (TP), monomeric anthocyanins (MA), and antioxidant capacity. (**A**) First latent variable (LV1) for TP, MA, and AC PLSR models; (**B**) second latent variable (LV2) for TP, MA, and AC PLSR models; and (**C**) third latent variable (LV3) for TP, MA, and AC PLSR calibration models.

and between anthocyanins and antioxidant capacity ($r = 0.63$) at a significance level (α) of 0.01. Our results support findings of Andre et al. (*3*) that determined that the presence of phenolic compounds largely accounts for the antioxidant capacity (ORAC) in potatoes. With regard to correlations between total phenolics, anthocyanins, and antioxidant capacity, Reyes et al. (*10*) reported a high positive correlation ($r \sim 0.88$) between antioxidant capacity, total phenolics, and anthocyanins levels.

Cross-validated (leave-one-out) PLSR models were developed based on IR spectra to determine total phenolics, antioxidant capacity, and monomeric anthocyanin content in various potato extracts. An optimal number of 20 factors were used in developing the calibration models that captured most of the relevant variance from complex spectra as related to specific chemical (polyphenolic compounds) levels. Generated models (**Table 2**) for total phenolics, monomeric anthocyanin, and antioxidant capacity showed excellent performance statistics with correlations (**Figure 2**) between the IR estimated concentrations and the reference analysis >0.99 for all three models (**Table 2**). The SECV, the magnitude of error expected when independent samples are predicted using the model, values were 4.17 mg gallic acid equiv/100 g DW, 0.87 mg anthocyanin/100 g DW, and 130.8 μmol Trolox equiv/100 g DW of potato powder for total phenolics, monomeric anthocyanins, and antioxidant capacity, respectively. The predictive ability of the models was evaluated on an independent set of potato extract samples. Models provided reliable predictions as compared to reference values (**Table 3**) with deviations lower than 2% from the target value.

Examination of the loading spectra indicated which bands of the spectrum were associated with the highest variation in the calibration set (*32*). **Figure 3** shows the spectroscopic contributions of the first three latent variables for all three PLSR models based on the spectroscopic information of potato extracts. The first latent variable (LV1) explained ~40% of the variance for each of the calibration models. LV1 showed similar loading plots for all three analyses (**Figure 3A**), possibly due to the high Pearson correlation values obtained for total phenolics, anthocyanins, and antioxidant capacity. Spectroscopic features of LV1 were characteristic of vibration modes related to the benzene ring (976 cm^{-1}), aromatic molecules (1689 and 1630 cm^{-1}), and primary alcohols [1014 cm^{-1}, possibly from sugar substitutions of glycosylated anthocyanins (*20*)]. Differences in the loading plot among the PLSR models were prominent for LV2 (**Figure 3B**) and LV3 (**Figure 3C**). These latter latent variables (LV) explained an additional 20% of the variance in the models and showed strong absorption bands in the 1400−1750 cm^{-1} region, possibly associated to signal from stretching modes of the C−C of the benzene ring (1400−1600 cm^{-1}) and aromatic C═O of keto- and/or carboxylic groups (1640−1750 cm^{-1}) (*36*). Additional important spectroscopic features observed for LV2 and LV3 included absorption bands in the 1000−1350 cm^{-1} region associated with the O−H in plane bend of alcohols/phenols, C−C−O asymmetric and symmetric stretching of alcohols, and C−O stretching of carboxylic acids. Phenolic compounds are characterized by hydroxylated aromatic rings with diverse substitution patterns, and the position and degree of hydroxylation in their molecules account largely for their reported antioxidant capacity (*37*). Absorption bands in the 2800−2950 cm^{-1} region, associated to C−H asymmetric and symmetric stretching of CH_3 and CH_2, were prominent in all loading plots (**Figure 3**).

