Journal of the Science of Food and Agriculture

Phytochemicals determination and classification in purple and red fleshed potato tubers by analytical methods and near-infrared spectroscopy

Journal:	*Journal of the Science of Food and Agriculture*
Manuscript ID:	JSFA-15-0181.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	TIERNO, ROBERTO; NEIKER, PLANT PRODUCTION LOPEZ, AINARA; UPNA, AGRICULTURAL PROJECTS Patrick, Riga; Basque Institute for Agricultural Research and Development (NEIKER), Department of Plant Production and Protection Arazuri, Silvia; Universidad Pública de Navarra, Proyectos e Ingeniería Rural:Mecatrónica Agraria Jaren, Carmen; Universidad Pública de Navarra, Deprtament of Proyect and Rural Engineering BENEDICTO, LEIRE; NEIKER, PLANT PRODUCTION RUIZ DE GALARRETA GOMEZ, JOSE IGNACIO; NEIKER, PLANT PRODUCTION
Key Words:	soluble phenolics, monomeric anthocyanins, carotenoids,, vitamin C, NIRS, Solanum tuberosum L.

Phytochemicals determination and classification in purple and red fleshed potato tubers by analytical methods and near-infrared spectroscopy

R. Tierno[1], A. López[2], P. Riga[1], S. Arazuri[2], C. Jarén[2], L. Benedicto[1], J. I. Ruiz de Galarreta[1]

[1]NEIKER-Tecnalia. The Basque Institute for Agricultural Research and Development. PO. Box. 46 E-01080. Vitoria. Spain.

[2]Department of Agricultural Projects and Engineering, Universidad Pública de Navarra, Campus de Arrosadía 31006. Navarra. Spain.

Corresponding author.

Jose Ignacio Ruiz de Galarreta

Tel: +34-945-121313

Fax: + 34-945-281422

E-mail: jiruiz@neiker.net

Abstract

BACKGROUND

Over the last two decades, the attractive colours and shapes of pigmented tubers and the increasing concern about the relationship between nutrition and health are contributing to the expansion of their consumption and specialty market. Thus, we have quantified the concentration of health promoting compounds such as soluble phenolics, monomeric anthocyanins, carotenoids, vitamin C, and hydrophilic antioxidant capacity, in a collection of 18 purple and red fleshed potato accessions.

RESULTS

Cultivars and breeding lines high in vitamin C, such as Blue Congo, Morada and Kasta, have been identified. Deep purple cultivars Violet Queen, Purple Peruvian and Vitelotte showed high levels of soluble phenolics, monomeric anthocyanins, and hydrophilic antioxidant capacity, whereas relatively high carotenoid concentrations were found in partially yellow coloured tubers, such as Morada, Highland Burgundy Red, and Violet Queen.

CONCLUSION

The present characterization of cultivars and breeding lines with high concentrations of phytochemicals is an important step both to support the consideration of specialty potatoes as a source of healthy compounds, and to obtain new cultivars with positive nutritional characteristics. Moreover, by using near-infrared spectroscopy a non-destructively identification and classification of samples with different levels of phytochemicals is achieved, offering an unquestionable contribution to potato industry for future automatic discrimination of varieties.

Keywords: soluble phenolics, monomeric anthocyanins, carotenoids, vitamin C, NIRS, *Solanum tuberosum* L.

Highlights:

Total soluble phenolics and monomeric anthocyanins found in certain purple potato cultivars were comparable to those of blueberries

Soluble phenolics, monomeric anthocyanins and hydrophilic antioxidant capacity of potato tubers were correlated.

Specialty potatoes can improve human diet

A classification of varieties according to their phytochemicals content is achieved by using non-destructive near-infrared reflectance spectroscopy coupled with chemometrics.

INTRODUCTION

Potatoes are an excellent source of vitamins and phytochemicals, such as vitamin C, phenolic acids, flavonoids and carotenoids, which may provide a wide range of health benefits [1]. Despite the fact that there are important differences between genotypes, potatoes may contribute to about 44 % of the recommended daily intake of vitamin C [2] depending on the diet. Among other things, this potent reducing agent acts as an antioxidant in the body, cofactor for numerous enzymes and also plays an important role in increasing the bioavailability of iron [3].

Phenolic compounds are considered to be health promoting phytochemicals as they have shown beneficial properties. Phenolics are commonly classified into three important groups: phenolic acids, flavonoids and tannins [4]. The major phenolic acid in potato is chlorogenic acid [5]. Anthocyanins are phenolic pigments which constitute the main subclass among flavonoids. The most common anthocyanidins (the de-glycosylated forms of anthocyanins) found in potatoes are malvidin, petunidin, delphinidin and peonidin in purple tubers and pelargonidin in red ones [6]. Different aglycones and sugar moieties determine their bioavailability and potential health effects [7]. A study of 74 Andean potato landraces revealed a wide variability for total phenolic compounds from 1.12 to 12.4 g GAE kg^{-1} DW and antioxidant capacity from 0.0283 to 0.251 mol TE kg^{-1} DW [8] thus showing the important genotype effect on the content of health promoting phytochemicals.

Carotenoids are a widespread family of lipophilic organic pigments [9]. Based on epidemiological studies a positive link is suggested between higher dietary intake and tissue concentrations of carotenoids and lower risk of chronic diseases [10]. The orange flesh colour found in some potatoes is due to zeaxanthin whereas lutein concentration correlates with

the intensity of yellow coloration. Cultivated diploid potatoes derived from *Solanum stenotonum* Juz. & Bukasov and *Solanum phureja* Juz. & Bukasov have been reported to be a great source of zeaxanthin and lutein [11-13].

Purple and red fleshed potato cultivars are attractive to consumers. Besides their exotic pigmentation, coloured genotypes show significantly higher contents of phenolic compounds [14-15]. Specialty potato food products, such as coloured chips, crisps, purees, canned potatoes, and ready meals are becoming more and more widespread. With varying degrees of success, some companies and research centres are trying to develop new coloured cultivars to support increasing demand of specialty potato market, giving particular emphasis on their superior nutritional profile, attractive colours and different textures. To this end, the selection of appropriate parents to be used in artificial crosses is one of the main decisions faced by breeders. Thus the identification of promising genotypes is a key step in potato breeding.

On the other hand, the analytical methods commonly employed to determine main compounds of potatoes in order to identify valuable genotypes require a lot of time and are destructive. Therefore, these methods seem to be not suitable for in-line applications in the food industry [16]. Nowadays, there are non-destructive available technologies to perform qualitative and quantitative analysis of food and food products. In this respect, Near-infrared spectroscopy (NIRS) is considered one of the most advanced techniques concerning non-destructive quality control of agricultural and food products [17]. It has been successfully used for the quantitative analysis of many agricultural and food products during the last decades [18-19]. However, NIR can also be used for qualitative analysis, where the aim is to classify samples on the basis of its spectral features rather than estimate the components present in them [20].

NIR applications for the prediction of potato constituents are a common practice nowadays; however, in this study taking into account that the number of samples is rather small and they are very specific, it becomes very challenging to develop a robust model for the prediction of phytochemicals. Other authors were able to determine the total phenolic content of whole and lyophilized tubers while working with a large number of samples including yellow, red and purple varieties [21-22].

Despite that quantitative analyses of potatoes are widely extended, literature concerning qualitative analysis is not as much to the best of our knowledge. Some authors have focused in the classification of potato tubers according to their chemical components. Thus, in a research developed by Fernández-Ahumada et al. [23], a discriminant analysis was performed in order to classify samples in two categories regarding to their protein content. The results obtained demonstrated the accuracy of NIR to classify potato samples in groups of low and high protein content.

In this study, we have analysed the concentration of phenolic compounds, monomeric anthocyanins, carotenoids, vitamin C, and hydrophilic antioxidant capacity by analytical methods in a collection of 18 purple and red fleshed potato cultivars to classify samples in 3 groups (low, mid and high) according to their phytochemicals content by NIRS combined with chemometric tools.

MATERIALS AND METHODS

Plant material and experimental conditions

A collection of 18 purple and red fleshed potato cultivars and breeding lines were selected on the basis of the contrasting flesh colour of tubers (Table 1). Tubers were selected from potato accessions (Potato Germplasm Collection, NEIKER) grown during the year 2013 in a precise field trial in Arkaute (Alava) in the northeastern of Spain (550 masl) with humid climate and annual rainfall of about 800 L m^{-2}. The soil with a clay loam texture was previously subjected to conventional wheat cropping. Plants were grown from mid May to mid October 2013 after pre-sowing fertilization with 800 kg ha^{-1} (NPK 4-8-16). Watering was performed using an automatic spray irrigation system. After harvesting, potatoes were stored at 4 °C in a darkened cold room for one month.

NIR spectral acquisition

Prior to lyophilisation NIRS measurements were made in whole intact tubers (unpeeled and free of soil). No previous preparation of the samples was accomplished for spectral acquisition. Since the main objective in the near future is to implement these techniques for real time measurements in potato handling lines, it is considered essential to analysed the tubers as they are currently being manipulated in those lines. NIR spectral data were collected using a Luminar 5030 Miniature "Hand held" AOTF-NIR Analyzer (Brimrose, Baltimore, Maryland). It covers a spectral range between 1100 and 2300 nm with 601 points (2 nm steps) and was used in reflectance mode. Samples were scanned at four different points along the equatorial area and the average spectrum was used for the analysis. Each spectrum was an average of 50 scans.

By measuring the samples in reflectance mode we obtained information from not only the skin but also from the flesh. Lammertyn et al. [24] proved that when a sample is irradiated in the near-infrared spectral range, the light penetrates through the skin and the NIR reflectance spectrum of samples contains information of the background. Moreover, they concluded that the amount of information coming from the background exceeded that coming from the skin. Those authors established a NIR penetration depth between 2 and 3 mm in the 900–1900 nm range for apple. In another study developed by López et al. [25] authors also found that NIR reflectance spectra of unpeeled potatoes in the 1100-2300 nm range contained information of both skin and flesh.

