JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

ARTICLE

pubs.acs.org/JAFC

Rapid Estimation of Taro (*Colocasia esculenta*) Quality by Near-Infrared Reflectance Spectroscopy

Vincent Lebot,*[,†] Roger Malapa,[‡] and Marion Bourrieau[§]

[†]CIRAD, P.O. Box 946, Port-Vila, Republic of Vanuatu

[‡]CTRAV, P.O. Box 231, Luganville, Santo, Republic of Vanuatu

[§]ENSIA-SIARC, Ministère de l'Agriculture, Montpellier 34398 cédex 5, France

Supporting Information

ABSTRACT: The aim of the present study is to develop a methodology for the rapid estimation of taro (*Colocasia esculenta*) quality. Chemical analyses were conducted on 315 accessions for major constituents (starch, total sugars, cellulose, proteins, and minerals). NIRS calibration equations, developed on a calibration set composed of 243 accessions, showed high explained variances in cross-validation (r^2_{cv}) for starch (0.89), sugars (0.90), proteins (0.89), and minerals (0.90) but poor response for amylose (0.44) and cellulose (0.61). The predictions were tested on an independent set of 58 randomly selected accessions. The r^2_{pred} values for starch, sugars, proteins, and minerals were, respectively, of 0.76, 0.74, 0.85, and 0.85 with ratios of performance to deviation (RPD) of 3.41, 4.01, 3.78, and 3.64. New calibration equations developed on 303 accessions confirmed good RPD values for starch (3.30), sugars (4.13), proteins (3.61), and minerals (3.74). NIRS could be used to predict starch, sugars, proteins, and minerals contents in taro corms with reasonably high confidence.

KEYWORDS: near-infrared spectroscopy, NIRS, quality, starch, taro

INTRODUCTION

Taro (*Colocasia esculenta* (L.) Schott) is a traditional food of cultural importance throughout the World. Corms may be roasted, baked, boiled, steamed, or fried. They may be processed into fresh or fermented paste, canned corm portions, flour, beverage, chips, and flakes. Hundreds of different varieties exist, and depending on their chemical composition different products or dishes can be prepared.[1]

Internal color of raw taro corms ranges from white, yellow, orange, to dark purple and may include combinations of two or more colors. The texture of corms varies after cooking: some are acrid, and some contain high proportions of oxalate. Oxalate is toxic in relatively low concentrations, and the calcium salt may be present as needle crystals, which irritate the mouth membranes. There is, however, a nonidentified cofactor, most likely a protein, which contributes to acridity.[2] Chemically, there is also significant variation of the major nutritional constituents between varieties. These differences are well-known from consumers and traders, and the wrong choice of a variety can cause a serious constraint to the development of taro for food processing purposes. A preliminary survey of the physicochemical characteristics of 31 varieties planted and harvested the same day in the same plot revealed great variation for proteins, sugars, minerals, starch, amylose, and dry matter contents.[3] The large range of values found for a particular trait such as proteins, carotenoids, or anthocyanins shows that there is considerable potential for taro quality improvement through breeding.[4−6]

Taro breeding is progressing, and germplasm collections are being characterized with morpho-agronomic descriptors and molecular markers. Molecular studies have revealed the presence of two distinct genepools in Asia and the Pacific and the need to use germplasm from both genepools to broaden the base of breeding programs.[7] DNA markers allow an assessment of genetic distances, maximizing chances of getting significant variation in progenies,[8] and a first genetic map of taro has been developed.[9] Taro seeds can be generated in large quantities and allow the intense screening of thousands of highly variable hybrid seedlings. Visual tools are then used at an early stage to screen progenies for morphological characteristics.[10]

However, the chances of getting a high yielding hybrid with excellent eating quality are very low, and they become much lower when the selection procedure includes resistance against diseases. Corm yield and corm quality appear to be negatively correlated. Soft corms, with high water content, generally characterize high yielding hybrids. Unfortunately, the physicochemical characteristics determining the quality of the corms are very expensive and laborious to assess. Low-cost methods for evaluation of numerous accessions need to be developed. NIRS (Near-infrared reflectance spectroscopy) has been used to predict major constituents contents in maize,[11] rapeseed,[12] sorghum,[13] sugar beet,[14] malt,[15] wheat,[16] potato,[17] and tropical root and tuber crops.[18] These various studies indicate that NIRS could be a useful tool for taro breeding, selection, and quality control.

In the present study, we attempt to elucidate the potential relationships between taro corm quality and variation in major constituents composition of the varieties. We also investigate the potential of NIRS as an alternative method for predicting these

Received: June 9, 2011
Accepted: August 2, 2011
Revised: August 2, 2011
Published: August 02, 2011

major constituents. The results and their practical applications for improving the quality of the taro corm are discussed.

MATERIALS AND METHODS

Blind Tasting Panel of Taro Varieties. Taste evaluation of different taro varieties was conducted using the germplasm collection maintained by the VARTC (Vanuatu Agricultural Research and Training Centre) in Santo. Accessions planted 1 × 1 m in a common plot were harvested when mature 7−9 months later. Blind tasting panel was conducted on 340 taro varieties, and evaluated for their eating quality in 17 tasting sessions. The central part of the corm of each variety was cut into slices 3 cm thick and boiled in water. Cooking time depended on the variety, and readiness was assessed by pricking the slice flesh to appreciate its consistency; it varied from 19 to 50 min. This assessment is identical to the one conducted by regular taro consumers. It is done by frequently pricking the piece of corm with a fork to verify that it goes through but that there is resistance and firmness; if it goes through easily, it indicates that the piece is overcooked. During each tasting session, 20 varieties were cooked and evaluated by a panel of 10 individuals. The taste panel was composed of five women and five men of age ranging from 20 to 60 years, all indigenous Vanuatu citizens (Melanesians) and all regular taro consumers. Each variety slice was then cut into five pieces, resulting in 100 pieces for degustation in a session. Each taster received two plates at a time, with 10 pieces of boiled taro, each piece corresponding to a different variety.[19]