Classification. Classification models for potato extracts were developed using SIMCA, a supervised classification algorithm that develops separate models based on PCA for each training set and allows category predictions by building distinct confidence regions around each class. The SIMCA Coomans plot (**Figure 4A**), a statistically derived plot that establishes boundaries between two classes, showed well-separated clusters among

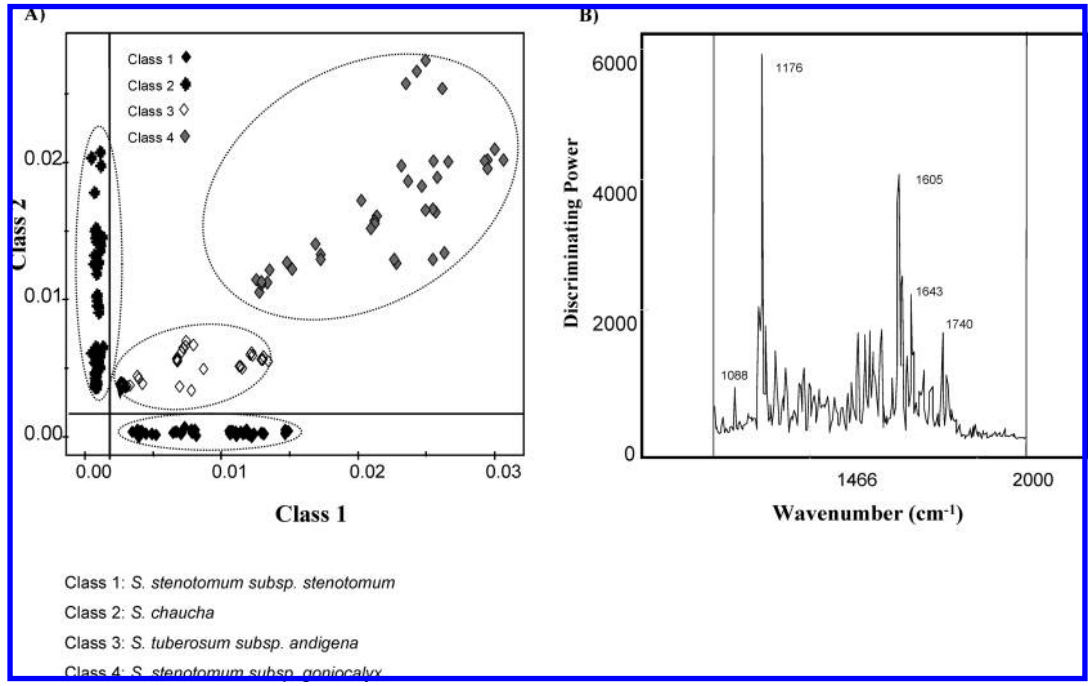

Figure 4. SIMCA Cooman's plot projection, illustrating the different clusters formed by four different potatoes species *S. stenotomum* ssp. *stenotomum* (class 1), *S. chaucha* (class 2), *S. tuberosum* ssp. *andigena* (class 3), and *S. stenotomum* ssp. *goniocalyx* (class 4). The discriminating power of the SIMCA model is also presented. (**A**) Cooman's plot class projections of transformed (second derivative, five-point window) FTIR spectra for polyphenolic extracts of potato using a three-bounce ATR-IR crystal. (**B**) Discriminating power of the SICMA model for polyphenolic extracts of potato.