Sample preparation for chemical analysis

Five raw tubers of each accession were washed and patted dry with paper towels, and subsequently peeled and diced. Diced tubers were divided into two equal parts. One part was frozen with liquid nitrogen, kept frozen at -80°C, and later freeze dried, milled by an automatic mortar grinder, and stored at -30°C until analysis. Lyophilisation was used for the determination of dry matter content. Total soluble phenolics and total monomeric anthocyanins were analysed using freeze dried tubers. Total carotenoids, vitamin C, and hydrophilic antioxidant capacity were determined immediately using the fresh portion. All assays were performed in triplicate.

Extraction and quantitative determination of vitamin C (VC)

Tubers were homogenized in oxalic acid 4% (1:1 w/v), until a homogeneous puree is obtained (1 min) in a Palson blender at maximum speed (1200 W). To optimize the homogenization, 20 g of puree were weighed in a 50 mL tube and a teaspoon of steel balls

was added to each tube and re-homogenized in Bullet Blender (Next Advance) 15 minutes at power 9. Then, samples were centrifuged at 3350 *g*, 10 min. The supernatant was preserved at -80 °C until analysis. Under these conditions, vitamin C was stable for at least two months (data not shown). On the day of analysis, the samples were thawed submerging in cold tap water for 15 minutes. Samples were shacked in a vortex, centrifuged 5 min at 372 *g* and filtered with glass fibre filter of 1 µm. The filtered extract was oxidized with 0.5% aqueous diclorophenol-indophenol solution and subsequently mixed with 2% 2, 4-dinitrophenylhydrazine in 70% sulphuric acid to allow the formation of hydrazones. The hydrazones were extracted with ethyl acetate: acetic acid 98:2 and this orange coloured extract was applied directly onto the HPTLC plates (nano silica gel F 254, 20 x 10 cm, Fluka) by the mean of the semi-automatic sampler LINOMAT 5 (CAMAG). Chromatography was developed in horizontal chamber as described in the application notes A. 10.5 [26]. The total ascorbic acid was measured at 510 nm by CAMAG TLC Scanner 3 and was quantified with the Wincats software. Results were expressed as g of vitamin C per kg fresh weight.

Extraction and quantitative determination of soluble phenolics (TSP)

Phenolics were extracted from lyophilized powder (1 g) with 10 mL MeOH: H_2O (70:30, v/v). The solid was suspended by shaking in a vortex for 1 min. The mixture was centrifuged at 7730 *g* for 10 min at 4 °C, and the supernatant containing the extracted phenolics was collected. The extraction operation was repeated twice again with the pellet and the final volume carried to 30 mL. Phenolic compounds were quantified following the

method described by Medina [27]. Results were expressed as g Gallic acid equivalents per kg fresh weight.

Extraction and quantitative determination of monomeric anthocyanins (TMA)

Monomeric anthocyanins were extracted and quantified by the pH differential method [28]. Anthocyanins were extracted from lyophilized powder (0.25 g) with 10 mL MeOH: HCl (99:1, v/v). The solid was suspended by shaking in a vortex for 1 min. The mixture was centrifuged at 7730 *g* for 10 min at 4 °C, and the supernatant containing the extracted anthocyanins was collected. The extraction operation was repeated twice again with the pellet and the final volume carried to 30 mL. After the determination of the appropriate dilution factor, two solutions of each test sample were prepared, one with pH = 1.0 aqueous buffer (Potassium chloride, 0.025 M) and other with 4.5 pH aqueous buffer (Sodium acetate, 0.4 M). Total monomeric anthocyanin concentration was calculated measuring the optical density at 520 and 700 nm at two different pH values (pH = 1.0 and 4.5). Absorbance was measured within 20–50 min of preparation vs MeOH: HCl (99:1, v/v) at pH 1 and 4.5 using the same dilution factor. The calculation of monomeric anthocyanin concentration, expressed as g cyanidin 3-O-glucoside equivalents per kg fresh weight, as follows:

$$TMA \text{ (g CGE kg}^{-1} \text{ FW)} = (A \times DF \times MW \times V \times \text{DM}) / (\varepsilon \times W \times l \times 10)$$

Where $A = (A_{520\ nm} - A_{700\ nm})$ pH = 1.0 - $(A_{520\ nm} - A_{700\ nm})$ pH = 4.5; *DF* = Dilution factor; *MW* (Molecular weight) = 449.2 g mol for cyanidin 3-O-glucoside; *V* = total extraction volume (mL); DM = g dry matter per kg fresh weight; ε (Molar extinction

coefficient) = 26,900 l cm^{-1} M^{-1} in aqueous solution; *W* = sample weight (g); *l* = pathlenght (cm).

Extraction and quantitative determination of total carotenoids (TC)

Total carotenoids were extracted and quantified according to Lachman et al. [29] with some modifications. Total carotenoids were extracted from fresh sample (15 g) after cryogenic grinding with 10 mL of chilled acetone. Borosilicate tubes with acetonic extracts were covered with tinfoil to prevent light activity and stored 3 days at 4 °C. After this period, the borosilicate tubes were put in an ultrasound bath and sonicated for 20 min and centrifuged at 7730 *g* for 10 min at 4 °C. The supernatant was collected, 10 mL of cold acetone were added to the pellet, the solid was re-suspended by shaking in a vortex for 1 min and the mixture was centrifuged. The extraction process was repeated once again and the filtrates were made up to 25 mL. The absorbance was measured at 444 nm using acetone as blank and the total carotenoid content was expressed as g lutein equivalents per kg FW from the equation:

$$TC\ (g\ LE\ kg^{-1}\ FW) = (A_{444\ nm} \times V \times 15)/\ (0.259 \times W \times 10^3)$$

Where $A = A_{444\ nm}$; V = total extraction volume (mL); W = sample weight (g).

Extraction and quantitative determination of hydrophilic antioxidant capacity (HAC)

Hydrophilic antioxidant capacity was analysed following two methods, ABTS (2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) and DPPH (2,2-diphenylpicrylhydrazyl),

according to Choong et al. [30]. These two indicator radicals were neutralized either by direct reduction via electron transfers or by radical quenching via H atom transfer, respectively. The ABTS assay measures the relative ability of antioxidants to scavenge the radical ABTS generated in aqueous phase compared with a Trolox standard (vitamin E analogue). The DPPH assay is based on the loss of absorption of radical DPPH when reduced by antioxidants. Both methods are widely used to determine antioxidant capacity of fruits, vegetables and beverages [31]. Fresh samples (2.5 g) were grinded in a mortar in liquid nitrogen. Hydrophilic antioxidants were extracted with 10 mL MeOH: H_2O (70:30, v/v). The mixture was centrifuged at 7730 *g* for 10 min at 4 °C, and the supernatant containing the extracted antioxidant was collected. The solid was suspended by shaking in a vortex for 1 min and the extraction operation was repeated twice again with the pellet and the final volume carried to 30 mL. The $ABTS^{\bullet+}$ solution was prepared by mixing 8 mM of ABTS salt with 3 mM of potassium persulfate in 25 mL of DIH_2O. The solution was held at room temperature in the dark for 16 h before use. The $ABTS^{\bullet+}$ solution was diluted with MeOH: H_2O (70:30, v/v) in order to obtain an absorbance between 0.8 and 0.9 at 734 nm. Antioxidant or standard solutions, 20 µL, were mixed with 980 µL of diluted $ABTS^{\bullet+}$ solution and incubated at room temperature. A reaction of 30 min was used for all the ABTS assays. In the DPPH assay, aliquots of the hydrophilic extracts were diluted (1:10, v/v) and 0.1 mL of the diluted sample was added to 3.9 mL of $DPPH^{\bullet}$ MeOH: H_2O (70:30, v/v) solution (6×10^{-5} mol L^1) to initiate the reaction. Absorbance was measured at 516 nm. A reaction of 3 h was used for all the DPPH assays. In both methods MeOH: H_2O (70:30, v/v) was used as blank and trolox MeOH: H_2O (70:30, v/v) dilutions were used as standard (0,100, 200, 300, 400 and 500 µM). The antioxidant activity was reported in mol trolox equivalents per kg fresh weight.

Statistical analysis

ANOVA and correlation matrix

One-way ANOVA was used to analyse the differences in the concentration of total soluble phenolics, total monomeric anthocyanins, total carotenoids, hydrophilic antioxidant capacity and vitamin C among coloured potato genotypes. A p value ≤ 0.05 was considered to be significant. Correlation analyses between parameters were calculated by using the CORR procedure of the SAS package [32].

PLS-DA

Partial Least Squares Discriminant Analysis (PLS-DA) is a linear supervised classification technique based on the PLS regression algorithm [33]. However, in PLS-DA a dummy response matrix Y must be introduced to account for classification problems. This matrix must be binary-coded (0, 1) with the same number of rows as X and same columns as groups object of study [34]. PLS-DA explains the variability presented in the data by creating linear combinations of the originals variables, called Latent Variables (LVs). The first LV covers the most variation in the data while the second most of the remaining and so on [35]. In this study, 3 PLS-DA models were performed to classify the varieties into three groups according to their phytochemical concentration level and called: low content (LC), mid content (MC) and high content (HC). The first PLS-DA model corresponded to varieties grouped according to their content of total soluble phenolics (TSP), total monomeric anthocyanins (TMA) and hydrophilic antioxidant capacity (HAC), since these three appeared to be well correlated in this study. Correlation between these three

phytochemicals was also found by other authors [14]. The second PLS-DA performed comprised groups categorised by their total carotenoids (TC) content. Finally, a third PLS-DA was accomplished covering varieties grouped by means of VC level.

