Amylose Quantification in Fresh Grated Cassava Clones using NIRS

High-Throughput Phenotyping Protocols (HTPP), WP3

Kampala, Uganda, 03/10/2022

Ephraim NUWAMANYA, National Crops Resources Research Institute (NaCRRI), Namulonge, Uganda

Babirye Fatumah NAMAKULA, NaCRRI, Namulonge, Uganda

Fabrice DAVRIEUX, Centre de coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), Saint Pierre, La Réunion, France

Michael KANAABI, NaCRRI, Namulonge, Uganda

Enoch WEMBABAZI, NaCRRI, Namulonge, Uganda

Ivan LYATUMI, NaCRRI, Namulonge, Uganda

Robert KAWUKI, NaCRRI, Namulonge, Uganda

This report has been written in the framework of RTBfoods project.

To be cited as:

Ephraim NUWAMANYA, Babirye Fatumah NAMAKULA, Fabrice DAVRIEUX, Michael KANAABI, Enoch WEMBABAZI, Ivan LYATUMI, Robert KAWUKI, (2023). *Amylose Quantification in Fresh Grated Cassava Clones using NIRS; High-Throughput Phenotyping Protocols (HTPP), WP3.* Kampala, Uganda: RTBfoods Calibration Report, 11 p. https://doi.org/10.18167/agritrop/00730

Ethics: The activities, which led to the production of this document, were assessed and approved by the CIRAD Ethics Committee (H2020 ethics self-assessment procedure). When relevant, samples were prepared according to good hygiene and manufacturing practices. When external participants were involved in an activity, they were priorly informed about the objective of the activity and explained that their participation was entirely voluntary, that they could stop the interview at any point and that their responses would be anonymous and securely stored by the research team for research purposes. Written consent (signature) was systematically sought from sensory panelists and from consumers participating in activities.

Acknowledgments: This work was supported by the RTBfoods project https://rtbfoods.cirad.fr, through a grant OPP1178942: Breeding RTB products for end user preferences (RTBfoods), to the French Agricultural Research Centre for International Development (CIRAD), Montpellier, France, by the Bill & Melinda Gates Foundation (BMGF).

This document has been reviewed by:	
Babirye Fatumah NAMAKULA (NaCCRI) Fabrice DAVRIEUX (CIRAD)	28/09/2022 19/11/2022
Final validation by:	
Fabrice DAVRIEUX (CIRAD)	08/12/2022

CONTENTS

Table of Contents

ABSTRACT

Context: This scientific report concerns data analysis of two matrices of measured data on fresh grated cassava 1) physico-chemical data and 2) spectral data. The data were collected on fresh cassava in NaCRRI, Uganda.

Place: Uganda, Réunion

Date: 19/11/2022

Authors: Fabrice DAVRIEUX, Fatumah Namakula BABIRYE and Ephraim NUWAMANYA

Content:

The analyses concern 94 cassava genotypes harvested in 2022 in Namulonge, Uganda. Genotypes came from NaCRRI Cassava breeding program. For this study 247 cassava roots were analysed; 3 uniformly sized non-necrotic roots per clone were sampled. Amylose quantification was achieved in NaCRRI physico-chemical laboratory in Namulonge. Near infrared spectra were scanned in NIR laboratory of NaCRRI in Namulonge. The protocol measurement follows the SOP protocol described for fresh grated cassava: https://doi.org/10.18167/agritrop/00669.

Amylose content, expressed as % of starch, vary between 15,2% and 33%. On the basis of the 247 samples 2 data sets were constituted: one training set (n = 173) and one test set (n =74) for this 70% of the samples were picked randomly for the training set while the remaining 30% were kept as test set.

Amylose content was calibrated using Modified Partial Least Squares Regression, for the spectral range: 400 nm - 2500nm, that is to say Visible and NIR regions. Calibrations were done on non-pretreated spectra and pretreated spectra using different pre-treatments. The best model was obtained with no pre-treatment applied to the spectra, the R^2 was 0,937 with an SECV equal to 0,801%.

This model was applied to predict samples from the test set, the standard error of prediction (SEP) is equal to SEP = 1,023% and the R^2 for prediction is 0,915. The ratio performance to deviation RPD is equal to 2,947.

The developed NIRs model for quantification of amylose content of cassava root presents an accuracy good enough to enable cassava selection based on amylose content expressed as % of dried starch. The error of the model is SEP = 1,02% which means that a predicted value could be defined with a confidence interval of +/- 2,04 % associated with 95 % of confidence. Furthermore, the database refers to 94 different genotypes representative of the variability of the starch properties, in particular for amylose content, thus the model is robust enough for amylose quantification in fresh cassava. The procedure can be applied in cassava breeding programs based on amylose content.