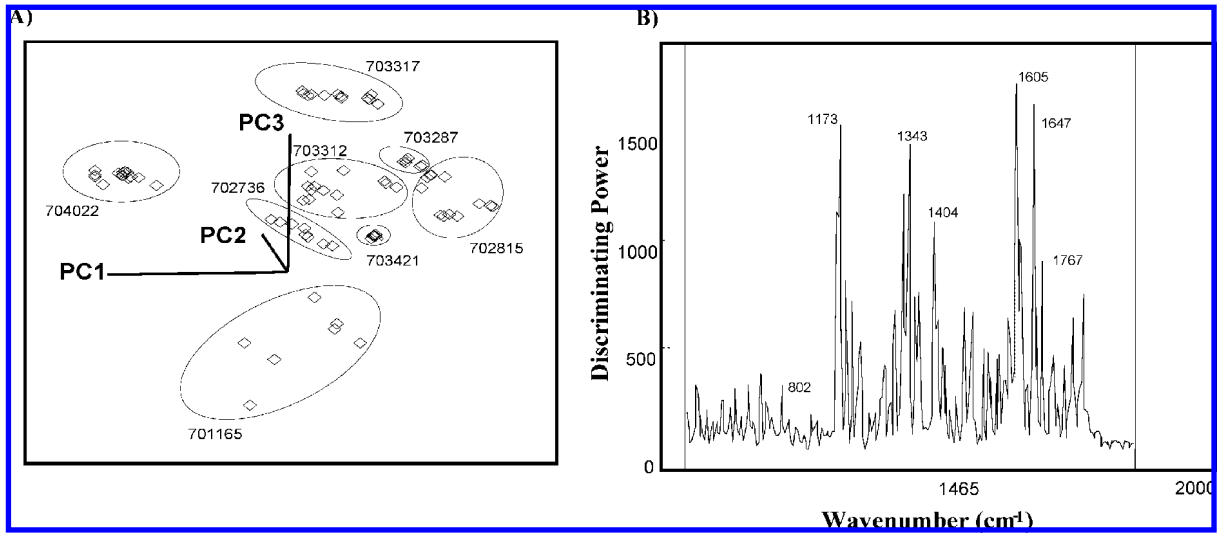

Figure 5. SIMCA class projection illustrating the different clusters formed by eight different potato varieties within the species *S. stenotomum* ssp. *goniocalyx*. The discriminating power of the SIMCA model is also presented. (**A**) SIMCA class projections of transformed (second derivative, five-point window) FTIR spectra for polyphenolic extracts of potato using a three-bounce ATR-IR crystal. (**B**) Discriminating power of the SIMCA model for polyphenolic extracts of potato.

four different species of potato: *S. tuberosum* ssp. *Andigena*, *S. stenotomum* ssp. *goniocalyx*, *S. chaucha*, and *S. stenotomum* ssp. *stenotomum*. In the plot, the two axes represent the distance of each sample from a specific category, so that each class model is drawn as a rectangle corresponding to the critical distance ($p = 0.05$) from the class (*35, 36*). Any sample having a distance to the corresponding centroid greater than the critical distance is considered as being outside the specific category. Moreover, samples plotted onto the lower left square of the diagram are assigned to both classes (*35, 36*). By using the IR spectroscopic region of 1800 to 900 cm^{-1}, the Cooman's plots showed a marked separation of the classes (potato species) based on principal components (**Figure 4A**) and demonstrated that the potato species evaluated did not share multivariate space, providing validation for the class separation. Another statistic that provides information regarding the ability of SIMCA to discriminate among classes is the interclass distance. Interclass distance between groups of samples is a measure of the distance between classes in the multivariate space. Interclass distances >3 are generally considered significant (*21*). The SIMCA model showed interclass distance (IC) values ranging from 6.79 to 22.97 for discrimination among potato species, further demonstrating the ability of IR spectroscopy combined with chemometrics to achieve reliable resolution of unique spectroscopic markers for differentiation of samples at the species level. The predictive ability of SIMCA model at species level was tested

by using four potato samples, included in the data set utilized to create SIMCA models, which were labeled as "unknowns". The SIMCA model successfully predicted the species of all masked samples. The discriminating power plot (**Figure 4B**) showed that major bands influencing the discrimination of classes were related to C−H bending in plane of the benzene ring (1176 cm^{-1}), C═C stretching of aromatic rings (1605 and 1643 cm^{-1}), and C═O stretching of unconjugated ketone and/or aldehyde groups (1740 cm^{-1}).

