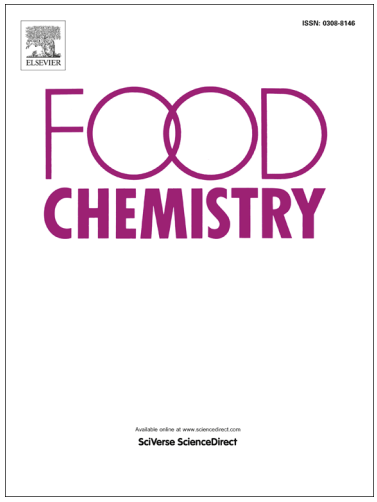


Detection of quinoa flour adulteration by means of FT-MIR spectroscopy combined with chemometric methods

Silvio D. Rodríguez, Guido Rolandelli, M. Pilar Buera

PII: S0308-8146(18)31559-0
DOI: https://doi.org/10.1016/j.foodchem.2018.08.140
Reference: FOCH 23482

To appear in: *Food Chemistry*

Received Date: 27 March 2018
Revised Date: 30 July 2018
Accepted Date: 31 August 2018

Please cite this article as: Rodríguez, S.D., Rolandelli, G., Pilar Buera, M., Detection of quinoa flour adulteration by means of FT-MIR spectroscopy combined with chemometric methods, *Food Chemistry* (2018), doi: https://doi.org/10.1016/j.foodchem.2018.08.140

Detection of quinoa flour adulteration by means of FT-MIR spectroscopy combined with chemometric methods.

Silvio D. Rodríguez[1,2,*], Guido Rolandelli[1], M. Pilar Buera[1]

1 Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales. Buenos Aires, Argentina.

2 CONICET – Universidad de Buenos Aires. Instituto de Biodiversidad y Biología Experimental y Aplicada (IBBEA). Buenos Aires, Argentina.

Abstract

Quinoa flour has been receiving an increasing attention as a substitute for wheat flour in bread formulations due to immuno-nutritional features. This growing interest in quinoa has increased the demand and consequently the prices, being a target for possible adulterations with cheaper cereals. Fourier transform Mid-infrared spectroscopy (FT-MIR) was used in the present work as a fingerprinting technique to detect the presence of three adulterants (soybean, maize and wheat flours). Partial least squares discriminant analysis (PLS-DA) and soft independent modelling of class analogy (SIMCA) models were used to classify pure from adulterated samples. 414 samples were measured, including pure quinoa flour, pure adulterant flours and adulterated quinoa flours using three different proportions (10, 5 and 1% w/w). PLS-DA showed better classification results than SIMCA, with error rates from 2 to 8% for the three strategies used to detect the presence of adulterants.

Keywords: Quinoa flour adulteration; FT-IR; chemometric methods; PLS-DA; SIMCA

* Corresponding author:
Dr. Silvio David Rodríguez.
Instituto de Biodiversidad y Biología Experimental y Aplicada (IBBEA). CONICET – Universidad de Buenos Aires. Intendente Güiraldes 2160, Pabellón 2, 4to Piso, Ciudad Universitaria, Buenos Aires, Argentina.
Tel: +5411 52859037
Mail: silviodavidrodriguez@gmail.com
ORCID number: 0000-0001-5559-8124

1.0 Introduction

Quinoa (*Chenopodium quinoa* Will.) is a dicotyledonous plant, considered a pseudocereal and native in the Andean region of South America. It is a granifer species domesticated by indigenous people about 5000 years ago (Filho et al., 2017). Quinoa is still widely cultivated in several countries in South America and it has been introduced and cultivated for commercial purposes in Europe, North America, Asia and Africa (Ferreira, Pallone, & Poppi, 2015). It is considered a crop with a great nutritional value and unique resistance to weather climate and soil conditions. The most studied, economically relevant and scientifically interesting edible part of the plant are its seeds, which are small, flat, and oval-shaped, varying in colors, from pale yellow to light red or even black (Aluwi, Murphy, & Ganjyal, 2017). Quinoa grains contain more protein than other cereals, with high levels of methionine and lysine. It contains a substantial amount of fiber, minerals, such as calcium and iron, and polyphenols (Nowak, Du, & Charrondière, 2016). Moreover, quinoa flour (among other Andean crop flours), has been received an increasing attention as a substitute to wheat flour in bread formulations due to its immuno-nutritional features, such as improving intestinal absorption of iron or modulating the hepatic production of inflammatory biomarkers (Laparra & Haros, 2018).

In recent years, the raising interest in quinoa, due to its nutritional value, has promoted a demand for both, quinoa seeds and quinoa flour (González-Muñoz, Montero, Enrione, & Matiacevich, 2016). This fact has triggered the exportation of quinoa being traded from 600 ton in 1992 to 68000 ton in 2015 and quinoa prices quadrupled from 1 USD/kg in 2007 to an average of 4 USD/kg in the last decade (ALADI and FAO, 2014; Ruiz et al., 2014; Stevens, 2017). Furthermore, the increment in the consumption of quinoa in the European Union has led a growing concern about legislation, safety policies, microbiological, chemical and physical controls (Ojinnaka, 2016).

The adulteration of a food product is a fraud that involves a deliberate addition, substitution or removal of a food ingredient without purchaser's knowledge, for economic gain. Food fraud is a cause of public health risk. Thus, consumers, governments and industries must invest in preventing

food adulteration (Su & Sun, 2017; Verdú et al., 2016). Food adulteration is not new, and a great number of cases were reported in diverse products: fresh or preserved meat and fish, edible oils, processed fruits and vegetables, dairy products, sauces, soups, syrups and honey, spices, or even cereal grains and flours (Collins, 1993; Knödler, Most, Schieber, & Carle, 2010; Verdú et al., 2016; Ziegler et al., 2016). Several efforts have been made to assess the quality of food products by the detection of possible exogenous added ingredients. A widespread strategy for this assessment is known as the fingerprint approach, which is based in obtaining a signal or image by certain analytical tools, such as electrophoresis fingerprinting (protein or DNA detection techniques), spectral fingerprinting (Nuclear magnetic resonance, infrared spectroscopy, UV-Vis spectroscopy, Raman spectroscopy, mass spectrometry and spectral imaging) or chromatographic fingerprinting (gas chromatography, liquid chromatography or high performance liquid chromatography) (Ellis et al., 2012; Manning & Soon, 2014; Ropodi, Panagou, & Nychas, 2016).

One of the fingerprinting techniques widely used to characterize food ingredients and detect possible adulterants is Fourier transform infrared spectroscopy (FT-IR) in the middle range (450 to 4000 cm^{-1}, FT-MIR) or near range (4000 cm^{-1} to 10000 cm^{-1}, FT-NIR). FT-MIR is a rapid, non-destructive technique and it is considered a green analytical method (Pallone, Caramês, & Alamar, 2018). FT-MIR monitors the fundamental vibrational and rotational stretching modes of molecular bonds, which produce a chemical profile of the sample and provide a greater amount of chemical information than FT-NIR (Lohumi, Lee, Lee, & Cho, 2015). Thus, FT-MIR was used to detect adulterants in raw milk (Botelho, Reis, Oliveira, & Sena, 2015), hazelnut oil (Ozen & Mauer, 2002), raw and cooked beef, honey and jams, among many others (Karoui, Downey, & Blecker, 2010; Rodriguez-Saona & Allendorf, 2011). Even more, FT-MIR has been used to classify cereal flours, such as wheat, oats and buckwheat (Cocchi et al., 2004), to discriminate different types of wheat flours from rye and triticale flours (Sujka, Koczoń, Ceglińska, Reder, & Ciemniewska-Żytkiewicz, 2017) or to determine components of quinoa flour (González-Muñoz et al., 2016).

Spectroscopic techniques, such as FT-MIR, are enhanced in a greater way when used in combination with multivariate statistical methods, also known as *chemometric* methods, in which each measurement is associated with multiple variables (Pomerantsev & Rodionova, 2012). Spectroscopic variables are the intensities associated to the frequency of each obtained spectrum for each sample. Usually, these variables are highly correlated and may also contain systematic variations, such as, additive compensations due to chemical interferences or related to the equipment (drift). These complications could be overcome by using chemometric methods, reducing the negative impact of non-relevant variables, improving the interpretation of the data and/or by obtaining simpler and more robust models (van den Berg, Lyndgaard, Sørensen, & Engelsen, 2013).

Two supervised chemometric methods, partial least squares discriminant analysis (PLS-DA) and soft independent modelling of class analogy (SIMCA) were extensively used for classification purposes of food samples scanned by spectroscopic methods (Berrueta, Alonso-Salces, & Héberger, 2007; López, Trullols, Callao, & Ruisánchez, 2014; Luna et al., 2017). In this work, FT-MIR in combination with PLS-DA or SIMCA models were used to differentiate pure quinoa flours from those adulterated with three other flours (soybean, maize and wheat), using three different adulterant proportions (1, 5 and 10% w/w).The aim of this study is to detect the three possible adulterants of quinoa flour with a reasonable low classification error for the three adulterant proportions.