Tasters were separated in two groups with tasters 1−5 receiving varieties under a first set of numbers and tasters 6−10 receiving the same under a different set of numbers. Tasters were seating at tables laid out on a "U" shape in the room where tasters from the first group were seating next to a taster from the second group. Each variety was identified under a different number on each plate. These numbers did not correspond to the real variety number and changed from one taster to the other. The tasters could drink water between tests and taste the variety as many times as they wished. Each variety was tested five times by different tasters. The mean of five scores was recorded for each variety, and the standard variation of the mean was used to assess scores reliability. Varieties were classified according to their mean scores into four groups: poor quality (2), acceptable (3), good (4), or very good (5). Tasters were also requested to record when varieties were acrid (presence or absence).

Chemical Analyses. Overall, 315 accessions representing varieties from various geographical origins as well as hybrid lines were chemically analyzed. One full corm was peeled and cut. Approximately 0.5−1 kg of fresh weight, corresponding to the central part of the corm, was manually sliced into chips and oven-dried at 60 °C for 48 h. Dry matter samples were split into two subsamples: one subsample was used for chemical analysis, and the other was used for NIRS. Samples of 200 g were sent to Laboratoire d'Analyses Agricoles Teyssier, Bourdeaux, France, for chemical analyses. Samples of approximately 50 g of dried chips were milled into flour just after oven drying, and dried chips were ground in a stainless kitchen steel mill (SEB, France) prior to NIRS analysis in Vanuatu.

Major constituents (starch, amylose, sugars, cellulose, total N, and ash) were analyzed according to AFNOR (Association Française, the French standards association) and EEC methods.[20] Following NF (Norme Française) V 18-109 for dry matter (DM) determination, samples were dried again to remove residual moisture (measured as % of total dry weight), and the powder was analyzed on an oven-dried air basis. Moisture was therefore expressed as a measurement of the sample prior to drying. All measurements were then expressed in %DM, and the data were adjusted by the residual moisture following oven drying.

Starch was quantitated using Ewers protocol (NF ISO 10-520) corresponding to hydrolysis in HCl, filtration, and polarimetric measurement (specific rotation: 185.7°). Amylose measurement was done by colorimetric analysis using a standard iodine solution (NF EN ISO 6647-1) (on 200 accessions only for budget reasons). Total sugars were quantitated through the colorimetric method of Luff Schoorl (CEE 98\54\CE). Crude cellulose (total fibers) was measured by the Weende method (NF V 03-040), which corresponds to nonsoluble organic residue obtained by sulfuric acid and alkaline treatments. Total N content (considered as equivalent total proteins) was calculated using the Kjeldahl method (NF V 18-100). Estimation of total minerals content was obtained from ashes produced at 550 °C (NF V 18-101). All analyses were performed in duplicate with accepted mean coefficient of variation (SEL) of ±3% for starch, amylose, sugars, cellulose, and residual moisture and ±2% for proteins (equivalent N) and ashes (minerals).

NIRS Measurements and Data Pretreatment. Dry matter samples were milled into flour, and granules size was homogenized using four sieves with decreasing diameters until granules passed through the 106 μm sieve. An ASD LabSpecPro spectrophotometer from Analytical Spectral Devices Inc. (ASD Inc., Boulder, Colorado, CO) fitted with a "muglight" or High Intensity Source Probe (HISP) (ASD Inc.) was used for the measurement of all spectra over the wavelength range of 350−2500 nm. On average, 6 g of homogenized taro flour was placed in an individual cell and compacted with a tea spoon to eliminate air voids within the sample. Each spectrum was obtained by averaging three different cells (repetitions) per sample with 25 scans for each. A reference reading (baseline) was taken when starting a session and another every 30 min. All of the spectra were recorded in diffuse reflectance as $\log(1/R)$ with respect to a Labsphere's Spectralon material reflectance standard (Labsphere Inc., North Sutton, New Hampshire), which is a Lambertian reflective PTFE (thermoplastic resin) with high overall reflectance. For each sample, corresponding to individual accession, three subsamples were scanned 25 times each and then averaged. The resulting averaged spectrum was recorded for the accession. Overall, 303 spectra were recorded and converted to absorbance using the Indico software (ASD Inc.). To assess the performance of the calibration, samples were separated into two sets: the calibration and the prediction sets. The prediction set was created by randomly selecting 58 accession numbers (approximately 20% of total 303 acc.) and the calibration set contained 245 samples (for amylose, the calibration set included 160 acc. and the validation set included 40 acc.).

Data Analysis. Major constituents chemical data were subjected to multivariate analysis using XLSTAT (version 6.02, 2009). Multivariate analysis (Principal Component) of the spectra was conducted with GRAMS/AI (version 8.0). The spectra and reference data were mathematically modeled using PLSPlus/IQ spectroscopy software (Thermo Electron Corp., OH). Using the values obtained with chemical analyses as the analyte value, a separate calibration was made for each of the six major constituents. Calibration of residual moisture was not attempted because spectra were recorded in Vanuatu, just after oven drying the samples, while residual moisture was measured in France on hygroscopic dry raw material. Partial least-squares (PLS) regression technique was used to develop a predictive model of the near-infrared part of the spectra (1000−2500 nm). The optimum number of PLS factors used for prediction was determined by full cross-validation and PRESS (prediction residual error sum of squares). Additionally, light scattering effects due to particle size differences were corrected by multiplicative scatter correction (MSC). The data were mean-centered, and the average spectrum was calculated from all of the calibration spectra and then subtracted from every calibration spectrum.