Once the prediction ability at species level was confirmed, a new SIMCA model at potato variety level (**Figure 5A**) was developed for the species *S. stenotomum* ssp. *goniocalyx*. The scores plot, a projection of the original data onto the principal components, allowed the visualization of well-separated clustering among the eight varieties. Interclass distances ranged from 4.59 to 28.02, indicating the ability of SIMCA to predict samples at the variety level. The discriminating power (**Figure 5B**) showed that in addition to the previously described bands at 1176, 1605, 1643, and 1740 cm^{-1} that were important for species classification, additional bands at 1404 and 1343 cm^{-1} that have been associated to C−H deformation vibration stretching of aromatic ring and O−H deformation vibration of alcohols, respectively, arose as important bands in understanding the clustering of samples.

Our results show the ability of IR spectroscopy combined with multivariate analysis to determine simultaneously multiple nutritional traits in a fast, simple, and accurate procedure, based on unique spectroscopic information associated with potato extracts. Furthermore, once the instrument is purchased, there is minimal operational cost involved in performing the technique. PLSR models generated from IR spectra provided excellent predictive abilities for total phenolics, monomeric anthocyanins content, and antioxidant capacity. Furthermore, IR allowed for the rapid characterization of potato varieties at species and variety levels. The total time required for both quantitative and qualitative sample analysis was less than 5 min, as compared to the several hours required for total phenolics, monomeric anthocyanins, and antioxidant capacity analyses using reference methods. High-throughput potato screening techniques will support the breeder's ability to develop potato varieties with enhanced nutritional traits through the selection of sustainable crops and the development of robust systems to support rural and urban agriculture.

LITERATURE CITED

(1) International Year of the Potato. Hidden treasure. Initiative launched by The United Nations and FAO, 2008; http://www.potato2008.org/en/potato/index.html (accessed May 2008).

(2) Food and Agriculture Organization of the United States—FAO. Buried treasure: The potato, 2006; www.fao.org/ag/magazyine/0611sp1.htm (accessed May 2008).

(3) Andre, C.; Ghislain, M.; Bertin, P.; Oufir, M; Herrera, M.; Hoffmann, L.; Hausman, J. F.; Larondelle, I.; Evers, D. Andean potato cultivars (*Solanum tuberosum* L.) as a source of antioxidant and mineral micronutrients. *J. Agric. Food Chem.* **2007**, *55* (2), 366–378.

(4) Lachman, J.; Hamouz, K. Red and purple coloured potatoes as significant antioxidant source in human nutrition—A review. *Plant Soil Environ.* **2005**, *51*, 477–482.

(5) Jansen, G.; Flamme, W. Coloured potatoes (*Solanum tuberosum* L.)—Anthocyanin content and tuber quality. *Genet. Resour. Crop Evol.* **2006**, *53*, 1321–1331.

(6) Campos, D.; Noratto, G.; Chirinos, R.; Arbizu, C.; Roca, W.; Cisneros-Zevallos, L. Antioxidant capacity and secondary metabolites in four species of Andean tuber crops: Native potato (*Solanum* sp.), mashua (*Tropaeolum tuberosum* Ruiz & Pavón), oca (*Oxalis tuberose* Molina) and olluco (*Ullucus tuberosus* Caldas). *J. Sci. Food Agric.* **2006**, *86*, 1481–1488.

(7) Brown, C. R.; Wrolstad, R. E.; Durst, R.; Yang, C.; Clevidence, B. A. Breeding studies in potatoes containing high concentrations of anthocyanins. *Am. J. Potato Res.* **2003**, *80*, 241–250.

(8) Rodriguez-Saona, L. E.; Giusti, M. M.; Wrolstad, R. E. Anthocyanin pigment composition of red-fleshed potatoes. *J. Food Sci.* **1998**, *63*, 458–465.

(9) Handique, J. G.; Baruah, J. B. Polyphebolic compounds: An overview. *React. Funct. Polym.* **2002**, *52*, 163–188.

(10) Reyes, L. F.; Miller, J. C.; Cisneros-Zevallos, L. Antioxidant capacity, anthocyanins and total phenolics in purple and red-fleshed potato (*Solanum tuberosum* L.) genotypes. *Am. J. Potato Res.* **2005**, *82*, 271–277.