2.0 Materials and methods

2.1 Samples

Quinoa flours were obtained from nine different brands (Q1 to Q9) of quinoa seeds purchased at a local market. Seeds were grinded into a benchtop mill during 30 seconds at room temperature to obtain quinoa flours. Pure samples of soybean, maize and wheat flours were obtained from a local

milling industry (Dietética Científica SA, Buenos Aires, Argentina). The twelve flours (9 from quinoa and the 3 adulterants) were passed through a sieve (mesh 420 µm) prior analysis and sample preparation. Flours were analyzed to quantify the amount of major dietary constituents following the methods of AOAC numbers:925.10 for moisture content, 920.87 for total protein, 923.05 for lipids and 923.03 for ash (AOAC, 2016). Total carbohydrate content was determined by difference as: 100 – [weight in grams × (protein + lipids + moisture + ashes)] in 100 g of quinoa flour. Table 1 shows the average percentage (dry basis except for moisture) of three independent determinations with the standard deviation value of each constituent. Three adulterations were made for each quinoa flour brand (Q1 to Q9) adding 10% w/w, 5% w/w and 1% w/w per adulterant flour. For example, in the case of soybean, 27 samples were produced, 9 for the adulteration of each quinoa brand with 10% of soybean flour, 9 for 5% and 9 for 1%. This pattern was kept with the other two adulterant flours (maize and wheat), giving a total of 81 adulterated samples. For adulterated samples both flours were vigorously mixed in a mortar until homogeneity. To assure the homogeneity, all samples were passed through a sieve with a mesh of 420 µm.

2.2 Fourier Transform Middle Infrared Spectroscopy (FT-MIR)

Each type of sample, pure quinoa flour from Q1 to Q9, adulterated samples with soybean (10, 5 and 1%), adulterated samples with maize (10, 5 and 1%), adulterated samples with wheat (10, 5 and 1%) and the three pure adulterant flours (Soybean, maize and wheat) were scanned at least by quadruplicate in a FT-MIR (Spectrum 400, Perkin Elmer Inc., Shelton CT, USA) equipped with a DTGS detector and attenuated total reflectance accessory (ATR, PIKE Technologies, Inc., Madison WI, USA). Each powdered sample was placed on the ATR crystal (diamond/ZnSe one reflectance at 45°) and pressed until the desired signal intensity. Spectra were collected from 600 to 4000 cm^{-1} with a resolution of 4 cm^{-1} and an accumulation of 64 spectra per sample. Every spectrum was base-line corrected, transformed to absorbance units and normalized using Spectrum Software ver. 6.3

(Perkin Elmer, Inc.). Due to the ATR crystal absorption all the spectra showed a highly noisy region from 1800 to 2500 cm^{-1}, so this region was not considered for further data analysis.

2.3 Organization of data sets for the classification using chemometric methods

The number of measurements acquired were 414 and FT-MIR spectra are available in a Microsoft Excel file in the *Supplementary material section*. With these measurements, different data sets were constructed before classification analysis. The first approach was to detect three separate cases, one for each adulterant. These cases were labeled as QS1 (Quinoa + Soybean), QM1 (Quinoa + Maize) and QW1 (Quinoa + Wheat). The purpose of the first approach (QS1, QM1 and QW1) was to classify data in three classes, pure quinoa flour, pure adulterant flour and adulterated quinoa flours in three different proportions: 10, 5 and 1% w/w of adulterant. The reason of including pure flours of the adulterants was to test the feasibility of FT-MIR to detect the differences between pure quinoa flour from pure adulterant flour. In case this did not occur, the classification method would find many difficulties to discriminate pure quinoa flour from the adulterated quinoa flours.

The second approach in the present work was to detect if a measurement is recognized as pure quinoa flour or adulterated quinoa flour selecting the possible adulterant flour without the class of pure adulterant flour. For these purposes three data sets were built: QS2, QM2 and QW2 which are the same from the first approach excluding pure adulterants.

Finally, the major challenge for this study was to detect pure quinoa flour from adulterated quinoa flour with any of the three adulterants (soybean, maize or wheat). The first case was labeled as QSMW1 containing the three proportions of adulterants (10, 5 and 1% w/w). The second, labeled as QSMW2 included only two proportions (10 and 5% w/w). And finally, the third case, labeled as QSMW3, which included only the adulterated quinoa flour only with 10% w/w.

The number of measurements for each type of sample is shown in Table SM1 (*Supplementary material section*).

2.4 Partial least squares discriminant analysis (PLS-DA)

PLS-DA is a supervised chemometric method widely used for pattern recognition of FT-MIR spectra and it is based on the PLS regression algorithm. In PLS-DA, a multivariate X matrix was constructed using the spectra for all the samples to be tested. The X matrix (with size of $I \times J$) contains I samples per J variables (FT-MIR wavenumbers, also known as original variables). As the PLS-DA is a supervised method, the information of the class for each sample was arranged in the Y binary variable previously defined. The goal of the algorithm is to find a new space (defined with new latent variables, also known as components) for each (X and Y) trying to maximize the linear relation between the components from X and Y and to minimize the root mean square error (RMSE). The parameters to optimize are the scores, the loadings and the residuals, which can be arranged into matrices. The optimization of these parameters could be described with two matrix decompositions as follows and it is performed by the algorithm in a first step, also called calibration step (also known as fitting step or training step): $X = TP' + E$ and $Y = UQ' + F$, where X is the spectra matrix ($I \times J$), T($I \times N$) is the matrix of scores for X, P'($J \times N$) is the transpose matrix of loadings for X, E($I \times J$) is the matrix of residuals for X, U($I \times N$) is the matrix of scores for Y(I,K), Q'(K,N)is the matrix of loadings for Y and F($I \times K$) is the matrix of residuals for Y. In PLS-DA the relevant sources of data variability are modelled by the components (score matrices), which are linear combinations of the original variables and allow graphical visualization (in case a reduction of the dimension occurs). The loading values are the coefficients of variables in the linear combination which determine the components and they can be interpreted as the influence of each original variable on each latent variable (component). The second step of the algorithm is the validation step (after calibration step), which consists in using the regressors obtained in the fitting

step and predict the class for samples not included into the fitting step. After the validation step, the algorithm compares the predicted class with the real class for each sample and the errors made by the prediction were used to calculate several error parameters described in *Section 2.6*.In this study, the number of components (*N*) was selected by a previous cross validation error approach, in which PLS-DA was run with a variable number of components (from 1 to 20) and the number of components with minimum validation error was selected. The validation procedure selected in all the PLS-DA cases were venetian blinds cross validation with 10 groups. PLS-DA was run into Classification Toolbox ver. 5.0 (Ballabio & Consonni, 2013) under GNU Octave for windows ver. 4.2.1.

2.5 Soft independent modelling class analogy (SIMCA)

SIMCA is a supervised chemometric method for pattern recognition based on principal component analysis (PCA). In fact, a PCA was modeled for each class from the proposed data set, and only the relevant number of components was used for each class (reaching the minimum cross validation error). The resulting principal components (PCs) were used to fix the boundaries of each class (by calculating Q-statistic and Hotelling T^2) and to assign a new sample (not included in the training step) into a class. SIMCA determines the class distance, calculated as the Euclidean distance from the principal component models. SIMCA results could be plotted in the so-called *Coomans* plot, which shows the distances for each sample from the model for class 1 against that from class 2. Four zones are defined with a confidence level (usually 95%), the zone of samples assigned to class 1, the zone of samples assigned to class 2, an overlap zone for both classes and the zone far for both classes. This is considered a soft method because a sample can be classified into a class or not assigned to any class and it can be extended for multiple dimensions, in case that more than two classes were modelled. The outcomes for SIMCA model (as the same way of PLS-DA) are the samples assignments to each class for both calibration and validation steps (Berrueta et al.,

2007; Luna, Pinho, & Machado, 2016). In a similar way used in PLS-DA, the number of components (*N*) was selected by a previous cross validation error approach, in which the algorithm was run with a variable number of components (from 1 to 20) and the number of components with minimum validation error was selected for each class. The validation procedure selected in all the SIMCA cases were venetian blinds cross validation with 10 groups. SIMCA was run into Classification Toolbox ver. 5.0 (Ballabio & Consonni, 2013) under GNU Octave for windows ver. 4.2.1.