Consequently, a 3-column response Y matrix was introduced for each PLS-DA in which varieties defined as having low content of phytochemicals were described by the dependent vector [1 0 0], varieties belonging to MC group, by the vector [0 1 0] and the ones belonging to the HC group by the vector [0 0 1]. Samples of the different varieties were randomly divided into calibration and prediction sets corresponding to 70% and 30% of samples respectively. Only the calibration data set was used to build the classification model while the prediction data set was used to externally evaluate its capability to classify new samples. Data were pre-processed by Multiplicative Scatter Correction (MSC) and Mean centre (MnC). Pre-processing methods are commonly used to reduce or avoid the influence of unwanted effects in the data that could negatively affect the consistency of the model [36-37]. MSC reduces scatter effects in the data [38] while MC reduces systematic noise [37]. The cross-validation (CV) method employed was Venetian Blinds with 10 data (splits) subsets since this method is considered simple and easy to implement. In this study the accuracy of the models was evaluated by the percentage of correctly classified samples in each group in calibration (Cal), cross-validation (CV) and Prediction (Pred). Both pre-processing of data and PLS-DA were performed in the PLS-Toolbox (Eigenvector Research Inc, Wenatchee, USA) working under Matlab R2014a (The Mathworks, MS, Natick, USA).

RESULTS

Chemical analysis

The tested potato cultivars showed significant differences among each other in terms of total soluble phenolics (TSP), total monomeric anthocyanins (TMA), total carotenoids (TC) and vitamin C (VC) ($p \leq 0.05$), thus genotype greatly affected all measured phytochemicals (Table 2). TSP concentrations, ranging from 0.140 ± 0.0208 to 2.78 ± 0.0512 g GAE kg^{-1} FW, were higher in the genotypes Violet Queen, Purple Peruvian and Highland Burgundy Red, while lowest TSP contents were found in NK-08/360, NK-08/349, NK-08/362 and Morea. TMA values ranged from 0.000100 ± 0.0000100 to 1.33 ± 0.0111 g CGE kg^{-1} FW. Highest TMA concentrations were also measured in the cultivars Violet Queen, Purple Peruvian, Highland Burgundy Red and Vitelotte, while lowest TMA concentrations were found in NK-08/360, Kasta, Rosa Roter and Morea. There is a six-fold variation in TC values which ranged from 0.00915 ± 0.00135 to 0.0590 ± 0.00425 g LE kg^{-1} FW. TC concentrations were higher in the cultivars Morada, Highland Burgundy Red, Rouge de Flandes and Rosa Roter, whereas Bleu de La Manche, Fenton and Blue Congo showed the lowest TC values. VC concentrations ranging from 0.0366 ± 0.0149 to 0.107 ± 0.0204 g kg^{-1} FW showed a three-fold variation among the collection of purple and red fleshed tubers. VC values were higher in Blue Congo, Morada and Kasta, while Rosa Roter, NK-08/349 and Highland Burgundy Red showed the lowest VC concentrations.

Statistical analyses revealed considerable differences in hydrophilic antioxidant capacity (HAC) between cultivars at $p \leq 0.05$ (fig. 4). The highest HAC value was measured in the cultivar Violet Queen (HAC_{ABTS} = 0.00939 ± 0.000171; HAC_{DPPH} = 0.00944 ± 0.0000656 mol TE kg^{-1} FW), while Morea showed the lowest HAC value

(HAC$_{ABTS}$ = 0.00420 ± 0.000315; HAC$_{DPPH}$ = 0.00255 ± 0.000136 mol TE kg^{-1} FW). The HAC measured in Violet Queen was between 125 and 270% higher than the HAC of the cultivar Morea, depending on the analytical method. The statistical analysis revealed very strong correlations ($p \leq 0.001$) and correlation coefficients higher than 0.78 among the following pairs of variables: TSP - TMA, TSP - HAC.ABTS, TMA - HAC.DPPH, TSP – HAC.DPPH and TMA - HAC-DPPH (Table 3). The highest r value was found for the HAC measured by ABTS and DPPH assays (r = 0.91), while other pairs of variables, such as TC – HAC, VC - TSP, VC - TMA, VC - HAC and VC - TC, showed r values close to 0.

PLS-DA

In this study, in order to perform a PLS-DA, once the phytochemical content of each variety was determined, samples were arbitrary divided into three groups according to the level of those compounds. As previously mentioned, 3 different PLS-DA models were performed.

Total soluble phenolics (TSP), total monomeric anthocyanins (TMA) and hydrophilic antioxidant capacity (HAC) model

Fig. 1 shows the varieties included in each of the three groups in the first PLS-DA carried out. It should be mentioned that only 16 varieties were included in this model, excluding Kasta and Rosa Roter since those two did not show correlation between TSP, TMA and HAC values. Samples (n= 429) were randomly divided between calibration and validation sets (n$_{cal}$= 302, n$_{val}$= 127). As it is shown in Fig 1, LC group comprised 165 samples belonging to 6 categories defined as having less than 0.900 g GAE kg^{-1} FW of TSP, 0.125 g

CGE kg^{-1} FW of TMA and 0.00400 mol TE kg^{-1} FW of HAC. It should be noted that variety NK-08/362 was included in this group only attending to their TSP and HAC content, because if according to its TMA content, it must have been included in the second group (MC).

The second group, MC, covered 132 samples belonging to 5 varieties with levels of TSP, TMA and HAC of 0.900 – 1.20 g GAE kg^{-1} FW, 0.125 – 0.240 g CGE kg^{-1} FW and 0.00400 – 0.00600 mol TE kg^{-1} FW respectively. Lastly, HC group also contained 132 samples from 5 different varieties with levels above 1.20 g GAE kg^{-1} FW, 0.240 g CGE kg^{-1} FW and 0.00600 mol TE kg^{-1} FW of TSP, TMA and HAC respectively.

As revealed before, PLS-DA models were evaluated in terms of correctly classified samples in each of the three groups (LC, MC and HC). In Table 4 the confusion matrix obtained with the PLS-DA can be seen. It shows the percentage of correctly classified samples of each of the 3 groups in calibration (Cal), cross-validation (CV) and prediction (Pred). The diagonal of each confusion matrix represents the percentage of samples correctly classified into the group they belong to while the values outside it correspond to wrongly classified samples, into a different group. A perfect classification corresponds to a matrix with a diagonal full of 100% surrounded by 0.00% rates.

In this study, a total of 9 LVs were used explaining 99.9% of the total variance. Slightly better classification rates were obtained for the LC group than for the rest of the groups. This was expected since samples into LC group presented values far away from the other two groups that had more similar values between each other. In any case, good rates above 80.0% of correctly classified samples were achieved for Cal and CV sets. However, LC group was the best classified for the three Cal, CV and Pred sets with more than 90.0% of samples correctly classified. It should be mentioned that the percentages of samples from

LC group that were badly classified into the other two groups corresponded to Valfi and NK-08/362 varieties, the ones with the highest values of phytochemicals inside this group. Moreover, as mentioned before if only taking into account TMA content of samples, NK-08/362 should be included in MC group, therefore, some misclassification of this variety was expected.

Some misclassification was also found in MC group in which a few samples belonging to Blue Star and Fenton varieties, with the lowest values of phytochemicals, were classified into LC group. On the other hand, a few samples belonging to British Columbia Blue and Blue Congo were identified as having HC. Finally, in HC group, a small number of samples from Vitelotte and Rouge de Flandes were incorrectly classified into MC. In this group, again, those varieties were the ones with the lowest values of phytochemicals of the group.

The results obtained demonstrated that in this study, PLS-DA classification technique enabled identification of different varieties of potato with low, mid and high content of TSP, TMA and HAC.

Fig 5 is the graphical representation of the confusion matrix for each group included in this study. The horizontal black line indicates the threshold above a sample is assigned to a particular class. Therefore, in Fig 5 (a), all samples located above the threshold are classified as belonging to LC group. It can be seen that there are a few samples belonging to MC and HC groups that according to this plot were classified as LC. Similarly, all the samples above the threshold in Fig 5 (b) and 5 (c) were classified as belonging to MC and HC respectively. In these plots, the prediction set of samples appeared inside a circle for the sake of easier visualization.

Total carotenoids model

Fig. 2 shows the varieties included in each of the three groups in the second PLS-DA. All 18 varieties were included in this model. 70.0% of the total number of samples (n= 471) were randomly selected for the calibration set (n_{cal}= 315) and 30.0% for validation (n_{val}= 156). As it is observed in Fig. 2, LC group comprised 192 samples belonging to 7 categories with TC content below 0.0200 g LE kg^{-1} FW, while the second group (MC) has a total of 132 samples from to 5 varieties with TC values between 0.0200 and 0.0300 g LE kg^{-1} FW, lastly, HC group is formed by 147 samples from 6 different varieties with TC levels above 0.0300 g LE kg^{-1} FW.

A total of 8 LVs were selected in this PLS-DA model explaining the 99.96% of variance. Table 5 shows the percentage of correctly classified samples into the three groups. It is observed that it was not possible to obtain an accurate classification of both LC and MC groups since a small number of samples were correctly classified into the group they belonged to and the fact that a considerable percentage of samples were classified in either of the other two groups. Thus, 17.2% and 31.2% of samples belonging to LC were classified as MC and HC respectively while 39.1% and 30.43% of samples of MC group were classified as LC and HC respectively in the Pred set. On the other hand, a high percentage of correctly classified samples was achieved in the HC group of around 80%. These results, suggests that NIRS technology could be used for screening processes in order to identified samples with high content of TC.

Vitamin C model

In Fig. 3 the varieties included in each of the three groups in the third PLS-DA are shown. Once again, samples (n= 471) were randomly divided between calibration and validation

sets (n_{cal}= 315, n_{val}= 156). In this analysis, LC group comprised 174 samples belonging to 7 categories defined as having less than 0.0700 g kg^{-1} FW of VC, the second group, MC, covered 162 samples belonging to 6 varieties with VC content between 0.0700 and 0.0920 g kg^{-1} FW and finally, HC group contained 135 samples from 5 different varieties with levels above 0.0920 g kg^{-1} FW.

A total of 5 LVs were selected explaining the 99.8% of the variance. Low classification rates, below 63.0%, were obtained for this PLS-DA (data not shown) for Cal, CV and Pred sets. Besides, high misclassification was found between LC and HC groups. Around 30.0% of samples belonging to HC group were wrongly classified as LC in Cal, CV and Pred sets suggesting that results were not reliable. It is worth mentioning that vitamin C content in potatoes highly depends on many factors and can vary considerably from one campaign to another. Thus, it becomes very challenging to perform robust classification methods for this compound. Moreover, significantly lower contents of VC were found in this study compared to other authors as mentioned before, due to the effect of cold temperatures during storage. Therefore, we consider that these facts could be responsible for the low classification rates obtained in this work.