(11) Lewis, C.; Walker, J.; Lancaster, J.; Sutton, K. Determination of anthocyanins, flavonoids and phenolic acids in potatoes. I: Coloured cultivars of *Solanum tuberosum* L. *J. Sci. Food Agric.* **1998**, *77*, 45–57.

(12) Nestel, P.; Bouis, H.; Meenakshi, J. V.; Pfeiffer, W. Biofortification of staple food crops. In: Symposium food fortification in developing countries. *J. Nutr.* **2006**, *136*, 1064–1067.

(13) Halim, Y.; Schwartz, S.; Francis, D.; Baldauf, N.; Rodriguez-Saona, L. E. Direct determination of lycopene content in tomatoes (*Lycopersicon esculentum*) by attenuated total reflectance infrared (ATR-IR) spectroscopy and multivariate analysis. *JAOAC Int.* **2006**, *89*, 1257–1262.

(14) Karoui, R.; Mounem, A.; Dufour, E.; Pillonel, L.; Shcaller, E.; Picque, D.; De Baerdemaeker, J.; Bosset, J. A comparison and joint use of NIR and MIR spectroscopic methods for the determination of some parameters in European Emmental cheese. *Eur. Food Res. Technol.* **2006**, *223*, 44–50.

(15) Innawong, B.; Mallikarjunan, P.; Irudayaraj, J.; Marcy, J. The determination of frying oil quality using Fourier transforms infrared attenuated total reflectance. *Lebensm. Wiss. Technol.* **2004**, *37*, 23–28.

(16) Paradkar, M.; Irudayaraj, J. Determination of cholesterol in dairy products by infrared techniques: 1. FTIR spectroscopy. *Int. J. Dairy Technol.* **2002**, *55*, 3.

(17) Naumann, D. FT-infrared and FT-Raman spectroscopy in biomedical research. Practical spectroscopy series. In *Infrared and Raman Spectroscopy of Biological Materials*; Gremlich, H.-U., Yan, B., Eds.; Marcel Dekker: New York, United States, 2001; Vol. 24, pp 323−377.

(18) Giusti, M.; Wrosltad, E. Characterization and measurement of anthocyanins by UV-visible spectroscopy. In *Handbook of Analytical Chemistry*; Wrolstad, E., Acree, T., Decker, E., Penner, M., Reid, D., Schwartz, S., Shoemaker, C., Smith, D., Sporns, P., Eds.; John Wiley & Sons, Inc.: New Jersey, United States, 2005; Unit F1.2, pp 19−31.

(19) Waterhouse, A. Determinations of total phenolics. In *Handbook of Food Analytical Chemistry*; Wrolstad, E., Acree, T., Decker, E., Penner, M., Reid, D., Schwartz, S., Shoemaker, C., Smith, D., Sporns, P., Eds.; John Wiley & Sons, Inc.: New Jersey, United States, 2005; pp 463−470.

(20) He, J.; Rodriguez-Saona, L.; Giusti, M. Mid-infrared spectroscopy for juice authentication—Rapid differentiation of commercial juices. *J. Agric. Food Chem.* **2007**, *55*, 4443–4452.

(21) Bjorsvik, H. R., Martens, H. Data analysis: Calibration of NIR instruments by PLS regression. In *Handbook of Near Infrared Analysis*; Burns, D., Ciurczak, E., Eds.; Marcel Dekker: New York, United States, 1992.

(22) Cao, G.; Alessio, H.; Cutler, R. Oxygen-radical absorbance capacity assay for antioxidants. *Free Radical Biol. Med.* **1993**, *14*, 303–311.