2.6 Classification performance parameters for PLS-DA and SIMCA

For both models (PLS-DA and SIMCA), two vectors of assignations for each sample were calculated, one corresponding to the calibration step and the second corresponding to the validation step. With this assignation vector and the real class information for each sample, a confusion matrix for each step is computed. The confusion matrix (also known as contingency table) is a square matrix with size of $G \times G$ (where G is the number of classes, and an extra column can be added in the case of samples not assigned to any class)and shows the outcomes of the classification model. Each element on the main diagonal of the confusion matrix represents the number of correct class assignations, while off-diagonal elements represent the errors in classification. Based on the confusions matrix (calibration and validation), it is possible to calculate additional performance parameters for each class, such as sensitivity (SEN), specificity (SPEC), precision (PREC) and parameters for calibration or validation steps, such as accuracy (ACC), error and non-errors rates, ER and NER, respectively. SEN represents the rate of correctly recognized samples to a class and is calculated as TP/(TP + FN), where TP is the number of positive cases correctly classified and FN are the false negative assignments. SPEC represents the percentage that rejects samples of all other classes and is calculated as TN/(FP + TN), where TN is the number of negative cases correctly classified and FP are the false positive assignations. PREC is defined as the purity of a class, which

is the ability to avoid wrong predictions in that class and is defined as TP/(TP + FP). SEN, PREC and SPEC values vary between 1 and 0, representing a perfect classification and no class discrimination, respectively. ACC represents the rate of false and positive samples correctly classified among all positive and negative samples and is calculated as (TP + TN) / (TP + TN + FP + FN). The non-error rate (NER) is the average value of the sensitivity for all the classes, and ER = 1 - NER. ACC, NER and ER are considered global parameters giving information of the overall classification for each algorithm step (calibration or validation). NER and ACC parameters take values from 0 to 1, indicating a perfect classification when the values are 1 (Ballabio, Grisoni, & Todeschini, 2018; Luna et al., 2016).

Another analysis to visualize the relation between SEN and SPEC is Receiver Operating Characteristic curves (ROC curves). The first ROC curve shows a plot of SEN and SPEC values varying the threshold limit (boundary limits of the classes), where the optimal threshold is the value showing the highest SEN and SPEC values. The second ROC curve shows the relation between SEN vs. 1-SPEC paired values for all the threshold limits studied. The area under the curve (AUC), is used to compare the global or overall performance of different classification methods and AUC is close to 1 when the model shows maximum classification ability (López, Colomer, Ruisánchez, & Callao, 2014).

3.0 Results

The major dietary constituents (moisture, protein, lipids, ash and carbohydrates) of quinoa, soybean, maize and wheat flours are shown in Table 1. Nine brands of quinoa seeds were used to made quinoa flour, showing a protein content in a range between 13.81 and 20.67%, a lipid content from 6.94 to 11.98%, ash content from 2.31 to 5.55% and carbohydrate content from 65.28 to 75.86%. The variations in the constituents observed for quinoa seeds could be due to different seed cultivars and growing, harvest or post-harvest conditions. The three flours used as adulterant in the

present study were soybean, maize and wheat and each one has a particular constituent profile, different from quinoa flours. Soybean flour showed greater protein and lipids contents and lower values for carbohydrate and ash. On the other side, maize flour showed lower values of proteins and lipids, but higher for carbohydrates and ash, in comparison with quinoa flours. Finally, wheat flour showed lower protein, lipid and ash than quinoa flours, but a higher value of carbohydrate content.

Samples with the three adulterant flours (soybean, maize and wheat) were prepared as indicated in *Section 2.1*. Three different classes of adulterated samples were prepared, with 10, 5 and 1% w/w of each adulterant flour. FT-MIR measurements were carried out over pure quinoa samples, adulterated samples (three adulterants in three different proportions) and pure soybean, maize and wheat flours.

This study was divided into the approaches or strategies mentioned in *Section 2.3*, which are: detection of a specific adulterated quinoa sample with soybean, maize or wheat, in different proportions (10, 5 and 1% w/w) in the same data set with pure flour adulterants (Cases QS1, QM1 and QW1), detection of a specific adulterated quinoa sample (same adulterants and proportions used for the three cases mentioned above) without the presence of pure adulterants samples (Cases QS2, QM2, and QW2) and finally, the cases for the detection of the three adulterants used in quinoa flours in the proportions mentioned above (QSMW1, QSMW2, QSMW3).

3.1 Cases QS1, QM1 and QW1

The average FT-MIR spectra for QS1, QM1 and QW1 cases are displayed in Figure 1 a), b) and c), respectively. Figure 1a) shows the average of 37 spectra measurements for pure quinoa, the average of 115 adulterated quinoa flours with three different proportions of soybean flour and the average of 10 pure soybean flours. Averaged spectra are shown due to the high number of spectra obtained and to better visualize the main differences without spectra overlapping (the full spectra for QS1, QM1 and QW1 cases are included in *Supplementary material section*: Figure SM1). The spectral absorption bands observed in Figure 1 were described in previous works. The bands at

2800 to 3040 cm^{-1}, associated to C–H and $C–H_2$ symmetric and asymmetric stretching and the band with a maximum at 1745 cm^{-1} (stretching of the ester carbonyl group) were mainly attributed to bond vibrations of the lipids in the flours (Roa, Santagapita, Buera, & Tolaba, 2014). The bands associated to protein bond vibrations are those with the maximum at 1640 cm^{-1} (C=O stretching and N–H bending) and at 1540 cm^{-1} (C–N stretching, C–O stretching and C–C stretching), also known as amide I and amide II bands, respectively (Guzmán-Ortiz et al., 2015). In addition, a small broad band from 1200 to 1500 cm^{-1} is represented by CH_2OH side chain related mode, C–O–H bending, $C–H_2$ twisting, $C–H_2$ bending and C–O–O stretch. Moreover, spectra show a strong absorption band, from 900 to 1200 cm^{-1} due to C–O and C–C stretching (1145 cm^{-1}), C–O–H bending (1080 cm^{-1}) and C–H bending (1000 cm^{-1}), mainly present in carbohydrates. Additionally, below 900 cm^{-1}, small bands can be observed due to skeletal modes of the pyranose ring from carbohydrates (Warren, Gidley, & Flanagan, 2016). The broad absorption band from 3010 to 3750 cm^{-1}, attributed principally to hydroxyl groups due to O–H stretching, was not considered in this study, because it is a band that tends to show intensity variations due to moisture content of the sample (González-Muñoz et al., 2016). Due to the O–H stretching absorption band (3010-3750 cm^{-1}) and the noisy region of high absorption due to diamond/ZnSe crystal of ATR accessory, all spectra were plotted and analyzed in the regions of 600 to 1800cm^{-1} and 2750 to 3050cm^{-1}.

Figure 1 a) shows the three averaged spectra of pure quinoa, adulterated quinoa (10, 5 and 1% w/w) and pure soybean flours (case QS1). Pure soybean flour shows higher absorbance values in all frequencies, which is mainly associated with lipids (2800 to 3040 cm^{-1}) and higher and sharper bands at 1640 and 1540 cm^{-1} (amide I and II, respectively) associated to proteins. In addition, several intense bands are present in the region of 1200 to 1500 cm^{-1} (CH_2OH, C–O–H, $C–H_2$, $C–H_2$ and C–O–O vibrational modes) and shifted and lower bands at 1040 (C–H bending) and only lower at 830 cm^{-1} and at 770 cm^{-1} (pyranose ring from carbohydrates). The differences observed in soybean averaged spectrum in comparison with quinoa are mainly given by a different composition and differences in the interactions among them. Spectra for adulterated quinoa are similar with

slight differences from the ones observed in pure quinoa. Figure 1 b) shows averaged spectra for pure quinoa flour, adulterated quinoa flour with maize and pure maize flour (case QM1). In this case, averaged spectra obtained of pure maize and adulterated quinoa flours showed similar absorbance intensities except for the bands with maximum at 1745 (stretching of the ester carbonyl group), 1540 (amide II), 1145 (C–O and C–C stretching), 925 (a shoulder of C–H bending) and 860 (pyranose ring from carbohydrates) cm^{-1}. In QM1 case, pure quinoa flour shows lower absorbance intensities for all the frequencies. The lower subplot in Figure 1, labeled as c), displays the case QW1 with pure quinoa, pure wheat and quinoa adulterated with wheat flour. Pure wheat flour shows lower absorbance intensities at 2930, 2855 and 1745 cm^{-1} (due mainly to lipids), but higher intensity bands at 1640 (amide I), 1540 (amide II) and 925 (skeletal modes of pyranose) cm^{-1}. Adulterated quinoa with wheat shows higher intensities for the bands at 2930, 2855, 1745, 1640 and 1540 cm^{-1}.