Keywords : cassava fresh grated, NIRS, amylose, starch

1 DATA

1.1 Material

Cassava root samples used for analysis were planted and harvested at Namulonge, 247 samples were analyzed corresponding to 94 genotypes. Three uniformly sized non-necrotic roots per clone were sampled. Amylose quantification was achieved in NaCRRI physico-chemical laboratory in Namulonge. Near infrared spectra were scanned in NIR laboratory of NaCRRI in Namulonge.

1.2 Amylose quantification

The amylose content of cassava starch was determined using a spectrophotometric procedure based on iodine staining with slight modifications by adjusting the volume of analysis to a total working volume of 50mL (Juliano, 1971). The amylose standard curve was prepared using potato amylose (Sigma Chemicals). Repeated analyses were undertaken where the result standard deviation was higher than 5% to ensure consistency. The results were presented as percentage amylose on dry weight basis.

1.3 Near Infrared measurements

The NIR spectra acquisition was done on fresh grated cassava roots using a DS2500 (FOSS) spectrometer. Three roots per clone were sampled and scanned, one spectrum per root was performed. The protocol measurement follows the SOP protocol described for fresh grated cassava: https://doi.org/10.18167/agritrop/00669

2 RESULTS

2.1 Amylose Content

Table 1: Amylose content descriptive statistics

Statistic	No	Min	Max	Mean	SD
Amylose (% of starch)	247	15.17	32.96	24.52	3.37

Amylose content vary between 15,2% and 33% (table 1) which correspond to a wide range with a relative high dispersion with a standard deviation of 3.4% (fig. 1). Thus, the dataset was diverse enough to be representative of amylose content within cassava genotypes in selection. And the range and distribution of the amylose values are suitable for modelisation using NIR fingreprints.

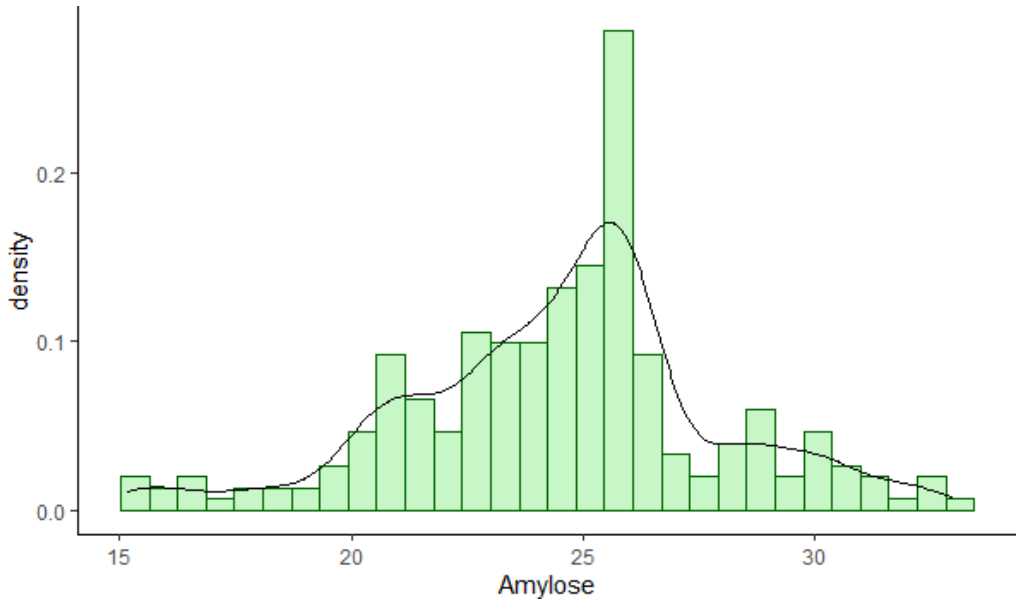

Figure 1: Amylose content (% of starch) distribution.

2.2 Statistics of Calibration and Validation sets

On the basis of the 247 samples 2 data sets were constituted: one training set (n = 173) and one test set (n =74) for this 70% of the samples were picked randomly for the training set while the remaining 30% were kept as test set.

The test and training sets are within the range of variability of the whole data (table 2). Similar dispersion is observed for both sets, which insure a correct evaluation of the model developed on the training set (fig. 2).