DISCUSSION

Phytochemical quantification

In the present study, the vitamin C concentrations were significantly lower than those reported by authors as Love et al. [39] or Han et al. [40], who found concentration values ranging from 0.115 to 0.420 g kg^{-1} FW in white or yellow fleshed commercial cultivars and breeding lines from North America. Similar to our data, Jimenez et al. [41] reported lower

vitamin C concentrations ranging from 0.0754 to 0.286 g kg^{-1} FW in a collection of seven Andean potato cultivars. According to the CD 2008/100/EC [42], an average edible portion (100 g of peeled tubers) contains about 4.6 to 13.3% of the recommended daily intake (RDA) of vitamin C. Moreover, considering the high losses during thermal processing from 25% to 40% [43-44] these cultivars should be considered only as a relatively poor source of vitamin C. However, numerous studies have shown that vitamin C levels are highly dependent on many factors, such as culture conditions, wounding and storage, which can alter rapidly it concentration in tubers [45-46]. According to Oba et al. [47] the vitamin C contents in potato tubers stored at 4 °C for one month may experience a decrease from 46 to 57%. Thus, the effect of cold storage on vitamin C could explain why our values were significantly lower than those reported by other authors.

TSP values showed wide variability among the collection of purple or red fleshed cultivars or breeding lines showing the great genotype effect on these phytochemicals. This is about a 20-fold variation in total soluble phenolics and the highest TSP values are generally comparable to those of blueberries (*Vaccinium corymbosum* L.) [48-49]. Accordingly, a study of 74 Andean potato landraces found about an 11-fold variation in TSP values [10]. Despite the fact that published data vary widely, those measured by Folin Ciocalteau method usually show lower value range from 0.0200 to 1.00 g GAE kg^{-1} FW [50-52] due to the limits of the method compared to the Fast Blue method used in this study [53]. Our results are in agreement with previous works reporting higher TSP values in coloured potato tubers, for instance, [54] found TSP levels about 0.250 and 2.87 g GAE kg^{-1} FW, while Stushnoff et al. [55] reported values between 0.900 and 4.00 g GAE kg^{-1} FW in a collection of coloured accessions.

Total monomeric anthocyanins (TMA) values found in the present tubers (from 0.000100 to 1.33 g CGE kg^{-1} FW) were similar to almost published total monomeric anthocyanin values in purple or red potato tubers. Reyes et al. [14] found TMA values between 0.210 and 0.550 g CGE kg^{-1} FW in red fleshed tubers, and between 0.110 and 1.75 g CGE kg^{-1} FW in purple fleshed ones. In fact, 1.75 g CGE kg^{-1} FW is an extremely high value for potato which has not been equalled in consulted bibliography. In this respect, it should be noted that potato peels contain higher anthocyanin levels and the evaluation of coloured cultivars and breeding lines by Reyes et al. [14] was performed using unpeeled tubers. Brown et al. [56] reported TMA values between 0.150 and 0.400 g CGE kg^{-1} FW in 18 pigmented clones and cultivars. Jansen and Flamme [57] also found TMA contents ranging from 0 to 0.800 g CGE kg^{-1} FW in 31 potato accessions. These highest values are nearly 2.8 and 1.5 times less than the TMA content of Violet Queen in each case. TMA content value of Violet Queen is comparable to that of *Vaccinum* L. berries, which are considered to be a one of the richest natural sources of anthocyanins with TMA contents ranging from 1.38 to 3.85 g CGE kg^{-1} FW [58]. However, the daily intake of potatoes usually is much higher than the daily intake of blueberries [59].

TMA values were well correlated with TSP and hydrophilic antioxidant capacity. The correlation between antioxidant capacity and phenolics has also been reported [8, 15]. Present values of hydrophilic antioxidant capacity were higher than that of white or yellow fleshed potatoes [58], but our values are similar to those obtained by Brown et al. [56] in coloured potato breeding lines. In relation to the total carotenoid concentrations, values ranging from 0.00910 to 0.0588 g LE kg^{-1} FW are comparable to those obtained by Lachman et al. [29] using the same colorimetric method (0.000200 – 0.0250 g LE kg^{-1} FW). The lipophilic character of most carotenoids can explain the lack of correlation between TC and HAC.

With the exception of Violet Queen, higher TC values were detected in tubers with a perceptible partially yellow coloration, such as Morada, Rosa Roter, Rouge de Flandes and Highland Burgundy Red. Relatively low TC values were found in most of the medium or deep purple cultivars. According to Kotíková et al. [61], deep purple potato cultivars generally have lower ability to synthesize and accumulate carotenoids when compared to yellow fleshed potato cultivars.

According to several authors, purple and red fleshed genotypes are a great source of certain minerals, phenolic compounds and antioxidant capacity. We have identified cultivars and breeding lines with very high concentrations of carotenoids (Morada, Highland Burgundy Red and Violet Queen), vitamin C (Blue Congo, Morada and Kasta), soluble phenolics, monomeric anthocyanins and hydrophilic antioxidant capacity (Violet Queen, Purple Peruvian and Vitelotte). Besides the commercial, gastronomical and medical importance of quality parameters in coloured tubers, the identification of phytochemical high producing germplasm is a key step to develop a potato breeding program for nutritional quality.

Sample classification by NIR spectroscopy

Some authors have studied the non-destructive determination of some chemical compounds in potatoes. Thus, polyphenols content of lyophilized potatoes was accurately predicted by using Fourier transform infrared spectroscopy achieving correlation values of 0.99 [62]. Moreover, other authors obtained good correlation coefficients when estimating the total and individual carotenoid concentration in potatoes by NIR spectroscopy [63].

As explained before, there is a little literature concerning qualitative analysis of potatoes by NIR spectroscopy despite its potential. Even so, some authors successfully

investigated the discrimination of two categories of potato samples regarding their recoverable protein content. An overall 87.5% of correctly classified samples was achieved in that study [23]. Other authors were able to classify potato chips by source of frying oil by combining NIR spectroscopy and Soft Independent Modeling of Class Analogy (SIMCA) [64].

In this study we have successfully classified samples into three levels according to their content of TSP, TMA and HAC. Figure 5 (a, b & c) confirms the results obtained, as it shows that the majority of the samples of each group are correctly classified into them. However, there was some misclassification as previously described, but we can say that overall, very good rates of discrimination were obtained.

Therefore, we can say that according to this study, NIR spectroscopy combined with PLS-DA was capable of accurately identifying samples containing different levels of TSP, TMA and HAC belonging to this collection of 18 purple and red-fleshed potatoes.

Regarding TC content, we found that NIRs technique was only capable of identifying samples with high content of this compound among the varieties analysed. Information that could be useful for screening processes.

Finally, the classification results achieved according to the content of Vitamin C suggested that it was not possible to obtained a reliable classification of varieties regarding their VC content by NIRS technology.

Nevertheless, these findings are of great importance considering the continuously increasing demand for quality control of food products among consumers and authorities [65]. The outcome from this study could be considered as a screening step for future potato breeding programs. Further research is advisable including a larger set of samples comprising not only coloured varieties but yellow skin tuber as well.

CONCLUSION

The identification of potato genotypes high in phenolic compounds, carotenoids and antioxidant capacity is a key step for both identifying phytochemical rich food products and for developing breeding lines with high concentrations of bioactive compounds. The automatic and non-destructive characterization of cultivars and breeding lines with different levels of bioactive compounds through near-infrared spectroscopy can be also suitable for discriminating and classifying potato tubers in terms of phytochemical content.

ACKNOWLEDGMENTS

This work was financed within the frame of INIA's project RTA2013-00006-C03 and the Basque Government.

REFERENCES

1 Navarre DA, Goyer A and Shakya R, Nutritional value of potatoes: Vitamin, phytonutrient, and mineral content. In: Advances in potato chemistry and technology, Singh J. and L. Kaur. (Eds.). Oxford, UK: Elsevier Publication. pp 395-424 (2009).

2 Storey M, The Harvested Crop: Nutritional Value. In: Potato biology and biotechnology. Advances and perspectives. Vreugdenhil D, J Bradshaw, CH Gebhardt, F Govers, DKL Mackerron, MA Taylos, and HA Ross (Eds.) Oxford, UK: Elsevier Publication. pp 442-443 (2007).

3 Grosso G, Bei R, Mistretta A, Marventano S, Calabrese G, Masuelli L, Giganti MG, Modesti A, Galvano F and Gazzolo D, Effects of vitamin C on health: a review of evidence. *Frontiers Biosci* **18**: 1017-1029 (2013).

4 Rice-Evans CA, Miller JM and Paganga G, Structure-antioxidant activity relationship of flavonoids and phenolic acids. *Free Rad Bio Med*. 20, 933-956 (1996).

5 Brown CR. 2005. Antioxidants in potato. *Am J Pot Res* **82**: 163-172.

6 Kita A, Bąkowska-Barczak A, Hamouz K, Kułakowska K and Lisińska G, The effect of frying on anthocyanin stability and antioxidant activity of crisps from red- and purple-fleshed potatoes (*Solanum tuberosum* L.). *J Food Compos Anal* **32**: 169-175 (2013).

7 Marko D, Puppel N, Tjaden Z, Jakobs S and Pahlke G, The substitution pattern of anthocyanidins affects different cellular signaling cascades regulating cell proliferation. *Mol Nutr Food Res* **48**: 318-325 (2004).

8 Andre CM, Ghislain M, Bertin P, Oufir M, Herrera MR, Hoffmann L, Hausman JF, Larondelle Y and Evers D, Andean potato cultivars (*Solanum tuberosum* L.) as a source of antioxidant and mineral micronutrients. *J Agric Food Chem* **55**: 366–378 (2007).