As part of the model process, a principal component analysis (PCA) was used to check for gross spectroscopic outliers. The Mahalanobis distance of each spectrum to the mean spectrum of the group was calculated, and the removal of outliers was based on distance $H > 3$ from the average spectrum of the file. Spectra and concentration outliers were removed, and PLS was run again until the highest r^2_{cv} (determination coefficient for cross validation) corresponding to the

Table 1. Results of Blind Tasting Panel Conducted by 10 Individuals on 340 Taro Accessions

quality	*n* var.[a]	mean score	mean std	mean CV %	*n* acrid varieties	%
very good (*n* = 26)	15	4.8	0.45	9.3	0	0.0
	11	4.6	0.55	11.9	0	0.0
good (*n* = 69)	13	4.4	0.55	12.4	0	0.0
	27	4.2	0.45	10.6	0	0.0
	29	4.0	0.00	0.0	2	6.9
acceptable (*n* = 181)	33	3.8	0.84	22.0	2	6.1
	43	3.6	0.89	24.8	6	14.0
	36	3.4	0.89	26.3	10	2.8
	34	3.2	0.84	26.1	6	17.6
	35	3.0	0.71	23.6	9	24.7
poor (*n* = 64)	24	2.8	0.84	29.9	11	45.8
	19	2.6	0.55	21.1	10	5.3
	18	2.2	0.45	20.3	5	27.8
	3	2.0	0.00	0.0	3	100.0
total	340		0.83	24.2	64	18.8

[a] Number of varieties with the same mean score.

smallest SECV (standard error of cross validation) was obtained. At that point, factor loadings were used to determine which wavelengths were important to correlate with concentrations to narrow down the spectroscopic region. The loading plots showed which wavelengths were important to correlate with concentrations. The loading weights showed how much each wavelength point contributed to explaining the response along each model component. For starch and amylose, the regions used were 1200−2200 nm, while for sugars, proteins, and minerals, the region was 1200−2400 nm, but for cellulose, the region was 1400−2000 nm. The PLS analysis was then conducted again on these new regions to obtain for each constituent equations with higher explanation of the total variability in the calibration values without increasing the number of PLS factors used.

Statistical parameters used to evaluate models performances included the standard error of calibration (SEC), the determination coefficient for cross validation (r^2_{cv}), the standard error of cross-validation (SECV), the determination coefficient for prediction (r^2_{pred}), and the standard error of prediction (SEP). SEC and SEP were calculated using an Excel spreadsheet by squaring the differences of the actual minus the predicted concentrations for each sample in the calibration (SEC) and test (SEP) sets. These values were then summed, and the sum was divided by the number of samples (*n*). The square root of this value was used for SEC and SEP. SEC describes the calibration set (243 acc.), and SEP describes the test set composed of 58 samples not included in the calibration set. The ratio of performance to deviation (RPD = SD/SECV) was also used to evaluate performances of the models (with SD as the standard deviation of the original chemical data in the calibration set).[21] Finally, new calibrations were computed on the total number of samples (243 + 58 = 303 acc.).

RESULTS

Out of 340 varieties tasted, 26 were found to be "very good", 69 "good", 181 "acceptable", and 64 were thought to be of "poor" quality. For each of these four groups of varieties, the mean standard variation and CV% of the mean score give an estimation of the reliability of the appreciations within each group (Table 1). The best 15 varieties were all rated 4−5−5−5−5. When three scores out of five were equal to "5", the variety was still considered as "very good". Overall, 26 varieties were rated "very good", and the low CVs (9.3% and 11.9%) of their mean scores indicate that these appreciations seem fairly reliable. In other words, most tasters agreed on what characterizes a "very good" taro variety. The "very good" varieties were subsequently selected as recommended varieties and propagated for distribution to farmers who also confirmed later their very good quality, strengthening the consensus. "Good" varieties also appear to be properly rated with an average CV% ranging from 12.4 to nil. However, ratings such as "acceptable" and "poor" quality appeared much more difficult to appreciate during the blind tasting panel exercise. Not less than 181 varieties were rated as "acceptable". This might be an indication of the difficulty of rating varieties with average quality and is confirmed by the high mean CVs (22−26.3%) of their mean scores.

Acridity was not rated evenly, and some consumers, especially women, were more susceptible than others. The best 66 varieties (mean scores from 4.2 to 4.8) were found to be free of acridity by all tasters. Out of 340 varieties consumed after boiling, 64 varieties were rated acrid, but these belong to different quality groups ranging from poor to good. Some of these varieties might need a longer cooking time and/or a different preparation to be palatable (Table 1).

Overall, 315 taro accessions were analyzed for the chemical variation of their major constituents. They originated from five SE Asian countries and 12 different islands of Vanuatu. Great care was taken to select morphologically distinct genotypes within each country of origin. Varieties from Asia and the Pacific were included in the sample as well as hybrids between the two genepools. "Poor" quality varieties as well as "very good" and recommended varieties were also included in the sample.

Results of the chemical analyses conducted on 315 accessions are presented in Table 2. Significant variation was observed for dry matter content (DM) and major constituents. The least variable constituent was starch (CV% = 6.73), and the most variable was total sugars (CV% = 79.51). Two subsamples assembling 26 "very good" and 38 "poor" varieties are also presented in Table 2. There are remarkable differences between their mean values with higher DM (40.66%) and starch (83.67%) contents and lower amylose (17.31%), sugars (2.09%), cellulose (2.48%), proteins (4.12%), and minerals (3.58%) values in "very good" varieties. Correlation coefficients calculated between major constituents indicate that starch content is positively correlated with %DM but negatively correlated with sugars, cellulose, proteins, and minerals contents (Table 3).

Principal component analysis conducted on the data matrix (315 acc. × 5 major constituents) reveals the respective contribution of the five variables to the projection, with axes 1 and 2 totalizing 69% of the total variance (Figure 1). Varieties rated "very good" present all very similar chemical compositions with starch content above 80% (Figure 2). Also, overall, varieties rated as "very good" seem to have a low amylose/starch ratio, but there is no clear-cut picture, and a few of these varieties also presented a high amylose content.