To classify samples of the cases QS1, QM1 and QW1, PLS-DA and SIMCA models were performed. The confusion matrices obtained are shown in *Supplementary material* Table SM2 for QS1 case, *Supplementary material* Table SM3 for QM1 and *Supplementary material section* Table SM4 for QW2. The rows entries represent the real class samples, while the columns, the prediction made by the models to a class. Class 1 represents pure quinoa flour, Class 2 adulterated samples with 10, 5 and 1% w/w of soybean flour and Class 3 pure soybean flour samples. The elements of the diagonal in the matrices are the number of miss-classified samples, for example, for QS1 the calibration steps for PLS-DA and SIMCA show 1 and 2 miss-classified samples, but 59 not assigned samples for SIMCA. This represent, in our opinion, a problem to calculate the overall performance parameters for SIMCA, because not assigned samples were not consider and high values do not represent a good classification into three classes. The second block of Table SM2 displays the confusion matrices for the cross-validation (CV) step for PLS-DA and SIMCA. CV results show more realistic values than the calibration step, with higher number of miss-classified samples. Thus, CV results were preferred to discuss PLS-DA and SIMCA performances. For QS1

in the CV step, 7 miss-classifications take place using PLS-DA and 3 samples are incorrectly classified, with 90 not assigned samples for SIMCA. These results show a trend, also observed in the cases QM1 and QW1 (see *Supplementary material section* Tables SM2 and SM3), with better performances from PLS-DA over SIMCA in both, calibration and validation steps. An important observation derived from confusion matrices for QS1, QM1 and QW1 is that none of the pure adulterant flours were incorrectly classified in the calibration or CV steps, suggesting that pure soybean, maize and wheat flours are quite spectrally different from quinoa or adulterated quinoa samples.

The recovered classification performance parameters for QS1, QM1 and QW1 are shown in the three first rows of Table 2. The non-error rates (NER) for PLS-DA were almost perfect for calibration step and reached a value of 0.96, which means a low percentage of errors (4%) for CV step in QS1, QM1 and QW1 cases. To confirm the excellent performance reached by PLS-DA, Table 2 also shows accuracy values (ACC) and the area under the curve values (AUC) of receiver operating characteristic curves (ROC curves) for both calibration and CV steps. ACC values obtained for QS1, QM1 and QW1 were close to those obtained for NER with a slight difference of 2% lower for QW1. An important observation can be pointed out regarding ACC values, which is that the feasibility of ACC to estimate a model performance strictly depends on the relative classes sizes. If the classes have a significantly different number of samples, ACC tends to the sensitivity value (SEN) of the most representative class (Ballabio et al., 2018). AUC values were recovered into Table 2 from ROC curves shown for QS1, QM1 and QW1, in *Supplementary material section* Figures SM27, SM4 and SM7, respectively. Figure SM27 (*Supplementary material section*) displays the ROC curves for PLS-DA of QS1 case and reports the variation of sensitivity (SEN) and specificity (SPEC), changing the threshold value limit for each class (see Figure SM27 b), d) and f)). The area under the curve (AUC) was obtained after that step, by plotting SEN and 1-SPEC recovered pairwise values and AUC was calculated from Figure SM27 a), c) and e) for the three classes in QS1 case. AUC values are reported in Table 2 and were equal to 1 for the three cases

QS1, QM1 and QW1. However, it is important to establish that AUC values represent a classification parameter only for the calibration step of the model.

On the other side, SIMCA shows a worse performance than PLS-DA for these cases. NER values for SIMCA are not quite different for those obtained by PLS-DA in the validation step, except for QW1 with a NER value 11% lower compared to PLS-DA (see Table 2). The amount of not assigned to any class samples were too high in the three cases (see Tables SM1, SM2 and SM3 in *Supplementary material section*) and only had an impact on AUC values, obtaining poor values (0.74, 0.77 and 0.68) for the adulterated samples (class 2) of QS1, QM1 and QW1, respectively (Table 2). Thus, AUC values take relevance in this situation when the model's outcome shows a high number of not assigned samples.

Additionally, Figure 2 shows the class responses for the calibration step (subplots a), c) and e)) and the CV step (subplots b, d and f) for all the samples tested in QS1 case. This graph reveals miss-classified samples overpassing the threshold value limit (shown as a dashed line) for each class. Calibration step shows almost not misclassification for QS1, QM1 and QW1, except for one sample in the QM1 case (see subplots of Figure 2 a), c) and e) and *Supplementary material section* Figures SM3 and SM6 a), c) and e)). The visual representation in Figure 2 for CV step confirms that the results presented in the confusion matrix of Table SM2 for the three classes, the pure quinoa flours and the pure adulterant samples were efficiently classified and a few samples from the adulterated ones were predicted as pure quinoa samples, in all the cases belonging to the 1% w/w added adulterant. A similar trend is shown for QM1 and QW1 in the figures included into *Supplementary material section*.

3.2 Cases QS2, QM2 and QW2

In these cases, pure adulterant samples were removed from the QS1, QM1 and QW1 data sets and new PLS-DA and SIMCA models were run to detect a specific adulteration with soybean (QS2), maize (QM2) or wheat (QW2) flour on quinoa flour in three different proportions of 10, 5

and 1% w/w. The three rows of the middle of Table 2 show the recovered classification parameters for cases QS2, QM2 and QW2 with the aim of classifying the samples into two classes, pure quinoa and adulterated quinoa. This task was successfully performed by PLS-DA, where in the validation steps show NER values of 0.94, 0.97 and 0.95 for QS2, QM2 and QW2, respectively. ACC and AUC values confirm the results observed using NER parameter, indicating that PLS-DA was a good choice to detect a flour used as adulterant until 1% w/w proportion, with an error of 6% for the CV step in the worst scenario. SIMCA shows a lower performance in comparison with PLS-DA. Apparently, NER and ACC are not able to account the samples not assigned to any class. In contrast, AUC parameter must be considered to describe SIMCA performance. For these cases, the AUC outcomes for class 1 (pure quinoa flour) are 0.82, 0.96 and 0.99 for QS2, QM2 and QW2, respectively, and are lower in the case of the prediction of class 2 (adulterated quinoa samples) with AUC values of 0.75, 0.91 and 0.89 for cases QS2, QM2 and QW2, respectively. Additional supporting figures (calculated responses of PLS-DA for both the calibration and CV steps and ROC curves for PLS-DA and SIMCA) and tables (Confusions matrices) for cases QS2, QM2 and QW2 are available in the *Supplementary material section*.

3.3 Cases QSMW1, QSMW2 and QSMW3

The major challenge of the present work was to detect any of the three adulterants used (soybean, maize and wheat flours) independently of the type of flour used in 10%, 5% and 1% w/w proportions together (case QSMW1), 10 and 5% w/w proportions together (case QSMW2), and 10% w/w adulteration proportion (case QSMW3). The classification performance parameters using PLS-DA and SIMCA for the two steps (calibration and validation) are shown in the last three rows of Table 2. Considering the difficulty of the task presented to the model, PLS-DA shows a low NER value of 0.83 for QSMW1, an acceptable value of 0.92 for QSMW2 and an excellent value of 0.98 for QSMW3. These values indicate that adulterations with proportions of 10 and 5% w/w were feasible to achieved by PLS-DA independently of the adulterant used in this work, but for lower

proportions than 5% the specific datasets should be preferred. Results did not improve by Using SIMCA model. AUC values for the class 2 (adulterated samples) were 0.60, 0.67 and 0.87 for QSMW1, QSMW2 and QSMW3, respectively (see Table 2). To support all the values recovered in Table 3, additional figures and tables are available in *Supplementary material section*.

With the aim of confirming if overfitting was occurring during calibration and CV steps, an additional set of FT-MIR measurements were performed. These values were analyzed with the previously validated algorithms for the prediction of their classes. 108 new measurements were made: 27 new spectra of pure quinoa flour and 81 spectra of adulterated quinoa flour with soybean, maize and wheat in the three different proportions (10, 5 and 1% w/w). Table 3 shows the recovered performance parameters for case QSMW1 for the prediction step using the previously validated PLS-DA and SIMCA models. NER and ACC values in both cases were 0.93 and 0.94, respectively, using PLS-DA, and 1.00 for both NER and ACC using SIMCA. NER values of 0.93 and 0.94 represent 3 and 4 misclassified samples respectively out of the 108 presented in the prediction data set. The prediction step results obtained from the new measurements, showed in Table 3, are quite similar from those found in the calibration and validation steps (Table 2).

3.4 Untargeted approach for QSMW1 case

In recent years, untargeted approach was gaining interest among chemometricians and food scientists to set the boundaries of a target class (i.e. pure quinoa flour) and once the boundaries were established (fixing a significance level and two distance parameters for each sample), predict the acceptance or not of an unknown sample to the target class (Granato et al., 2018). Due to a good performance in classification problems, PLS-DA is one of the most used algorithms to classify samples as adulterated or not in food adulteration/authentication situations. However, several authors consider PLS-DA or other discriminant analysis-based algorithms not appropriate for food adulteration/authentication problems (Oliveri & Downey, 2012; Rodionova et al., 2016). According to the facts showed by these authors, untargeted analysis was additionally performed to case

QSMW1 using the information extracted for the class of pure quinoa flour which lead to the outcomes given in Table 2. Figure 3 a) shows the scatter plot of Q residual values vs. Hotelling T^2 values for each sample of the case QSMW1, with the boundaries fixed for the class of pure quinoa flour. Only 13 measurements of adulterated quinoa flour were detected as pure quinoa flour and 3 measurements of pure quinoa flour were out of the confidence boundaries. This represents a NER value of 0.96 which leads to a similar result compared to targeted analysis (see Table 2). In addition, the boundaries found were used to predict the class of new measurements data set (prediction set) showing only 2 pure quinoa measurements out of the confidence limits (Figure 3 b)). These results show that in the objective of predict if a sample is adulterated with the three adulterant flours, both approaches (targeted and untargeted) support each other.