Table 2: Descriptive statistics of Amylose content per set of samples

Amylose	N	Min	Max	Mean	SD
Training Set	173	15.49	32.96	24.99	3.40
Test Set	74	15.17	29.82	23.41	3.06

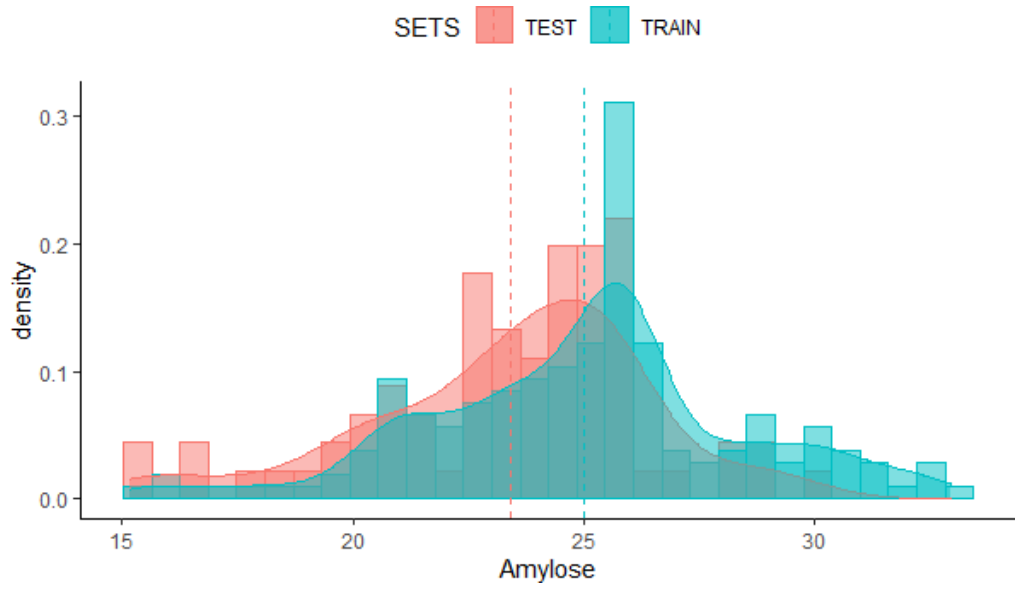

Figure 2: Amylose content distribution per set

2.3 Near Infrared Spectroscopy

The representation of the 247 spectra (fig.3) highlights the variability within the database in terms of response (absorbances), especially for water absorption bands (1500 and 1900 nm). All spectra presented similar patterns, no atypical spectra were detected.

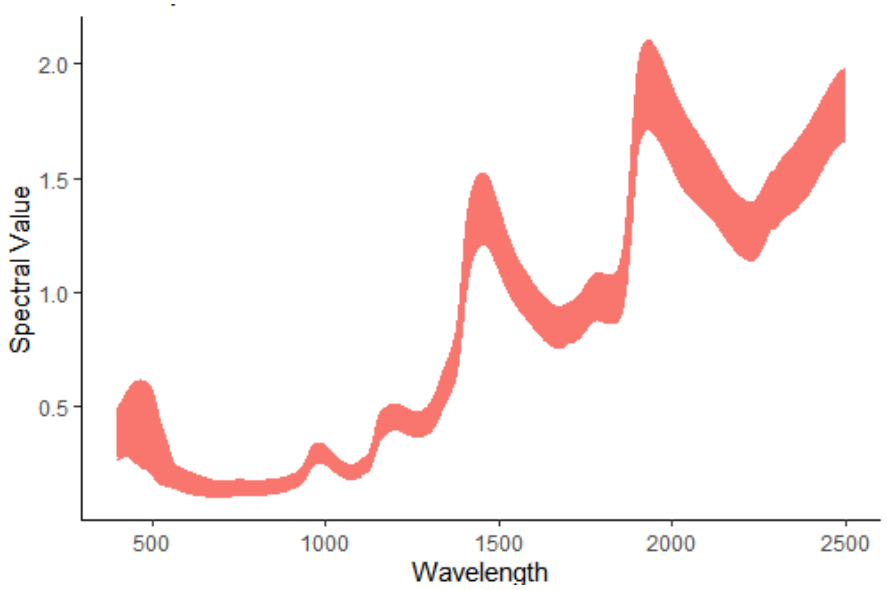

Figure 3: Absorbance spectra of the 247 samples

2.3.1 Spectra: Principal Components Analysis

A PCA calculated on the spectra (spectral range NIR) of the samples (247) led to 88.04% of variance explained by the 2 first PCs, (Figure 4). The Mahalanobis distances (GH) from the average spectrum were calculated on the PCs scores for the all the spectra, 3 spectra presented GH values higher than the limit (GH=3): P225_R3, P316_R1, and P225_R2. The maximum GH distance is observed for sample P225_R2 is 4.8. Considering these distances, all the samples are kept in the database.