9 Rodriguez-Amaya DB and Kimura M, HarvestPlus Handbook for carotenoid analysis. HarvestPlus Technical Monograph 2. Washington, DC and Cali. International Food Policy Research Institute (IFPRI) and International Center for Tropical Agriculture (CIAT) (2004).

10 Rao AV and Rao LG, Carotenoids and human health. *Pharm Res* 55: 207–216 (2007).

11 Lu W, Haynes K, Wiley E and Clevidence B, Carotenoid content and color in diploid potatoes. *J Am Soc Hort Sci* **126**: 722–726 (2001).

12 Griffiths W, Dale F, Morris W and Ramsay G, Effects of season and postharveststorage on the carotenoid content of *Solanum phureja* potato tubers. *J Agric Food Chem* **55**, 379–385 (2007).

13 Burgos G, Salas E, Amorós W, Auqui M, Muñoa L, Kimura M and Bonierbale M, Total and individual carotenoid profiles in *Solanum phureja* of cultivated potatoes: I. Concentrations and relationships as determined by spectrophotometry and HPLC. *J Food Compos Anal* **22**: 503-508 (2009).

14 Reyes LF, Miller JC and Cisneros-Zevallos L, Antioxidant capacity, anthocyanins and total phenolics in purple- and red-flesh potato *(Solanum tuberosum L.)* genotypes. *Am J Pot Res* **82**: 271-277 (2005).

15 Hamouz K, Lachman J, Dvořák P, Jůzl M and Pivec V, The effect of site conditions, variety and fertilization on the content of polyphenols in potato tubers. *Plant, Soil & Env* **52**: 407–412 (2006).

16 Chen JY, Zhang H, Yelian M and Asakura M, Nondestructive determination of sugar content in potato tubers using visible and near infrared spectroscopy. *JPN J Food Eng* **11**: 59-64 (2010).

17 Magwaza LS, Opara UL, Nieuwoudt H, Cronje PJR, Saeys W and Nicolaï B, NIR spectroscopy applications for internal and external quality analysis of citrus fruit- A Review. *Food Bioprocess Tech* **5**: 425-444 (2012).

18 Gunasekaran S and Irudayaraj J, Nondestructive Food Evaluation. Techniques to analyse properties and quality, optical methods: visible NIR and FTIR spectroscopy. New York: Marcel Dekker (2000).

19 Nicolaï BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI and Lammertyn J, Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol Tech* **46**: 99-118 (2007).

20 Lister SJ, Dhanoa MS, Stewart JL and Gill M, Classification and comparison of< i> Gliricidia</i> provenances using near infrared reflectance spectroscopy. *Anim Feed Sci Tech* **86**: 221-238 (2000).

21 López A, Jarén C, Arazuri S, Mangado J, Tierno R, Ruiz de Galarreta JI and Riga P, Estimation of the total phenolic content in potatoes by NIRS. In: International Conference of Agricultural Engineering (Zurich, Switzerland), P0640 (2014).

22 López A, Arazuri S, Jarén C, Tierno R, Ruiz de Galarreta JI and Riga P, Ability of NIRS to estimate total phenolic content of lyophilized potatoes. In: Núster, C.E.L., L.E.M. Rodríguez, and TV. Mosquera (Eds.) XXVI Congreso Latinoamericano de la papa – ALAP (Bogotá, Colombia), Calidad Nutricional, 105 (2014).

23 Fernández-Ahumada E, Garrido-Varo A, Guerrero-Ginel JE, Wubbels A, Van der Sluis C and Van der Meer JM, Understanding factors affecting near infrared analysis of potato constituents. *J Near Infrared Spec* **14**: 27-35 (2006).

24 Lammertyn JPA, De Baerdemaeker J and Nicolaï B, Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment. *Meas Sci Technol* **18**: 121-32 (2000).

25 López A, Arazuri S, Mangado J, Jarén C and Ruiz de Galarreta JI, Determinación de la penetración NIR en patata. VII Congreso Ibérico de Agroingeniería y Ciencias Hortícolas; Madrid (España): Fundación General de la Universidad Politécnica de Madrid (2013).

26 CAMAG, Vitamin C in fruit juice (Quantitative determinations). CAMAG Application Note A-10.5, CAMAG, Muttenz, Switzerland. Home page: http://www.camag.com/en/tlc_hptlc/camag_laboratory/methods.cfm, accessed January 2014. CAMAG, Application Notes (2012).

27 Medina MB, Simple and rapid method for the analysis of phenolic compounds in beverages and grains. *J Agric Food Chem* **59**: 1565–1571 (2011).

28 Giusti M and Wrolstad RE, Current protocols in food analytical chemistry, in Current Protocols in Food Analytical Chemistry, I. John Wiley and Sons, Editor. John Wiley and Sons, Inc (2001).

29 Lachman J, Hamouz K, Hejtmánková A, Dudjak J, Orsák M and Pivec V, Effect of white fleece on the selected quality parameters of early potato (*Solanum tuberosum* L.) tubers. *Plant, Soil & Env* **49**: 370-377 (2003).

30 Choong CT, Van-Den TA, Roger A, McFeeters F, Roger L, Thompson KV, Pecota G and Craig Y, Antioxidant activities, phenolic and β- carotene contents of sweet potato genotypes with varying flesh colours. *Food Chem* **103**: 829-838 (2007).

31 Floegel A, Kim DO, Chung SJ, Koo SI and Chun OK, Comparison of ABTS/DPPH assays to measure antioxidant capacity in popular antioxidant-rich US foods. *J Food Compos Anal* **24**: 1043-1048 (2011).

32 SAS, SAS/STAT 9.3. User'Guides Survey Datas Analysis. SAS Institute Inc, Cary, North Carolina, USA (2011).

33 Ballabio D and Consonni V, Classification tools in chemistry. Part 1: linear models. PLS-DA. *Anal Methods* **5**: 3790-3798 (2013).

34 Vitale R, Bevilacqua M, Bucci R, Magrì AD, Magrì SL and Marini F, A rapid and non-invasive method for authenticating the origin of pistachio samples by NIR spectroscopy and chemometrics. *Chemometr Intell Lab* **121**: 90-99 (2013).

35 Barker M and Rayens W, Partial least squares for discrimination. *J Chemometr* **17**: 166-173 (2003).

36 Amigo JM, Practical issues of hyperspectral imaging analysis of solid dosage forms. *Anal Bioanal Chem* **398**: 93-109 (2010).

37 Barbin DF, ElMasry G, Sun DW and Allen P, Predicting quality and sensory attributes of pork using near-infrared hyperspectral imaging. *Anal Chim Acta* **719**: 30-42 (2012).

38 Rinnan Å, Van den Berg F and Engelsen SB, Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends Anal Chem* **28**: 1201-1222 (2009).

39 Love SLST, Shafii B, Price WJ, Mosley AR and Thornton RE, Stability of expression and concentration of ascorbic acid in North American potato germplasm. *Hort Sci*. **39**: 156-160 (2004).

40 Han J, Kozukue N, Young K, Lee K and Friedman M, Distribution of ascorbic acid in potato tubers and in home-processed and commercial potato foods. *J Agric Food Chem* **52**: 6516–6521 (2004).

41 Jimenez ME, Rossi AM and Samman NC, Phenotypic, agronomic and nutritional characteristics of seven varieties of Andean potatoes. *J Food Compos Anal* **22**: 613-616 (2009).

42 CEC, CD 2008/100/EC of 28 October 2008 amending Council Directive 90/496/EEC on nutrition labelling for foodstuffs as regards recommended daily allowances, energy conversion factors and definitions. *Official Journal of the European Union* **285**: 9-12. Home page:http://eurex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:285:0009:0012:EN:PDF, accessed June 2014. CEC, Commission of the European Communities (2008).

43 Hagg M., Hakkinen U, Kumpulainen J, Ahvenainen R and Hurme E, Effects of preparation procedures, packaging and storage on nutrient retention in peeled potatoes. *J Sci Food Agric* **77**: 519–526 (1998).

44 Vanderslice JT, Higgs DJ, Hayes JM and Block G, Ascorbic acid and dehydroascorbic acid content of foods-aseaten. *J Food Compos Anal* **3**: 105–118 (1990).

45 Tudela JA, Espín JC and Gil MI, Vitamin C retention in fresh-cut potatoes. *Postharvest Biol Tech* **26**: 75–84 (2002).

46 Dale MFB, Griffiths DW and Todd DT, Effects of genotype, environment, and postharvest storage on the total ascorbate content of potato (*Solanum tuberosum* L.) tubers. *J Agric Food Chem* **51**: 244-248 (2003).

47 Oba K, Yamamoto A, Ohara A, Ishii G and Omemura Y, Effect of storage temperature on vitamin C levels and L-galactonolactone dehydrogenase activity of potato tubers. *J Jpn Soc Food Sci Technol* **45**: 510-513 (2010).

48 Zheng W and Wang SY, Oxygen radical absorbing capacity of phenolics in blueberries, cranberries, chokeberries, and lingonberries. *J Agric Food Chem* **51**: 502-509 (2002).

49 You Q, Wang B, Chen F, Huang Z, Wang X and Luo PG, Comparison of anthocyanins and phenolics in organically and conventionally grown blueberries in selected cultivars. *Food Chem* **125**: 201-208 (2010).

50 Hu C, Tsao R, Liu R, Sullivan JA and McDonald MR, Influence of cultivar and year on phytochemical and antioxidant activity of potato (*Solanum tuberosum* L.) in Ontario. *Can J Plant Sci* **92**: 485-493 (2012).

51 Aziz A, Yasin M, Randhawa MA, Yasmin A, Yahangir MA and Sohail M, Nutritional and antioxidant profile of some selected Pakistani potato cultivars. *Pak J Food Sci* **23**: 87-93 (2013).

52 Murniece I, Kruma Z, Skrabule I and Vaivode A, Carotenoids and Phenols of Organically and Conventionally Cultivated Potato Varieties. *IJCEA* **5**: 342-348 (2013).

53 Lester GE, Lewers KS, Medina MB and Saftner RA, Comparative analysis of strawberry total phenolics via Fast Blue BB vs. Folin–Ciocalteu: Assay interference by ascorbic acid. *J Food Compos Anal* **27**: 102–107 (2012).