4.0 Discussion

In the present work FT-MIR in combination with PLS-DA or SIMCA was successfully used to discriminate several cases of quinoa flour adulteration. In QS1, QM1 and QW1 cases the pure adulterant samples (soybean, maize and wheat) were clearly separated from quinoa or adulterated quinoa samples. FT-MIR showed to be a useful technique for the classification of other pure cereals (wheat and oats) and pseudo cereal (buckwheat), both raw and processed flours, using a wavelet transform method (WPTER) on FT-MIR spectra with a 100% of success (Cocchi et al., 2004). Moreover FT-MIR was used to determine dietary constituents in four types of wheat, three types of spelt, three types of rye and two types of triticale flours (Sujka et al., 2017) and moisture content in quinoa flour (González-Muñoz et al., 2016), showing a relation between the composition and the relative spectral intensities in the middle infrared (MIR) region (Ziegler et al., 2016).

The results described in sections 3.2 and 3.3 show the capability of PLS-DA and SIMCA models to classify or differentiate samples until 1% w/w of added adulterant flour when a specific adulterant was investigated, and until 5% w/w when these three possible adulterants were used. In

both approaches, FT-MIR with PLS-DA show acceptable ER values in the range of 2-8%. These results are in good agreement with those shown for the classification of edible oil adulterations until a 5% w/w proportion using FT-MIR (Gurdeniz & Ozen, 2009) or NIR imaging and spectroscopy for the detection of adulterants in wheat kernels, flours and bread formulations with an adulterant proportion in the range of 2.5-10% w/w (Verdú et al., 2016; Ziegler et al., 2016).

If targeted approach (including adulterated samples as another class) was performed, SIMCA showed poor global classification performances in comparison with PLS-DA in all cases presented in this work. This could be due to the model's calculation of the latent variables (components). SIMCA and PLS-DA are models based on computing a new space, defined by a previously selected number of components (also known as latent variables, which are a linear combination of original variables) as the variables of that new space. However, the rules to find these components are quite different. SIMCA is based on principal component analysis (PCA) and calculates the components in the direction of maximum variance of each modeled class. In the other hand, PLS-DA calculates the components by maximizing the component's correlation between *X* block (FT-MIR spectra for each sample) and *Y* response (the classes for each sample) for all classes (Berrueta et al., 2007). However, the results obtained using untargeted approach (considering the pure quinoa flour for set boundaries) by SIMCA were like those found with PLS-DA. This fact would encourage the use of untargeted approach in further research.

5.0 Conclusion

In the present work FT-MIR measurements on pure quinoa flour, adulterated quinoa flour and three adulterants flours (soybean, maize and wheat) in three different proportions (10, 5 and 1% w/w) in combination with PLS-DA or SIMCA models were explored. Two approaches with an acceptable error were provided. The first focused on the detection of a specific adulterant (cases QS2, QM2 and QW2) with a limit of 1% w/w of adulterant added. The second approach was to

detect unspecific adulterants (cases QSMW2 and QSMW3) with a limit of 5% w/w of adulterant proportion. This combination correctly classifies the three pure adulterants flours from pure quinoa flour. These results confirm that FT-MIR (in combination with PLS-DA and SIMCA) provides a rapid, non-destructive and accurately method to detect adulteration of quinoa flour with soy, maize and wheat flours in different proportions, for all the cases presented.

Acknowledgments

S.D.R. and M.P.B. are research staff members of the National Scientific and Technical Research Council (*CONICET*). G.R. is recipient of a *CONICET* fellowship. The authors are grateful for the financial support given by PICT 2013-1331 and UBACYT 2014-2017 20020130100443BA funds.

Conflict of interest statement

Funding: This study was funded by *Agencia Nacional de Promoción Científica y Tecnológica* (PICT 2013-1331) and *Programación UBACYT* (20020130100443BA).

Conflict of interest: Silvio David Rodríguez, Guido Rolandelli and María del Pilar Buera declare to have no conflict of interest.

Ethical approval: This article does not contain any study with human participants or animals performed by any of the authors.

References

Aluwi, N. A., Murphy, K. M., &Ganjyal, G. M. (2017). Physicochemical Characterization of Different Varieties of Quinoa. *Cereal Chemistry Journal*, *94*(5), 847–856. https://doi.org/10.1094/CCHEM-10-16-0251-R

AOAC. (2016). Official Methods of Analysis of AOAC INTERNATIONAL 20th Ed., AOAC INTERNATIONAL, Gaithersburg, MD, USA, www.eoma.aoac.org

Asociación Latinoamericana de Integración (ALADI), Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO). (2014). Tendencias y Perspectivas del Comercio Internacional de Quinua. AccessedFebruary 2018, available at http://www.fao.org/3/a-i3583.pdf

Ballabio, D., & Consonni, V. (2013). Classification tools in chemistry. Part 1: linear models. PLS-DA. *Analytical Methods*, *5*(16), 3790. https://doi.org/10.1039/c3ay40582f

Ballabio, D., Grisoni, F., & Todeschini, R. (2018). Multivariate comparison of classification performance measures. *Chemometrics and Intelligent Laboratory Systems*, *174*, 33–44. https://doi.org/10.1016/j.chemolab.2017.12.004

Berrueta, L. A., Alonso-Salces, R. M., & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A*, *1158*(1–2), 196–214. https://doi.org/10.1016/j.chroma.2007.05.024

Botelho, B. G., Reis, N., Oliveira, L. S., & Sena, M. M. (2015). Development and analytical validation of a screening method for simultaneous detection of five adulterants in raw milk using mid-infrared spectroscopy and PLS-DA. *Food Chemistry*, *181*, 31–37. https://doi.org/10.1016/j.foodchem.2015.02.077

Cocchi, M., Foca, G., Lucisano, M., Marchetti, A., Pagani, M. A., Tassi, L., & Ulrici, A. (2004). Classification of Cereal Flours by Chemometric Analysis of MIR Spectra. *Journal of Agricultural and Food Chemistry*, *52*(5), 1062–1067. https://doi.org/10.1021/jf034441o

Collins, E. J. T. (1993). Food adulteration and food safety in Britain in the 19th and early 20th centuries. *Food Policy*, *18*(2), 95–109. https://doi.org/10.1016/0306-9192(93)90018-7

Ellis, D. I., Brewster, V. L., Dunn, W. B., Allwood, J. W., Golovanov, A. P., & Goodacre, R. (2012). Fingerprinting food: current technologies for the detection of food adulteration and contamination. *Chemical Society Reviews*, *41*(17), 5706. https://doi.org/10.1039/c2cs35138b

Ferreira, D. S., Pallone, J. A. L., & Poppi, R. J. (2015). Direct analysis of the main chemical constituents in Chenopodium quinoa grain using Fourier transform near-infrared spectroscopy. *Food Control*, *48*, 91–95. https://doi.org/10.1016/j.foodcont.2014.04.016

Filho, A. M. M., Pirozi, M. R., Borges, J. T. D. S., Pinheiro Sant'Ana, H. M., Chaves, J. B. P., & Coimbra, J. S. D. R. (2017). Quinoa: Nutritional, functional, and antinutritional aspects. *Critical Reviews in Food Science and Nutrition*, *57*(8), 1618–1630. https://doi.org/10.1080/10408398.2014.1001811

González-Muñoz, A., Montero, B., Enrione, J., & Matiacevich, S. (2016). Rapid prediction of moisture content of quinoa (Chenopodium quinoa Willd.) flour by Fourier transform infrared (FTIR) spectroscopy. *Journal of Cereal Science*, *71*, 246–249. https://doi.org/10.1016/j.jcs.2016.09.006

Granato, D., Putnik, P., Kovacevic, D.B., Sousa Santos, J., Calado, V., Silva Rocha, R., Gomes Da Cruz, A., Jarvis, B., Rodionova, O.Y., &Pomerantsev, A. (2018). Trends in Chemometrics: Food Authentication, Microbiology, and Effects of Processing. Comprehensive Reviews in Food Science and Food Safety, *17(3)*, 663–677. https://doi.org/10.1111/1541-4337.12341