The comparison of the NIRS spectra (Figure 3) and the chemical values allowed the establishment of equations of calibration for the prediction of starch, sugars, proteins (equivalent N), and minerals. The results are presented in Table 4. For starch, the SECV (1.56%) and SEC (1.56%) values are identical, indicating robust fitting. The SEP (2.14%) is not too distant, and the r^2_{pred} of 0.76 indicates an acceptable estimation of the equation accuracy on the validation samples. Deviations of single samples are visualized in a scatter plot between measured and predicted starch

Table 2. Major Constituents Analyzed in 315 Taro Accessions (as % of DM Adjusted to Moisture Content)[a]

varieties		DM	starch	amylose	ratio a/s	sugars	cellulose	proteins	minerals
	min	19.51	55.88	10.71	0.13	0.21	1.40	2.13	1.47
	max	53.74	89.46	49.30	0.78	21.80	7.30	14.79	8.85
all	mean	35.81	78.88	19.64	0.25	3.80	3.07	4.61	4.22
$n = 315$	std error	7.01	5.31	5.19	0.08	3.02	0.90	1.66	1.09
	CV%	*19.56*	*6.73*	*26.43*	*32.51*	*79.51*	*29.21*	*36.06*	*25.81*
very good	mean	40.66	83.67	17.31	0.21	2.09	2.48	4.12	3.58
$n = 26$	std error	6.69	3.81	3.73	0.05	1.54	0.60	1.32	0.87
	CV%	*16.46*	*4.56*	*21.54*	*22.79*	*73.65*	*24.32*	*31.92*	*24.22*
poor	mean	33.59	68.66	25.54	0.37	8.61	4.03	5.23	5.22
$n = 38$	std error	5.33	3.83	8.32	0.13	4.39	0.95	1.77	1.53
	CV%	*15.87*	*5.28*	*32.58*	*35.33*	*51.02*	*23.45*	*33.90*	*29.28*
t test[b]		*7.78*	*4.57*	*4.58*	*3.58*	*3.47*	*1.59*	*2.47*	*2.59*

[a] Comparison between 26 "very good" and 38 "poor" accessions. [b] Student t value = 2.02 at $p = 0.005$.

Table 3. Correlation Coefficients Between Major Constituents for 315 Accessions (Pearson ($n - 1$))

variables	DM	starch	amylose[a]	sugars	cellulose	proteins
starch	+0.372[b]					
amylose	+0.320[b]	+0.542[b]				
sugars	−0.212	−0.721[b]	−0.207			
cellulose	−0.283[b]	−0.504[b]	−0.292[b]	+0.284[b]		
proteins	−0.179	−0.319[b]	−0.218	+0.107	+0.203	
minerals	−0.400[b]	−0.542[b]	−0.558[b]	+0.001	+0.407[b]	+0.181

[a] On 200 varieties. [b] r value at 1% = 0.254.

values of the 58 acc. in the test set (Figure 4A). In terms of predictive performance, the equations for starch could be considered as good with RPD parameters above 3. Some authors claim that a RPD value of at least 3 is necessary for efficient NIR reflectance predictions.[21]

Amylose statistical parameters revealed the poor performances of the equations with r^2_{cv}, r^2_{pred}, and RPD values of 0.44, 0.15, and 1.67, respectively. The total sugar model presents similar SECV (0.73) and SEC (0.75) values, but the SEP is much higher (1.50), although the r^2_{pred} is of 0.74. When SECV and SEP values differ significantly, this could be an indication that too many samples (HT > 3 = 21) were removed during the modeling process. The RPD value of 4.01 indicates, however, a good predictive potential for this equation. Deviations of single samples are visualized in a scatter plot between measured and predicted sugars values (Figure 4B). Cellulose could not be satisfactorily predicted, and a poor r^2_{cv} (0.61) was obtained, with very low r^2_{pred} (0.37) and RPD (2.05). Proteins (measured as total N equivalent) produce similar SECV, SEC, and SEP values (respectively, 0.46, 0.57, and 0.57) and a high r^2_{pred} of 0.85, indicating good and robust prediction with 85% of confidence. The RPD value above 3.5 confirms a very good potential of prediction for this model. Deviations of single samples are visualized in a scatter plot between measured and predicted proteins values (Figure 4C). Minerals are known to have a poor relationship with NIRS, but they presented similar SECV, SEC, and SEP values (respectively, 0.29, 0.35, 0.44) and could be predicted with 85% of confidence with a good RPD value of 3.64. Deviations of single samples are visualized in a scatter plot between measured and predicted minerals values (Figure 4D).

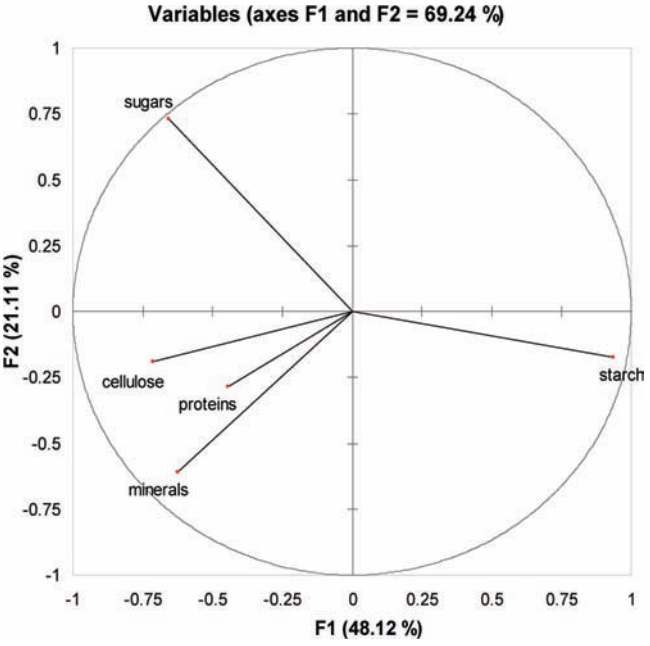

Figure 1. PCA of 315 accessions × 5 variables (starch, sugars, cellulose, proteins, minerals).