54 Madiwale G, Effect of genotype, storage and processing on the polyphenolic content, composition, in vitro anti-cancer activity and sensory attributes of colored-flesh potatoes (PhD Thesis). Colorado State University, Fort Collins, Colorado (2012).

55 Stushnoff C, Holm D, Thompson M, Jiang W, Thompson H, Joyce N and Wilson P, Antioxidant properties of cultivars and selections from the Colorado Potato Breeding Program. *Am J Pot Res* **85**: 267-276 (2008).

56 Brown CR, Culley D, Yang CP, Durst R and Wrolstad R, Variation of anthocyanin and carotenoid contents and associated antioxidant values in potato breeding lines. *J Amer Soc Hort Sci* **130**: 174-180 (2005).

57 Jansen G and Flamme W, Coloured potatoes (*Solanum Tuberosum* L.) - anthocyanin content and tuber quality. *Gen Res Crop Evol* **53**: 1321-1331 (2006).

58 Cevallos-Casals BA and Cisneros-Zevallos L, Stoichiometric and kinetic studies of phenolic antioxidants from Andean purple corn and red-fleshed sweetpotato. *J Agric Food Chem* **51**: 3313-3319 (2003).

59 FAO, FAOSTAT database. Home page: http://faostat.fao.org. FAO, Food and Agriculture Organization of the United Nations (2012).

60 Wu X, Beecher GR, Holden JM, Haytowitz DB, Gebhardt SE and Prior RL, Lipophilic and Hydrophilic Antioxidant Capacities of Common Foods in the United States. *J. Agric Food Chem* **52**: 4026-4037 (2004).

61 Kotíková Z, Hejtmánková A, Lachman J, Hamouz K, Trnková E and Dvorak P, Effect of selected factors on total carotenoid content in potato tubers (*Solanum tuberosum* L.). *Plant Soil & Env* **53**: 355 (2007).

62 Shiroma-Kian C, Tay D, Manrique L, Giusti MM and Rodriguez-Saona L, Improving the screening process for the selection of potato breeding lines with enhanced polyphenolics content. *J Agr Food Chem* **56**: 9835-9842 (2008).

63 Bonierbale M, Grüneberg W, Amoros W, Burgos G, Salas E, Porras E and Zum Felde. T, Total and individual carotenoid profiles in *Solanum phureja* cultivated potatoes: II. Development and application of near-infrared reflectance spectroscopy (NIRS) calibrations for germplasm characterization. *J Food Comp Anal* **22**: 509-516 (2009).

64 Shiroma C and Rodriguez-Saona L, Application of NIR and MIR spectroscopy in quality control of potato chips. *J Food Comp Anal* **22**: 596-605 (2009).

65 López A, Arazuri S, García I, Mangado J and Jarén C, A review of the application of near-infrared spectroscopy for the analysis of potatoes. *J Agr Food Chem* **61**: 5413-5424 (2013).

Figure captions:

Figure 1. Flowchart of varieties distribution according to their content of total soluble phenolics (TSP), total monomeric anthocyanins (TMA) and hydrophilic antioxidant capacity (HAC).

Figure 2. Flowchart of varieties distribution according to their content of total carotenoids (TC).

Figure 3. Flowchart of varieties distribution according to their content of vitamin C (VC).

Figure 4. Hydrophilic antioxidant capacity (HAC) measured by ABTS and DPPH methods. Means and standard deviations are expressed as mol trolox equivalents per kg FW. Average values (n = 3) in columns labeled with the same letter are not significantly different at $p \geq 0.05$ ($LSD_{ABTS} = 0.000474$; $LSD_{DPPH} = 0.000240$).

Figure 5. PLS-DA analysis of LC group (a), MC (b) and HC (c). Horizontal line (----) indicates the threshold above a sample is assigned to a particular group.

Table 1. Collection of 18 purple and red fleshed potato (*Solanum tuberosum* L.) cultivars or breeding lines.

Cultivar or breeding lines	Origin	Status	Skin/Flesh type[1]
Bleu de La Manche	France	Cultivar	P/P
Blue Congo	Sweden-UK	Cultivar	P/P
Blue Star	Netherlands	Cultivar	P/P
British Columbia Blue	Canada-UK	Cultivar	P/P
Fenton	Canada-UK	Cultivar	P/P
Highland Burgundy Red	France	Cultivar	R/R
Kasta	Spain	Cultivar	P/P
Morada	Spain	Cultivar	P/P
Morea	Spain	Cultivar	P/P
NK-08/349	Spain	Breeding line	P/P
NK-08/360	Spain	Breeding line	P/P
NK-08/362	Spain	Breeding line	P/P
Purple Peruvian	Peru	Cultivar	P/P
Rosa Roter	Peru	Cultivar	R/R
Rouge de Flandes	Belgium	Cultivar	R/R
Valfi	Sweden-UK	Cultivar	P/P
Violet Queen	Netherlands	Cultivar	P/P
Vitelotte	France	Cultivar	P/P

[1] Key to skin and tuber flesh types: R = Red, P = Purple.

Table 2. Total soluble phenolics (TSP), total monomeric anthocyanins (TMA), total carotenoids (TC) and vitamin C (VC) in peeled tubers of 18 purple and red fleshed potato cultivars and breeding lines. Means and standard deviations of 1) TSP are expressed as g gallic acid equivalents per kg fresh weight; 2) TMA are expressed as g cyanidin 3-glucoside equivalents per kg fresh weight; 3) TC are expressed as g lutein equivalents per kg fresh weight; 4) VC are expressed as g of vitamin C per kg fresh weight (n = 3).

Cultivar/ Breeding line	**TSP**	**TMA**	**TC**	**VC**
Bleu de La Manche	0.915±0.0159	0.230±0.0264	0.00915±0.00135	0.0831±0.00945
Blue Congo	1.18±0.0476	0.178±0.0235	0.00962±0.00125	0.107±0.0204
Blue Star	1.08±0.0117	0.129±0.0143	0.0193±0.000700	0.0907±0.00361
British Columbia Blue	1.12±0.0833	0.228±0.0316	0.0161±0.00200	0.0577±0.00971
Fenton	1.00±0.0914	0.172±0.0126	0.00915±0.00220	0.0847±0.0135
Highland Burgundy Red	2.14±0.0489	0.350±0.0102	0.0360±0.00360	0.0573±0.00957
Kasta	1.27±0.0530	0.00126±0.000300	0.0246±0.00235	0.0972±0.00455
Morada	0.514±0.00984	0.00686±0.00178	0.0590±0.00425	0.107±0.00612
Morea	0.400±0.0122	0.0320±0.00955	0.0292±0.00571	0.0912±0.0219
NK-08/349	0.376±0.0275	0.0545±0.0163	0.0236±0.00283	0.0501±0.0144
NK-08/360	0.140±0.0208	0.000100±0.0000100	0.0194±0.00152	0.0704±0.0120
NK-08/362	0.428±0.0252	0.158±0.00460	0.0326±0.00160	0.0869±0.0182
Purple Peruvian	2.15±0.0102	0.408±0.0170	0.0188±0.00113	0.0926±0.00973
Rosa Roter	1.09±0.0110	0.0359±0.0.06	0.0341±0.00515	0.0366±0.0149
Rouge de Flandes	1.81±0.0239	0.270±0.00584	0.0349±0.00155	0.0677±0.0109
Valfi	0.521±0.0162	0.123±0.00336	0.0203±0.00223	0.0589±0.0179
Violet Queen	2.78±0.0512	1.33±0.0111	0.0351±0.00117	0.0671±0.0110
Vitelotte	1.64±0.0901	0.436±0.00741	0.0282±0.00484	0.0931±0.0182
LSD (0.05)	0.107	0.0255	0.00475	0.0230

LSD: Least Significance Difference.

Table 3. Matrix for Pearson correlation coefficients (*r*) showing the linear relationship among total soluble phenolics (TSP), total monomeric anthocyanins (TMA), hydrophilic antioxidant capacity measured by ABTS and DPPH (HAC), total carotenoid content (TC) and vitamin C (VC).

	TSP	TMA	HAC.ABTS	HAC.DPPH	TC	VC
TSP	1.00					
TMA	0.789****	1.00				
HAC-ABTS	0.787****	0.871****	1.00			
HAC-DPPH	0.792****	0.799****	0.907****	1.00		
TC	0.0903	0.127	0.0371	-0.0681	1.00	
VC	-0.0466	-0.0651	-0.0225	0.111	-0.0310	1.00

Table 4. Confusion matrix of the three groups – low content (LC), mid content (MC) and high content (HC) - in the TSP, TMA and HAC PLS-DA model for Cal, CV and Pred.

			Actual Group (%)		
			LC	MC	HC
Calibration set (Cal)	**Predicted group (%)**	LC	**91.2**	2.47	0.00
		MC	4.90	**87.6**	9.09
		HC	3.92	9.88	**90.9**
Cross-validation set (CV)		LC	**90.2**	4.94	0.00
		MC	6.86	**82.7**	10.4
		HC	2.94	12.3	**89.6**
Prediction set (Pred)		LC	**97.6**	0.00	3.85
		MC	2.38	**83.8**	19.2
		HC	0.00	16.2	**76.9**

Table 5. Confusion matrix of the three groups – low content (LC), mid content (MC) and high content (HC) - in the TC PLS-DA model for Cal, CV and Pred.