Gurdeniz, G., &Ozen, B. (2009). Detection of adulteration of extra-virgin olive oil by chemometric analysis of mid-infrared spectral data. *Food Chemistry*, *116*(2), 519–525. https://doi.org/10.1016/j.foodchem.2009.02.068

Guzmán-Ortiz, F. A., Hernández-Sánchez, H., Yee-Madeira, H., San Martín-Martínez, E., Robles-Ramírez, M. del C., Rojas-López, M., … Mora-Escobedo, R. (2015). Physico-chemical, nutritional and infrared spectroscopy evaluation of an optimized soybean/corn flour

extrudate. *Journal of Food Science and Technology*, *52*(7), 4066–4077. https://doi.org/10.1007/s13197-014-1485-5

Karoui, R., Downey, G., & Blecker, C. (2010). Mid-Infrared Spectroscopy Coupled with Chemometrics: A Tool for the Analysis of Intact Food Systems and the Exploration of Their Molecular Structure−Quality Relationships − A Review. *Chemical Reviews*, *110*(10), 6144–6168. https://doi.org/10.1021/cr100090k

Knödler, M., Most, M., Schieber, A., & Carle, R. (2010). A novel approach to authenticity control of whole grain durum wheat (Triticum durum Desf.) flour and pasta, based on analysis of alkylresorcinol composition. *Food Chemistry*, *118*(1), 177–181. https://doi.org/10.1016/j.foodchem.2009.04.080

Laparra, J. M., & Haros, M. (2018). Inclusion of Whole Flour from Latin-American Crops into Bread Formulations as Substitute of Wheat Delays Glucose Release and Uptake. *Plant Foods for Human Nutrition*. https://doi.org/10.1007/s11130-018-0653-6

Lohumi, S., Lee, S., Lee, H., & Cho, B.-K. (2015). A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. *Trends in Food Science & Technology*, *46*(1), 85–98. https://doi.org/10.1016/j.tifs.2015.08.003

López, M. I., Colomer, N., Ruisánchez, I., & Callao, M. P. (2014). Validation of multivariate screening methodology. Case study: Detection of food fraud. *Analytica Chimica Acta*, *827*, 28–33. https://doi.org/10.1016/j.aca.2014.04.019

López, M. I., Trullols, E., Callao, M. P., & Ruisánchez, I. (2014). Multivariate screening in food adulteration: Untargeted versus targeted modelling. *Food Chemistry*, *147*, 177–181. https://doi.org/10.1016/j.foodchem.2013.09.139

Luna, A. S., da Silva, A. P., Alves, E. A., Rocha, R. B., Lima, I. C. A., & de Gois, J. S. (2017). Evaluation of chemometric methodologies for the classification of Coffea canephora cultivars via FT-NIR spectroscopy and direct sample analysis. *Analytical Methods*, *9*(29), 4255–4260. https://doi.org/10.1039/C7AY01167A

Luna, A. S., Pinho, J. S. A., & Machado, L. C. (2016). Discrimination of adulterants in UHT milk samples by NIRS coupled with supervision discrimination techniques. *Analytical Methods*, *8*(39), 7204–7208. https://doi.org/10.1039/C6AY01351A

Manning, L., & Soon, J. M. (2014). Developing systems to control food adulteration. *Food Policy*, *49*, 23–32. https://doi.org/10.1016/j.foodpol.2014.06.005

Nowak, V., Du, J., & Charrondière, U. R. (2016). Assessment of the nutritional composition of quinoa (Chenopodium quinoa Willd.). *Food Chemistry*, *193*, 47–54. https://doi.org/10.1016/j.foodchem.2015.02.111

Ojinnaka, D. (2016). Legislative control of quinoa in the United Kingdom and European Union. *Madridge Journal of Food Technology, 1*(1), 53–57. https://doi.org/10.18689/mjft.2016-108

Oliveri, P., & Downey, G. (2012). Multivariate class modeling for the verification of food-authenticiy claims. Trends in Analytical Chemistry, *35*, 74–86. https://doi.org/10.1016/j.trac.2012.02.005

Ozen, B. F., & Mauer, L. J. (2002). Detection of Hazelnut Oil Adulteration Using FT-IR Spectroscopy. *Journal of Agricultural and Food Chemistry*, *50*(14), 3898–3901. https://doi.org/10.1021/jf0201834

Pallone, J. A. L., Caramês, E. T. S., & Alamar, P. D. (2018). Green analytical chemistry applied in food analysis: alternative techniques. *Current Opinion in Food Science*. https://doi.org/10.1016/j.cofs.2018.01.009

Pomerantsev, A. L., & Rodionova, O. Y. (2012). Process analytical technology: a critical view of the chemometricians: PAT: a critical view of the chemometricians. *Journal of Chemometrics*, *26*(6), 299–310. https://doi.org/10.1002/cem.2445

Roa, D. F., Santagapita, P. R., Buera, M. P., & Tolaba, M. P. (2014). Ball Milling of Amaranth Starch-Enriched Fraction. Changes on Particle Size, Starch Crystallinity, and Functionality as a Function of Milling Energy. *Food and Bioprocess Technology*, *7*(9), 2723–2731. https://doi.org/10.1007/s11947-014-1283-0

Rodionova, O.Y., Titova, A.V., &Pomerantsev, A.L. (2016). Discriminant analysis is an inappropriate method of authentication. Trends in Analytical Chemistry, *78*, 17–22. https://doi.org/10.1016/j.trac.2016.01.010

Rodriguez-Saona, L. E., &Allendorf, M. E. (2011). Use of FTIR for Rapid Authentication and Detection of Adulteration of Food. *Annual Review of Food Science and Technology*, *2*(1), 467–483. https://doi.org/10.1146/annurev-food-022510-133750

Ropodi, A. I., Panagou, E. Z., & Nychas, G.-J. E. (2016). Data mining derived from food analyses using non-invasive/non-destructive analytical techniques; determination of food authenticity, quality & safety in tandem with computer science disciplines. *Trends in Food Science & Technology*, *50*, 11–25. https://doi.org/10.1016/j.tifs.2016.01.011

Ruiz, K. B., Biondi, S., Oses, R., Acuña-Rodríguez, I. S., Antognoni, F., Martinez-Mosqueira, E. A., … Molina-Montenegro, M. A. (2014). Quinoa biodiversity and sustainability for food security under climate change. A review. *Agronomy for Sustainable Development*, *34*(2), 349–359. https://doi.org/10.1007/s13593-013-0195-0

Stevens, A. W. (2017). Quinoa quandary: Cultural tastes and nutrition in Peru. *Food Policy*, *71*, 132–142. https://doi.org/10.1016/j.foodpol.2017.08.003

Su, W.-H., & Sun, D.-W. (2017). Evaluation of spectral imaging for inspection of adulterants in terms of common wheat flour, cassava flour and corn flour in organic Avatar wheat (Triticum spp.) flour. *Journal of Food Engineering*, *200*, 59–69. https://doi.org/10.1016/j.jfoodeng.2016.12.014

Sujka, K., Koczoń, P., Ceglińska, A., Reder, M., & Ciemniewska-Żytkiewicz, H. (2017). The Application of FT-IR Spectroscopy for Quality Control of Flours Obtained from Polish Producers. *Journal of Analytical Methods in Chemistry*, *2017*, 1–9. https://doi.org/10.1155/2017/4315678

van den Berg, F., Lyndgaard, C. B., Sørensen, K. M., & Engelsen, S. B. (2013). Process Analytical Technology in the food industry. *Trends in Food Science & Technology*, *31*(1), 27–35. https://doi.org/10.1016/j.tifs.2012.04.007

Verdú, S., Vásquez, F., Grau, R., Ivorra, E., Sánchez, A. J., & Barat, J. M. (2016). Detection of adulterations with different grains in wheat products based on the hyperspectral image technique: The specific cases of flour and bread. *Food Control*, *62*, 373–380. https://doi.org/10.1016/j.foodcont.2015.11.002

Warren, F. J., Gidley, M. J., & Flanagan, B. M. (2016). Infrared spectroscopy as a tool to characterise starch ordered structure—a joint FTIR–ATR, NMR, XRD and DSC study. *Carbohydrate Polymers*, *139*, 35–42. https://doi.org/10.1016/j.carbpol.2015.11.066

Ziegler, J. U., Leitenberger, M., Longin, C. F. H., Würschum, T., Carle, R., & Schweiggert, R. M. (2016). Near-infrared reflectance spectroscopy for the rapid discrimination of kernels and flours of different wheat species. *Journal of Food Composition and Analysis*, *51*, 30–36. https://doi.org/10.1016/j.jfca.2016.06.005

Figure Captions

Figure 1. Averaged FT-MIR spectra in the range of 3050 to 2750 cm^{-1} and 600-1800 cm^{-1}, for pure quinoa flour (long dashed blue line), adulterated quinoa (short dashed red line) with: a) soybean flour, b) maize flour, and c) wheat flour, and pure adulterant flour (solid green line): a) soybean, b) maize and c) wheat.