The r^2_{pred} values of starch, sugars, proteins, and minerals are high enough to allow good estimates of their contents. RPD between 3.41 and 4.01 for the starch and sugars models also allow good quantitative predictions to be made. Values above 2.5 for proteins are considered to be good models,[20] and the value here is above 3.5. The number of terms is also relatively low if we consider a general recommendation of 1 factor for every 10 samples in a model (Table 4).

Calibrations were modeled again on 303 samples by adding the calibration and validation sets. These new models (Table 5) could not be validated on an independent test set. However, their RPD values were high, and their SECV and SEC values were close enough to suggest fair and robust fitting for starch (respectively, 1.59 and 1.91), sugars (0.73 and 0.85), proteins (0.46 and 0.58), and minerals (0.29 and 0.38). The new model for proteins presented a RPD value above 3.5, and such a value indicates a very good predictive model.[20] RPD values for starch

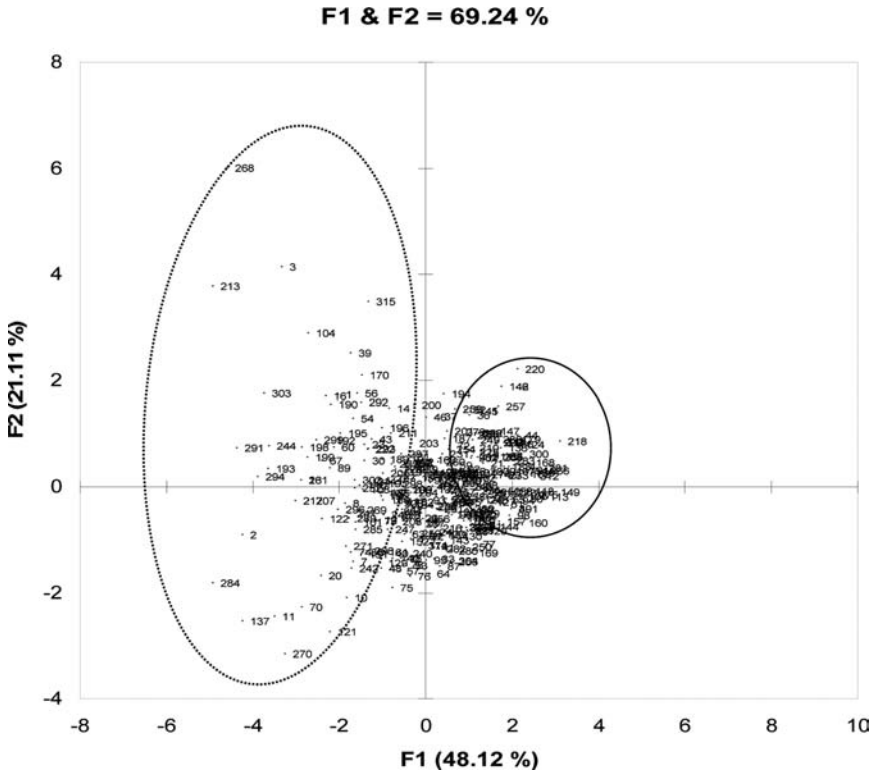

Figure 2. PCA of 315 accessions × 5 variables (starch, sugars, cellulose, proteins, minerals). Very good and good accessions with more than 80% starch are in the small circle. Poor quality accessions, with less than 70% starch, are in the dotted line circle.

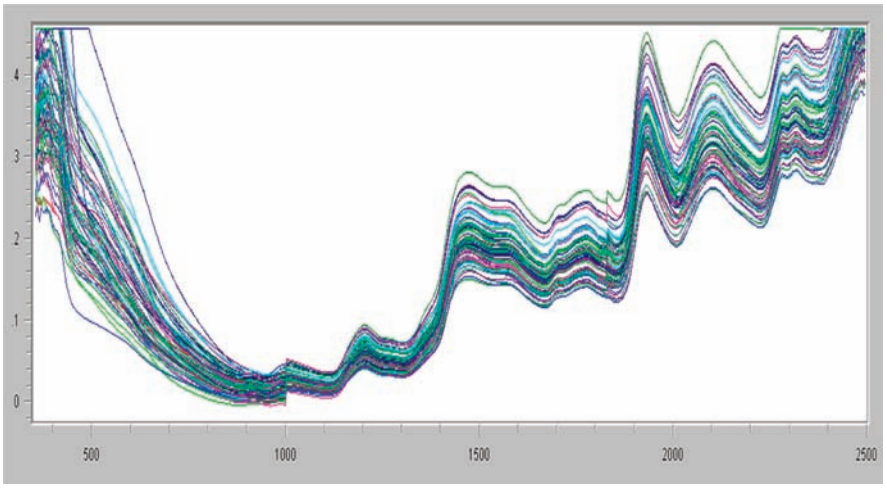

Figure 3. Infrared spectra of 303 taro varieties: *x*-axis, wavelengths; *y*-axis, absorbance. 350−750 nm is the visible range (variability due to colors of samples), and 750−2500 nm is the NIR range. The water peaks are at 1450 and 1940 nm.

and sugars were also very high, respectively, of 3.30 and 4.13. Again, the model for minerals is also very good with an RPD value of 3.74. Values for cellulose were slightly improved, but the low r^2_{cv} and RPD (0.56 and 2.11) do not allow retaining this model for prediction. Models for starch, sugars, proteins, and minerals present good potential but will need to be further tested on independent samples.