			Actual Group (%)		
			LC	MC	HC
Calibration set (Cal)	**Predicted group (%)**	LC	**45.6**	26.8	4.25
		MC	20.0	**43.9**	14.9
		HC	34.4	29.3	**80.8**
Cross-validation set (CV)		LC	**46.4**	25.6	4.25
		MC	20.0	**42.7**	14.9
		HC	33.6	31.7	**80.8**
Prediction set (Pred)		LC	**51.6**	39.1	6.52
		MC	17.2	**30.4**	13.0
		HC	31.2	30.4	**80.4**

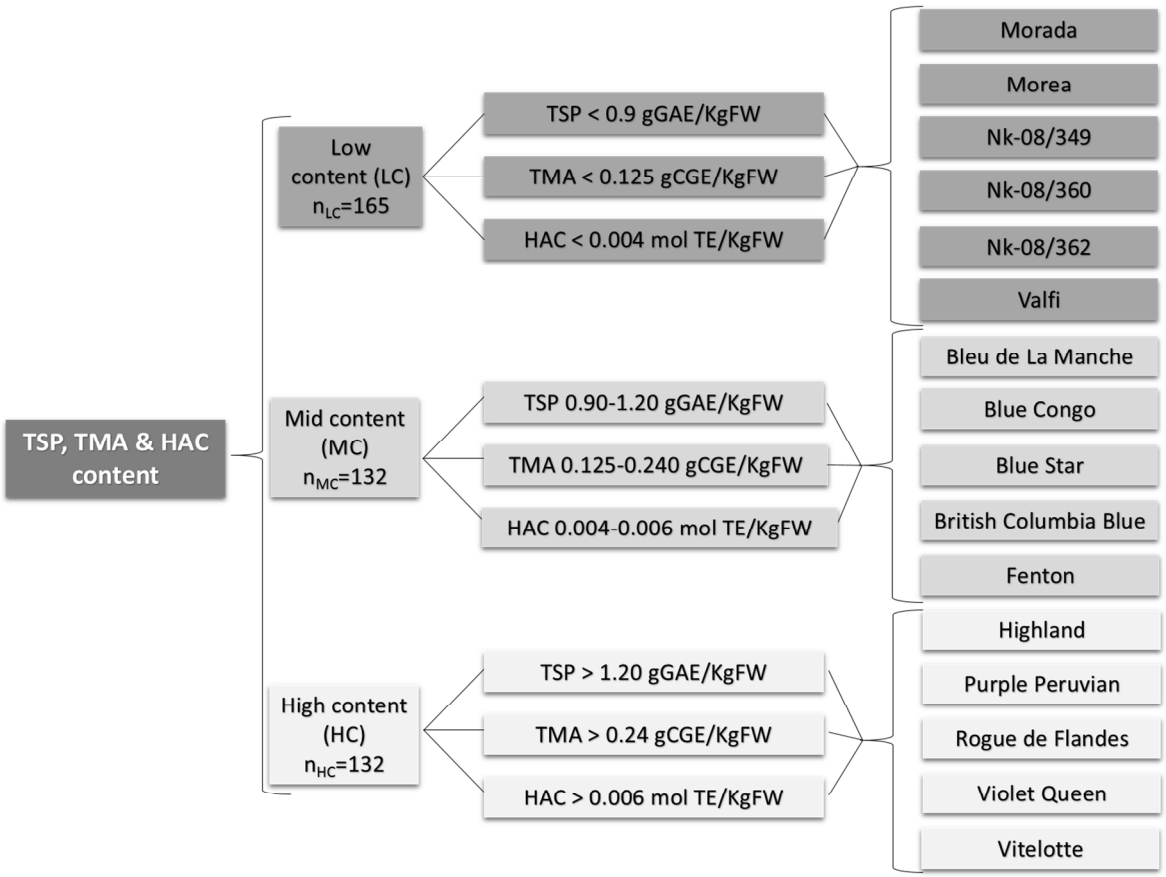
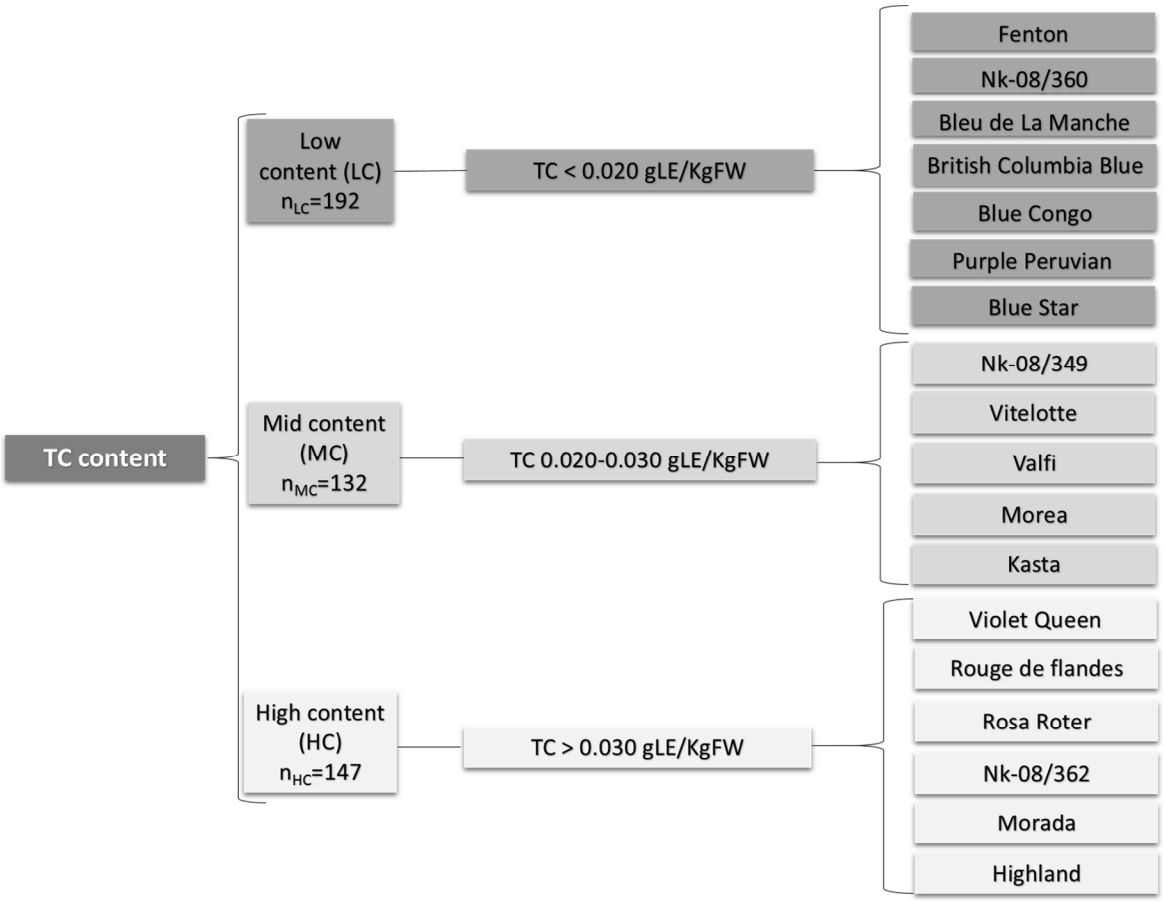

TSP, TMA & HAC content
Low content (LC) n_{LC}=165
Mid content (MC) n_{MC}=132
High content (HC) n_{HC}=132
TSP < 0.9 gGAE/KgFW
TMA < 0.125 gCGE/KgFW
HAC < 0.004 mol TE/KgFW
TSP 0.90-1.20 gGAE/KgFW
TMA 0.125-0.240 gCGE/KgFW
HAC 0.004-0.006 mol TE/KgFW
TSP > 1.20 gGAE/KgFW
TMA > 0.24 gCGE/KgFW
HAC > 0.006 mol TE/KgFW
Morada
Morea
Nk-08/349
Nk-08/360
Nk-08/362
Valfi
Bleu de La Manche
Blue Congo
Blue Star
British Columbia Blue
Fenton
Highland
Purple Peruvian
Rogue de Flandes
Violet Queen
Vitelotte