(23) Wold, S.; Albano, C.; Blomquist, G.; Coomans, D.; Dunn, W. J., III; Edlund, U.; Eliasson, B.; Hellberg, S.; Johansson, E.; Norden, B.; Sjöström, M.; Söderström, B.; Wold, H. Pattern recognition by means of disjoint principal component models (SIMCA). In *Proc. Symp. Anv. Stat.*; Haskuldson, A., Eds.; NEUCC, RECAU, and RECKU: Copenhagen, 1981; pp 475−487.

(24) Dunn, W. J., III; Wold, S. SIMCA pattern recognition and classification. In *Chemometric Methods in Molecular Design*; Van de Waterbeemd, H., Ed.; VCH Publishers: New York, 1995; pp 179−193.

(25) Rodriguez-Saona, L. E.; Wrolstad, R. E. Extraction, isolation, and purification of anthocyanins. In *Handbook of Food Analytical Chemistry*; Wrolstad, E., Acree, T., Decker, E., Penner, M., Reid, D., Schwartz, S., Shoemaker, C., Smith, D., Sporns, P., Eds.; John Wiley & Sons, Inc.: New Jersey, United States, 2005; pp 7−18.

(26) Escribano, M. T.; Santos-Buelga, C. Polyphenol extraction from foods. In *Methods in Polyphenol Analysis*; Santos-Buelga, C., Williamson, G., Eds.; The Royal Society of Chemistry: Cambridge, United Kingdom, 2003; Chapter 1.

(27) Xu, F.; Sun, J.; Sun, R.; Fowler, P.; Baird, M. Comparative study of organosolv lignins from wheat straw. *Ind. Crop Prod.* **2006**, *23*, 180–193.

(28) Lam, H.; Proctor, A.; Howard, L.; Jin-Cho, M. Rapid fruit extracts antioxidant capacity determination by Fourier transform infrared spectroscopy. *J. Food Chem.* **2005**, *70*, C545–C549.

(29) Sudiyani, Y.; Imamura, Y.; Doi, S.; Yamauchi, S. Infrared spectroscopic investigations of weathering effects on the surface of tropical Wood. *J. Wood Sci.* **2003**, *49*, 86–92.

(30) U.S. Department of Agriculture. Oxygen radical absorbance capacity (ORAC) of selected foods. Agricultural Research Service, 2007; www.ars.usda.gov/nutriendata (accessed May 2008).

(31) Rodriguez-Saona, L. E.; Wrolstad, R. E.; Pereira, C. Glycoalkaloid content and anthocyanin stability to alkaline treatments of red-fleshed potato extracts. *J. Food Sci.* **1999**, *64*, 445–450.

(32) Haaland, D. M.; Thomas, E. V. Partial least-squares methods for dpectral snalysis 1: Relation to other multivariate valibration methods and the extraction of multivariate information. *J. Anal. Chem.* **1988**, *60*, 1193–1202.

(33) Fakharedine, N.; El Hajjouji, H.; Ait Baddi, G.; Revel, J. C.; Mafidi, M. Chemical and spectroscopic analysis of organic matter transformation during aerobic digestion of olive-mill waste-waters. *Process Biochem.* **2006**, *41*, 398–404.

(34) Smith, B. *Infrared Spectroscopic Interpretation*; CRC Press LLC: Boca Raton, FL, 2000.

(35) Marini, F.; Magri, A.; Bucci, R.; Balestrieri, F.; Marini, D. Class-modeling techniques in the authentication of Italian oils from Sicily with a protected denominations of origin (PDO). *Chemom. Intell. Lab.* **2005**, *80*, 140–149.

(36) Blanco-Gomis, D.; Herrero-Sanchez, L.; Mangas, A. Characterization of apple cider cultivars by chemometric techniques using data from high-performance liquid chromatography and flow-injection analysis. *Analyst* **1998**, *123*, 1187–1191.

(37) Wang, H.; Cao, G.; Prior, R. Oxygen radical absorbing capacity of anthocyanins. *J. Agric. Food Chem.* **1997**, *45*, 304–309.

Received for review June 4, 2008. Revised manuscript received August 27, 2008. Accepted August 27, 2008.

JF801716B