Fresh taro corm quality is related to a certain chemical composition, and different varieties are processed and cooked into various preparations throughout the World.[22] In Vanuatu, it appears that "very good" varieties prepared after boiling have high DM (>35%) and starch contents (>80% DM). Consumers characterize these varieties as being "elastic", and this appreciation refers to the somewhat firm texture felt when chewing a piece of the boiled corm. The present analysis conducted on a large sample confirms the results obtained by a previous preliminary study on a much smaller sample of 31 varieties.[3] Blind panel tasters seem to reach a consensus when appreciating these varieties

The models developed in the present study show good accuracy, but it remains to be seen whether larger sample sets will improve them to enable more precise prediction. When comparing

Table 4. Statistical Parameters of the Calibration and Validation Sets

	calibration (n = 245)									validation (n = 58)	
constituents	mean % DM	SD	SEL ±	HT H > 3	PLS terms	r^2_{cv}	SECV	SEC	RPD	r^2_{pred}	SEP
starch	79.22	5.32	2.38	21	12	0.89	1.56	1.56	3.41	0.76	2.14
amylose[a]	19.58	4.71	0.59	15	3	0.44	2.88	3.66	1.67	0.15	5.76
sugars	3.74	2.94	0.11	19	12	0.90	0.73	0.75	4.01	0.74	1.50
cellulose	3.03	0.85	0.09	20	8	0.61	0.41	0.41	2.05	0.37	0.83
proteins	4.64	1.73	0.09	21	9	0.89	0.46	0.57	3.78	0.85	0.57
minerals	4.09	1.06	0.08	16	8	0.90	0.29	0.35	3.64	0.85	0.44

[a] Calibration set of 160 acc. and validation set of 40 acc.

the performance of the new calibration models (with n = 303), the high r^2_{cv} and RPD values were confirmed. Determination coefficients (r^2_{pred}) generally improve as the working range increases. Consequently, if more range is added in the same model, then it could improve coefficient values. Additionally, when different samples are added, a larger spectroscopic diversity is described, and, therefore, some samples might actually be better spectrally described as the number of samples in the calibration set increases. However, determination coefficients for the prediction set (r^2_{pred}) cannot reflect the whole situation because the range of the 58 accessions values affects the coefficient values. These values change according to the type and number of validation samples, and it is necessary to consider the long-term effects. Errors of prediction values have been shown to have uncertainties, and it is therefore recommended to be cautious while reporting prediction errors because they may change according to the validation set used.[23] SEP is, therefore, a better overall indicator. Also, the selection of the calibration set was done by removing at random 58 samples for the test set. The calibration set has not covered the whole variation of the data set (303 acc.). A better sample selection might be helpful by selecting, for example, on constituent concentration rather than a random selection of numbers. Further work should concentrate on validating the results over different years.

The models for starch, sugars, proteins, and minerals present potential for improvement if more samples could be added. The lowest RPD was obtained for amylose, but it is not the first time that this constituent is found difficult to analyze by NIRS.[13] The protein content calibration is particularly interesting as it can be further improved. Proteins content is usually estimated by multiplying the total N content by a standard conversion factor of 6.25. However, the nitrogen to protein ratio does vary according to the species considered and change with amino acid content and mineral nitrogen and nonprotein nitrogen. For the present study, we decided to present our results measured as total N as proteins. In the future, it would be of interest to improve the calibration models on the real protein content of taro, which vary according to amino acids. Once known, the values obtained by the Kjeldahl method could be converted into more accurate measurements for NIRS calibrations on taro.

In taro breeding programs, mass selection results in the rapid accumulation of suitable genes but has to be complemented with efficient screening techniques of hundreds of hybrids generated in controlled crosses. Correlation coefficients between major constituents indicate that breeding for increased DM and starch contents will reduce sugars, proteins, and minerals. These correlations do not present practical problems as "poor" quality varieties have been shown to present low DM and starch and high sugars, cellulose, proteins, and minerals. Obviously, NIRS could assist taro breeders in their choice and selection of the best genotypes, based on the chemical composition requested by consumers by predicting simultaneously starch, sugars, proteins, and minerals on a single sample. As starch is significantly negatively correlated with the other three major constituents, the simultaneous prediction of all four constituents allows for rapid estimation of the variety chemotype and therefore its quality.

Acridity is fairly difficult to score, but all varieties rated "very good" by tasters are not acrid. Previous studies have shown that raphides are not the only cause of acridity but that they could play a major role in penetration and carrying the acrid factor. The irritant is thought to be a proteinase or histamine, but more research is needed to elucidate the complexity of this trait.[2] However, it is reasonable to consider that from a genetic point of view, acridity is a wild trait, comparable to cyanogens in cassava or glycoalkaloids in potato, and used by the taro plant as repellent to herbivores. Traditional varieties have reduced levels of acridity, and it is possible to assume that this trait is negatively correlated with improved quality, in other words, with higher DM and starch contents. If this could be confirmed, it would be possible to select for non acrid type by selecting for high starch and low sugars, cellulose, proteins, and minerals. This question is important for breeding programs and deserves further research.

Variation in taro mucilage was observed but was not studied, although mucilage is also an important factor in the determination of taro quality. The main components of this mucilage have been shown to be galactose, arabinose, and an arabinogalactan-protein, all varying in great proportion according to the variety.[24] More studies are needed to elucidate the respective roles of these compounds for eating quality. Also, taro corms have been shown to present variable carotenoids, anthocyanins, and phenolic compounds, mostly responsible for different corm flesh colors. Regarding carotenoids, taro varieties present all-*trans*-β-carotene, small amounts of lutein, and two unknown substances.[5] Flavonols content is very high with not less than nine components including catechin and epicatechin, indicating an interesting source of healthy components in taro corms.[6] The varying colors, as shown by the visible range of the spectra (350−750 nm) (Figure 3), increased in intensity after boiling, but these colors were not scored by tasters. The relationships between secondary metabolites and quality remain to be demonstrated, but it is quite clear that consumers often prefer colorful taro corms rather than white flesh.