Figure 2. Calculated class responses for PLS-DA model used in the case QS1for the calibration step: a), c) and e) and validation step: b), d) and f). Blue circles represent samples of pure quinoa flour, red squares represent samples of adulterated quinoa flour and green diamonds represent pure soybean flour for class 1: a) and b); class 2: c) and d) and class 3: e) and f).

Figure 3. Scatter plot of Q residuals values vs. Hotelling T^2 values for the measurements in QSMW1 case using the class of pure quinoa flour to set the boundaries (confidence of 95%, solid black lines): a) calibration set and b) prediction set. Blue circles represent measurements of pure quinoa flour, red squares represent measurements of adulterate quinoa flour.

Table 1. Average values with standard deviation of the major dietary constituents of quinoa, soybean, maize and wheat flours. The values are expressed in % dry basis, except for moisture content, which are expressed in % of wet basis.

Flour type		Major dietary components (%)				
		Moisture	**Protein**	**Lipids**	**Ash**	**Carbohydrates**
Quinoa	**Brand 1**	14.3 ± 0.8	19.0 ± 0.7	7.5 ± 0.5	3.2± 0.3	70.3± 1.1
	Brand 2	13.4 ± 0.8	16.3 ± 0.8	6.9 ± 0.6	4.1 ± 0.4	72.6 ± 1.1
	Brand 3	16.2±1.0	20.7 ± 0.9	10.3 ± 0.5	3.7 ± 0.4	65.3± 0.9
	Brand 4	13.3 ± 0.7	13.8 ± 0.8	7.7 ± 0.5	2.6 ± 0.5	75.9± 1.00
	Brand 5	11.8 ± 1.1	13.8 ± 0.7	7.1± 0.5	4.2±0.4	74.9 ± 0.9
	Brand 6	12.9 ± 0.7	18.2 ± 0.5	6.9± 0.4	2.3 ± 0.3	72.6± 0.9
	Brand 7	15.8± 0.8	17.1 ± 0.7	8.7± 0.5	5.6± 0.3	68.7± 1.1
	Brand 8	16.7 ± 0.9	16.0 ±1.0	12.0± 0.5	5.0 ± 0.5	67.0 ± 1.2
	Brand 9	14.7 ± 0.8	18.4± 0.8	8.3 ± 0.6	4.3± 0.5	69.0 ± 0.9
Soybean		10.5± 0.7	41.9 ± 0.6	23.7 ± 0.6	2.9± 0.5	31.4 ± 0.8
Maize		13.3 ± 0.5	9.5± 0.8	4.7 ± 0.4	4.1 ± 0.5	81.7 ± 1.1
Wheat		12.2 ±1.0	14.7 ± 0.8	5.2 ± 0.4	1.3 ± 0.4	78.8±1.0

Table 2. Recovered classification performance parameters for PLS-DA and SIMCA models used in the adulteration cases.

Dataset	Number of classes	PLS-DA						SIMCA					
		Number of components	Calibration step			CV step		Number of components	Calibration step			CV step	
			NER	ACC	AUC	NER	ACC		NER	ACC	AUC	NER	ACC
QS1	3	19	0.99	0.99	C1: 1.00 C2: 1.00 C3: 1.00	0.96	0.96	C1: 9 C2: 1 C3: 1	0.99	0.98	C1: 0.87 C2: 0.74 C3: 1.00	0.96	0.96
QM1	3	20	1.00	1.00	C1: 1.00 C2: 1.00 C3: 1.00	0.96	0.96	C1: 8 C2: 9 C3: 2	1.00	1.00	C1: 0.97 C2: 0.77 C3: 1.00	0.94	0.95
QW1	3	20	1.00	1.00	C1: 1.00 C2: 1.00 C3: 1.00	0.96	0.94	C1: 8 C2: 6 C3: 2	0.97	0.96	C1: 0.91 C2: 0.68 C3: 1.00	0.85	0.92
QS2	2	13	0.96	0.97	C1: 1.00 C2: 1.00	0.94	0.95	C1: 3 C2: 1	0.84	0.85	C1: 0.82 C2: 0.75	0.86	0.87
QM2	2	19	1.00	1.00	C1: 1.00 C2: 1.00	0.97	0.98	C1: 8 C2: 19	1.00	1.00	C1: 0.96 C2: 0.91	0.97	0.99
QW2	2	19	1.00	1.00	C1: 1.00 C2: 1.00	0.95	0.96	C1: 11 C2: 17	1.00	1.00	C1: 0.99 C2: 0.89	0.99	0.98
QSMW1	2	20	0.96	0.99	C1: 1.00 C2: 1.00	0.83	0.95	C1: 11 C2: 1	0.99	0.98	C1: 0.99 C2: 0.60	0.78	0.95
QSMW2	2	20	0.99	1.00	C1: 1.00 C2: 1.00	0.92	0.96	C1: 11 C2: 1	0.99	0.98	C1: 0.99 C2: 0.67	0.78	0.93
QSMW3	2	20	1.00	1.00	C1: 1.00 C2: 1.00	0.98	0.99	C1: 8 C2: 17	1.00	1.00	C1: 0.95 C2: 0.87	0.97	0.99

* NER: non-error rate, ACC: accuracy, AUC: area under the curve (ROC analysis).
* CV step: Cross validation step.
* C1: class 1, C2: class 2, C3: class 3.

Table 3. Recovered classification performance parameters for the prediction of new samples using PLS-DA and SIMCA models in QSMW1 case.

Data set	Prediction step PLS-DA		Prediction step SIMCA	
	NER	ACC	NER	ACC
QSMW1	0.93	0.94	1.00	1.00

Figure 1. Rodríguez et al.

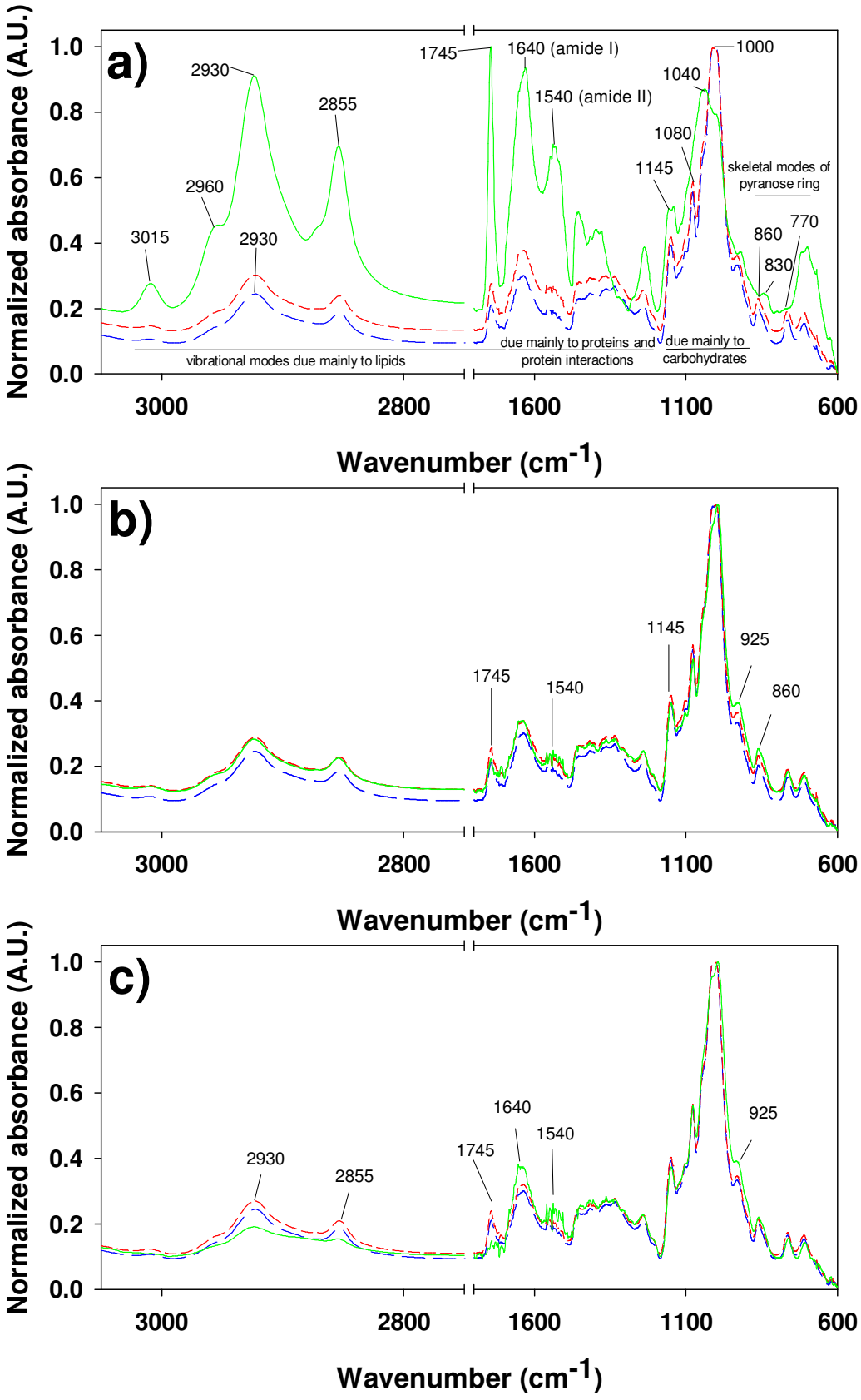


Figure 2. Rodríguez et al.