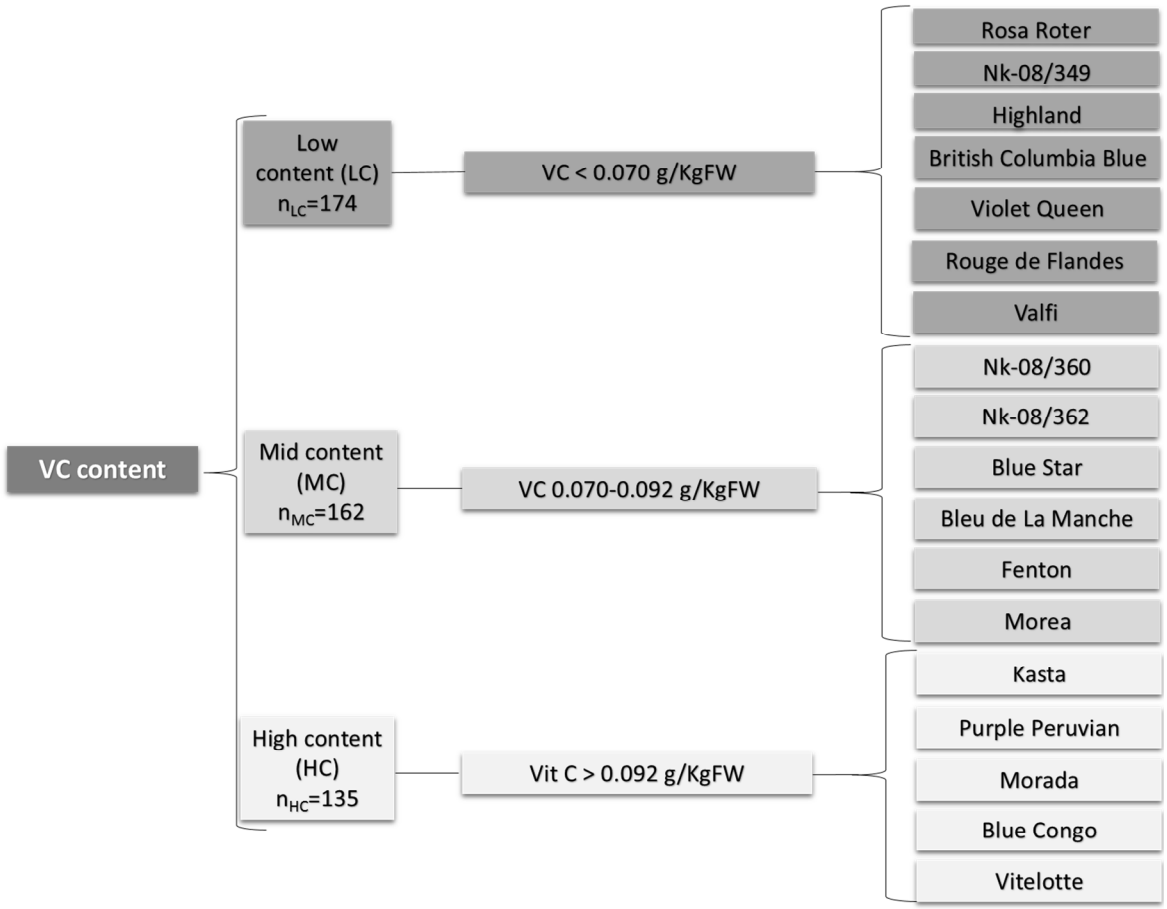
Fig.1

841



842

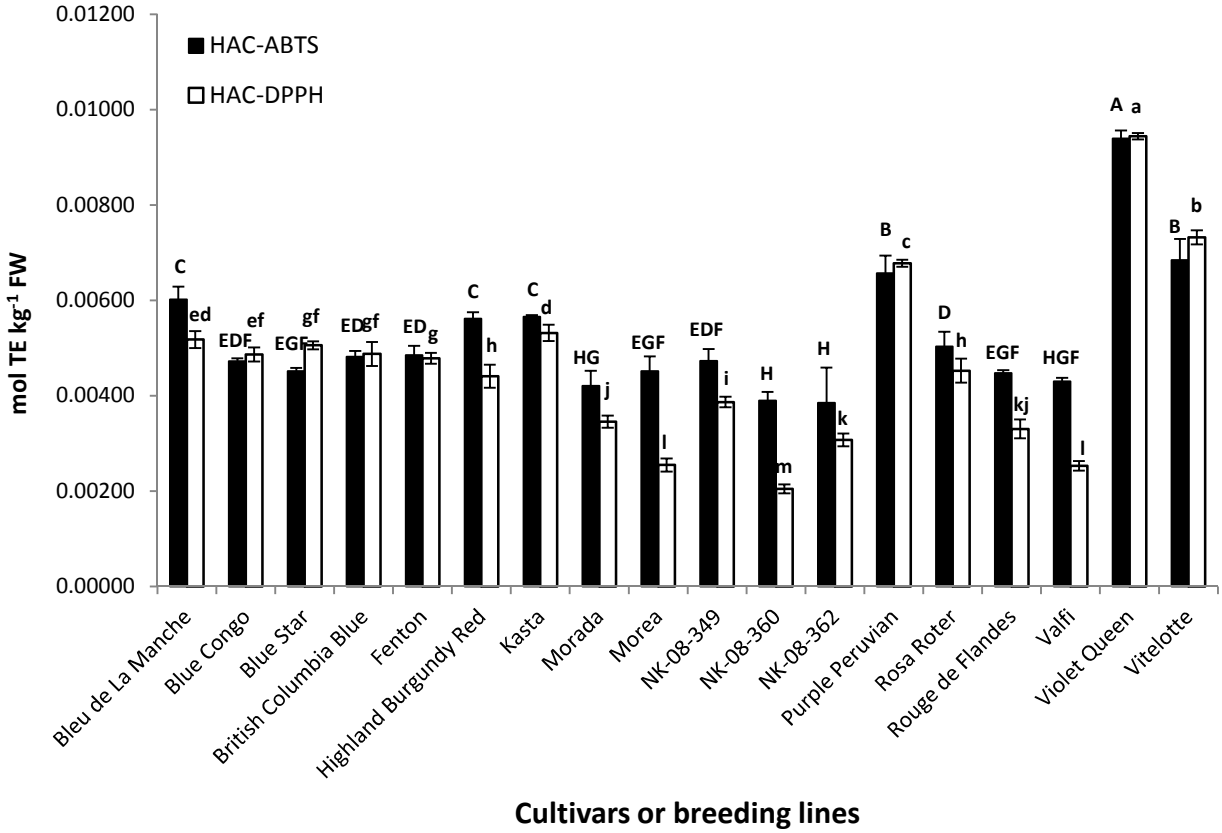
843 Fig. 2



844

845

846 Fig. 3



847

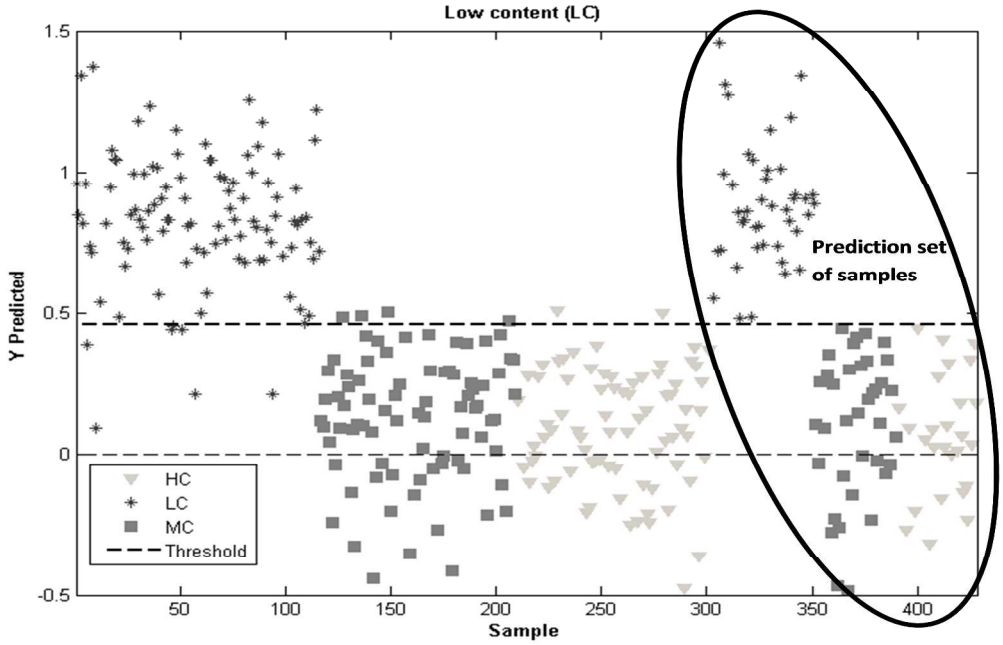
848

849

850 Fig.4

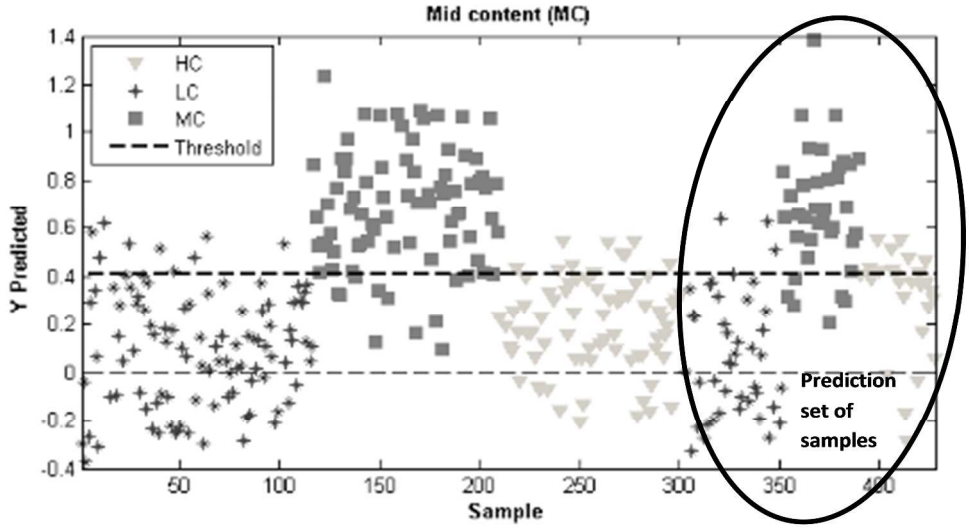
851

852 a)



853

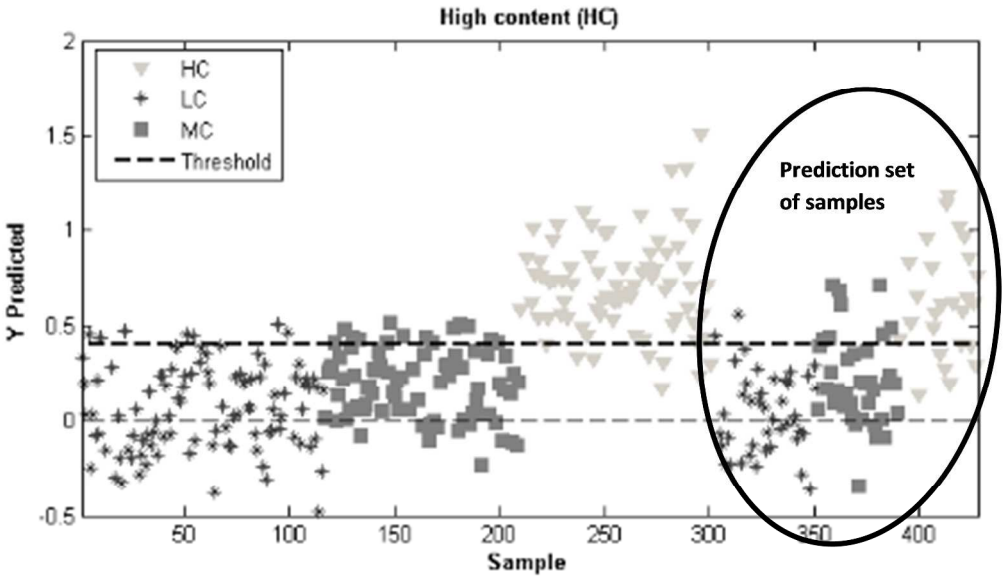
854 b)



855

856

857 c)



858

859

860 Fig.5

861

862