Some preliminary basic information now exists on the relationship between chemotypes and taste, but much more is needed to understand the chemical variation between varieties. Taro is a

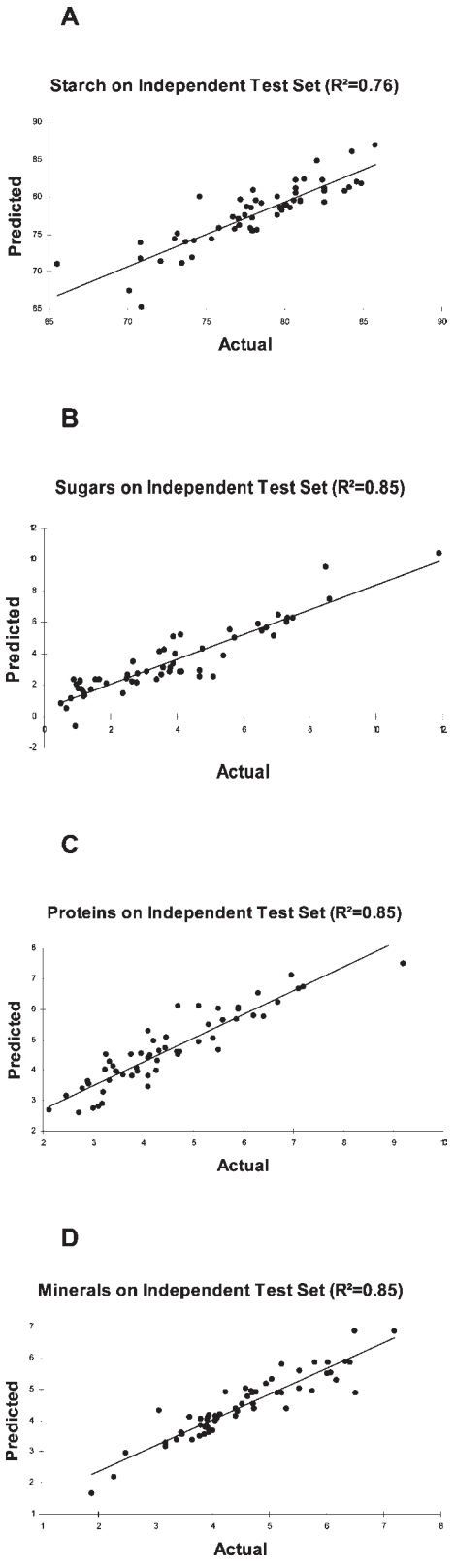

Figure 4. Validation results of the NIRS models for starch (A), sugars (B), proteins (C), and minerals (D).

diploid species, but nothing is known on the segregation of these major constituents. The problem is rather complex as these constituents are most likely controlled by many sets of different genes, and molecular tools can hardly be used for markers assisted selection and conventional selection of parents for breeding or selection. Other constituents such as carotenoids or flavonols should be investigated as well, but they present very low concentration in the DM. NIRS offer interesting perspectives for spectra assisted selection.

Growing urbanization implies that taro corm quality is controlled and that new products are developed. Increasing numbers of urban dwellers have the resources to buy taro corms presented in convenient supermarket-style packages. If processed into snack foods, vacuum packed products, and special flours, taro would be in demand in growing Asian or African cities. In addition, the gelling properties and general sensory characteristics of taro open up the possibility of a completely new range of products. This potential is enhanced by globalization, and consumers in Western countries are eager to try novel foods, especially those from the tropics. In many developing countries, there is now a strong desire to develop taro production by breeding for improved corm quality.

Table 5. New Calibration with 303 Accessions

constituents	mean % DM	SD	SEL ±	HT > 3	PLS terms	r^2_{cv}	SECV	SEC	RPD
starch	78.88	5.27	2.36	27	11	0.89	1.59	1.91	3.30
sugars	3.85	3.03	0.12	24	12	0.90	0.73	0.85	4.13
cellulose	3.07	0.89	0.09	28	9	0.56	0.42	0.54	2.11
proteins	4.61	1.68	0.09	28	11	0.89	0.46	0.58	3.61
minerals	4.21	1.1	0.08	25	11	0.90	0.29	0.38	3.74

■ ASSOCIATED CONTENT

Supporting Information. Table 6: Selected accessions analyzed for major constituents. Table 7: Chemical analysis of 315 accessions of taro (*Colocasia esculenta*). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: +678 25947. Fax: +678 25947. E-mail: vincent.lebot@cirad.fr.

Funding Sources

This research was financially supported by the Food Security Thematic Programme of the EU (grant no DCI-FOOD/2010/230-267) and the International Network for Edible Aroids (INEA, www.EdibleAroids.org).

■ ACKNOWLEDGMENT

We thank S. Stephens for laboratory assistance in preparing samples and taking spectra.

■ ABBREVIATIONS USED

SEC, standard error of calibration; SECV, the standard error of cross-validation; SEP, the standard error of prediction; SEL, the standard error of the laboratory analysis; r^2_{cv}, the determination coefficient for cross validation; r^2_{pred}, the determination coefficient for prediction; RPD, the ratio of performance to deviation; *H*,

Mahalanobis distance limit; HT, number of outliers removed; %DM, percentage of dry matter; CV%, coefficient of variation

REFERENCES

(1) Rao, R., Matthews, P. J., Eyzaguirre, P. B., Hunter, D., Eds. *The Global Diversity of Taro: Ethnobotany and Conservation*; Bioversity International: Rome, Italy, 2010; 202 p.