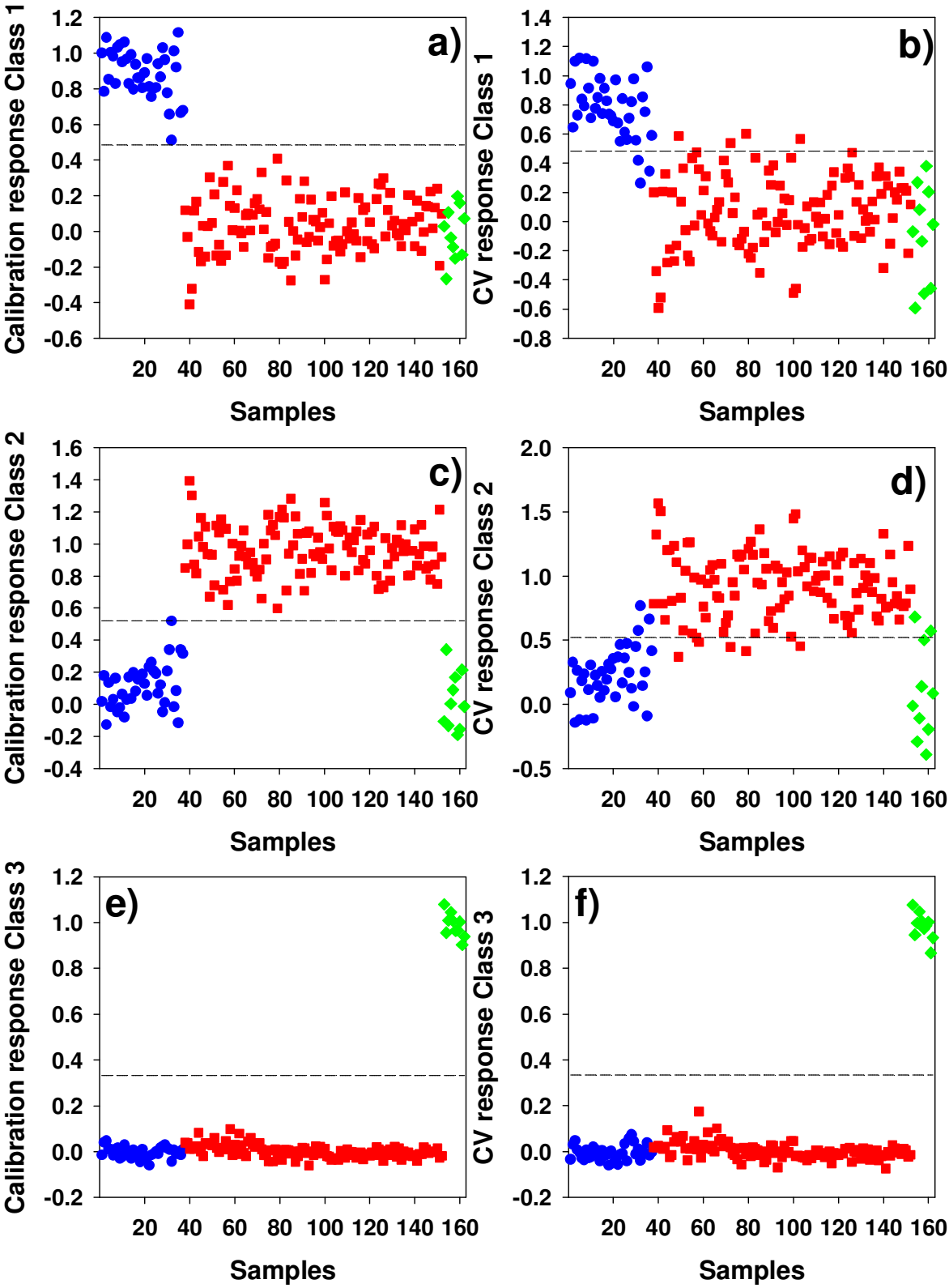
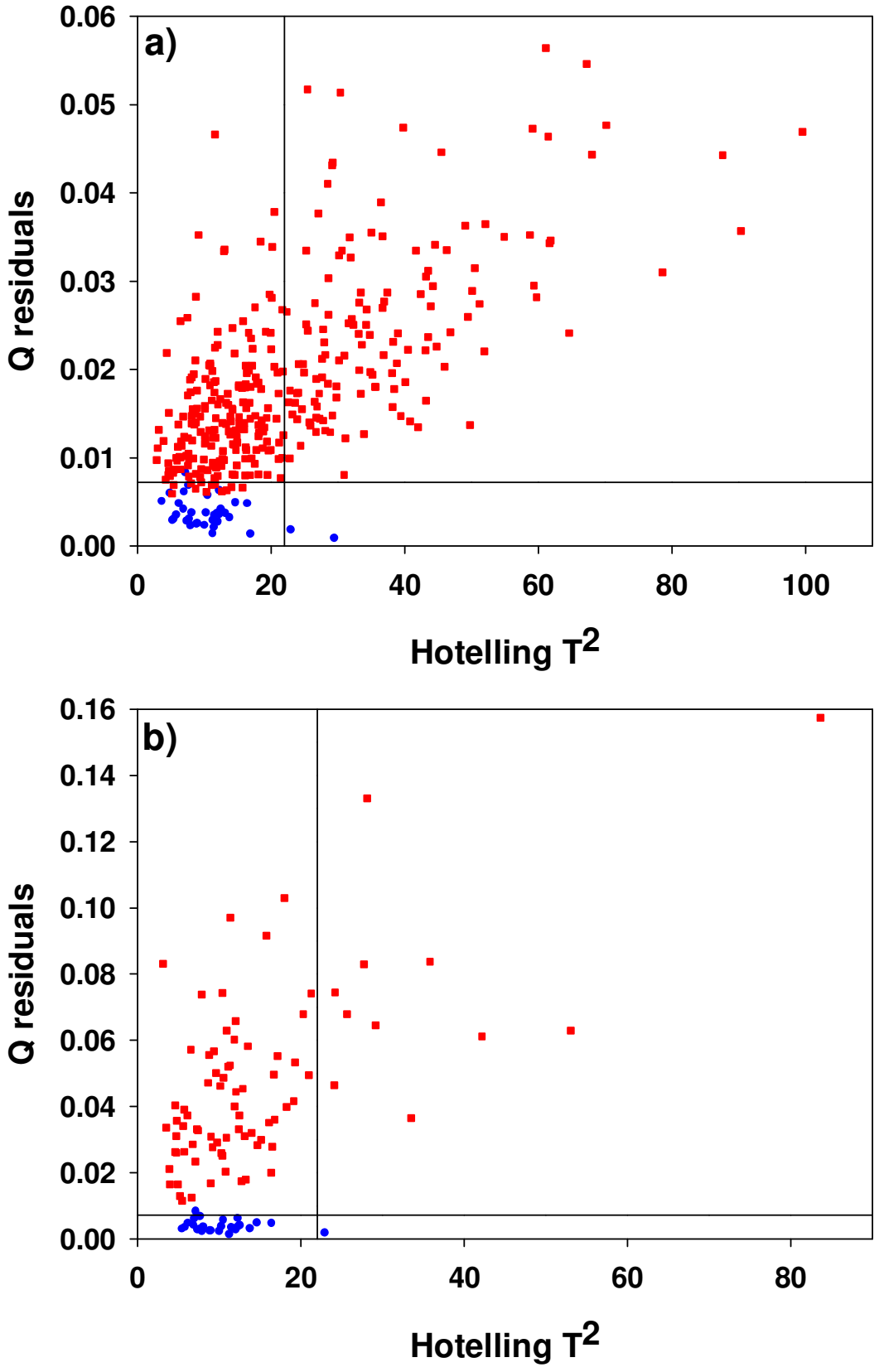


Figure 3. Rodríguez et al.



Highlights Rodríguez et al 2018

- FT-MIR was used to detect soybean, maize and wheat flours in pure quinoa flour.
- PLS-DA and SIMCA were useful tools for the detection of quinoa flour adulteration.
- Three different strategies for quinoa adulteration detection were discussed.