(2) Paull, R.; Tang, C. S.; Gross, K.; Uruu, G. The nature of the taro acridity factor. *Postharvest Biol. Technol.* **1999**, *16*, 71–78.

(3) Lebot, V.; Gunua, T.; Pardales, J. R.; Prana, M. S.; Thongjiem, M.; Viet, N. V.; Yap, T. C. Characterisation of taro (*Colocasia esculenta* (L.) Schott) genetic resources in Southeast Asia and Oceania. *Genet. Resour. Crop Evol.* **2004**, *51*, 381–392.

(4) Champagne, A.; Legendre, L.; Malapa, R.; Lebot, V. Chemotype profiling to guide breeders and explore traditional selection of root crops in Vanuatu, South Pacific. *J. Agric. Food Chem.* **2009**, *57*, 10363–10370.

(5) Champagne, A.; Bernillon, S.; Moing, A.; Rolin, D.; Legendre, L.; Lebot, V. Carotenoid profiling of tropical rootcrop chemotypes from Vanuatu, South Pacific. *J. Food Compos. Anal.* **2010**, *23*, 763–771.

(6) Champagne, A.; Bernillon, S.; Moing, A.; Rolin, D.; Legendre, L.; Lebot, V. Diversity of anthocyanins and other phenolic compounds among tropical root crops from Vanuatu, South Pacific. *J. Food Compos. Anal.* **2011**, *24*, 315–325.

(7) Kreike, C. M.; van Eck, H. J.; Lebot, V. Genetic diversity of taro, *Colocasia esculenta* (L.) Schott, in Southeast Asia and the Pacific. *Theor. Appl. Genet.* **2004**, *109*, 761–768.

(8) Mace, E.; Godwin, I. Development and characterisation of polymorphic microsatellite markers in taro (*Colocasia esculenta*). *Genome* **2002**, *45*, 823–832.

(9) Quero-García, J.; Courtois, B.; Ivancic, A.; Letourmy, P.; Risterucci, A. M.; Noyer, J. L.; Feldmann, P.; Lebot, V. First genetic maps and QTL studies of yield traits of taro (*Colocasia esculenta* (L.) Schott). *Euphytica* **2006**, *15*, 187–199.

(10) Ivancic, A.; Quero Garcia, J.; Lebot, V. Development of visual tools for selecting qualitative corm characteristics of taro (*Colocasia esculenta* (L.) Schott). *Aust. J. Agric. Res.* **2003**, *54*, 581–588.

(11) Berardo, N. G.; Mazzinelli, P.; Valoti, P.; Laganà Redaelli, R. Characterization of maize germplasm for the chemical composition of the grain. *J. Agric. Food Chem.* **2009**, *57*, 2378–2384.

(12) Amar, S.; Becker, H. C.; Möllers, C. Genetic variation in phytosterol content in winter rapeseed (*Brassica napus* L.) and development of NIRS calibration equations. *Plant Breed.* **2009**, *128*, 78–83.

(13) De Alencar Figueiredo, L. F.; Davrieux, F.; Fliedel, G.; Rami, J. F.; Chantereau, J.; Deu, M.; Courtois, B.; Mestres, C. Development of NIRS equations for food grain quality traits through exploitation of a core collection of cultivated sorghum. *J. Agric. Food Chem.* **2006**, *54*, 8501–8509.

(14) Roggo, Y.; Duponchel, L.; Huvenne, J. P. Quality evaluation of sugar beet (*Beta vulgaris*) by near-infrared spectroscopy. *J. Agric. Food Chem.* **2004**, *52*, 1055–1061.

(15) Marte, L.; Belloni, P.; Genorini, E.; Sileoni, V.; Perretti, G.; Montanari, L.; Marconi., O. Near-infrared reflectance models for the rapid prediction of quality of brewing raw materials. *J. Agric. Food Chem.* **2009**, *57*, 326–333.

(16) Rakszegi, M.; Boros, D.; Kuti, C.; Làang, L.; Bedō, Z.; Shewry, P. R. Composition and end-use quality of 150 wheat lines selected for the HEALTHGRAIN diversity screen. *J. Agric. Food Chem.* **2008**, *56*, 9750–9757.

(17) Haase, N. U. Rapid estimation of potato tuber quality by near-infrared spectroscopy. *Starch/Stärke* **2006**, *58*, 268–273.

(18) Lebot, V.; Champagne, A.; Malapa, R.; Shiley, D. NIR determination of major constituents in tropical root and tuber crop flours. *J. Agric. Food Chem.* **2009**, *57*, 10539–10547.

(19) Bourrieau, M. Valorisation des racines et tubercules tropicaux pour l'alimentation humaine en Océanie. M.Sc. Thesis in Food Science, ENSIA-SIARC, Montpellier, France, 2000; 110 p.

(20) AFNOR. http://www.boutique.afnor.org/NEL1AccueilNormeEnLigne.aspx; Date last accessed: August 1st, 2011.

(21) Williams, P. *Near-Infrared Technology: Getting the Best out of Light. A Short Course in the Practical Implementation of Near-Infrared Spectroscopy for the User*; PDK Projects, Inc.: Nanaimo, Canada, 2003.

(22) Matthews, P. J. An Introduction to the History of Taro as a Food. In *The Global Diversity of Taro: Ethnobotany and Conservation*; Rao, et al., Eds.; Bioversity International: Rome, Italy, 2010; pp 6−28, 202 p.

(23) Sileoni, V.; van den Berg, F.; Marconi, O.; Perretti, G.; Fantozzi, P. Internal and external validation strategies for the evaluation of long-term effects in NIR calibration models. *J. Agric. Food Chem.* **2011**, *59*, 1541–1547.

(24) Jiang, G.; Ramsden, L. Characterisation and yield of the arabinogalactan-protein mucilage of taro corms. *J. Sci. Food Agric.* **1999**, *79*, 671–674.