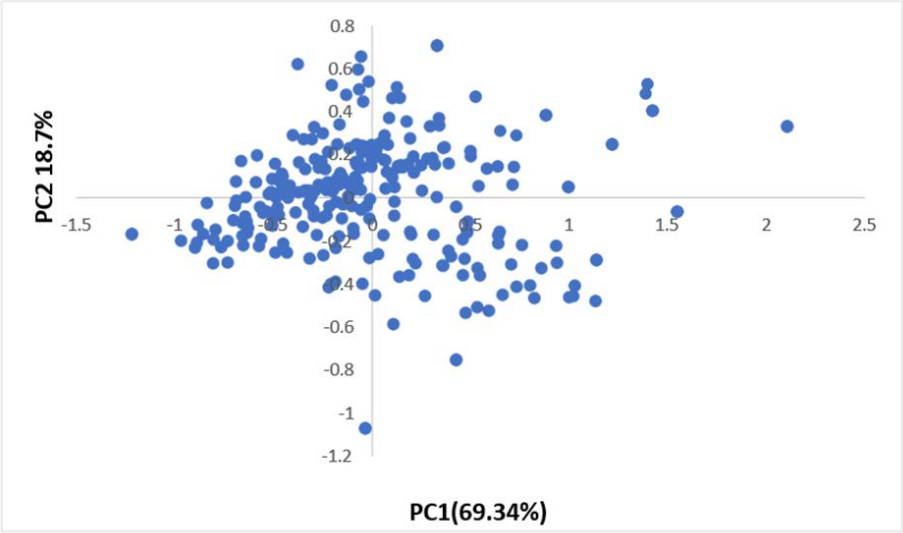

Figure 4: Scores for the 2 first PCs of the 247 spectra.

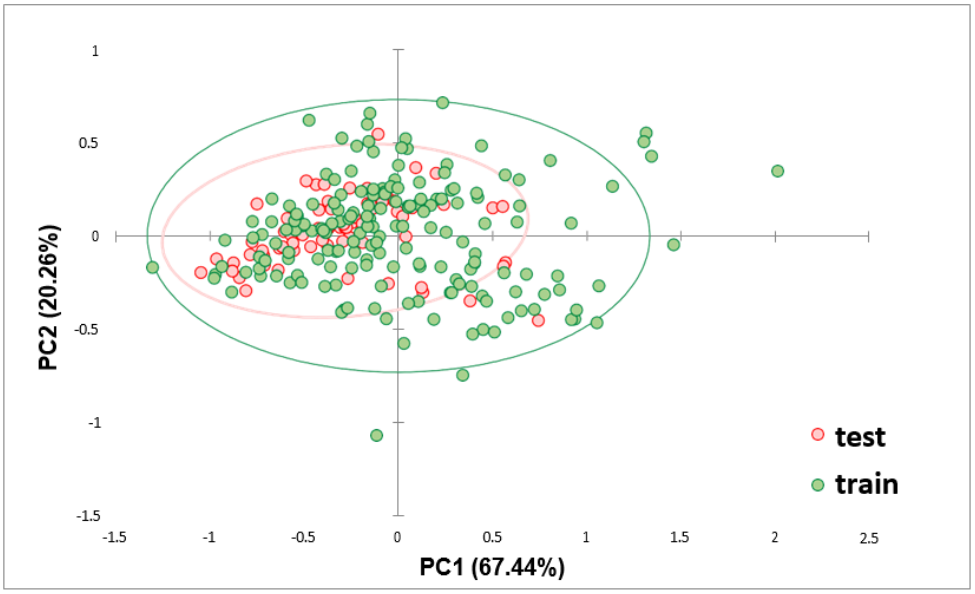

Figure 5: Scatter plot of test samples projections on PCA space of train samples

The projection of the test samples onto PCs space of the train samples (fig.5) confirms that the training set spectral variability is well representative of the total variability of the population.

2.4 Calibration

Amylose content was calibrated using Modified Partial Least Squares Regression, for the spectral range: 400 nm - 2500nm, that is to say Visible and NIR regions. Calibrations were done on non-pretreated spectra and pretreated spectra using different pre-treatments. The best model was selected based on highest R^2c, R^2cv, lowest SEC and SECV, minimum PLS factors and highest ratio SD/SEP (Table 3). The best model was obtained with no pre-treatment applied to the spectra, the R^2 was 0,937 with an SECV equal to 0,801%

Table 3: NIRS Calibration Model statistics for Amylose content

Constituent	N	Mean	SD	SEC	R^2c	SECV	R^2cv	SEP	R^2p	RPD
Amylose	164	24.96	3.04	0.764	0.937	0.801	0.930	1.023	0.915	2.97

This model was applied to predict samples from the test set, the standard error of prediction (SEP) is equal to SEP = 1,023% and the R^2 for prediction is 0,915. The ratio performance to deviation RPD is equal to 2,947. These performances indicated that this model is efficient to predict amylose content in fresh cassava roots. The scatter plot of predicted values versus laboratory values illustrates the quality of the fitting (fig.6)

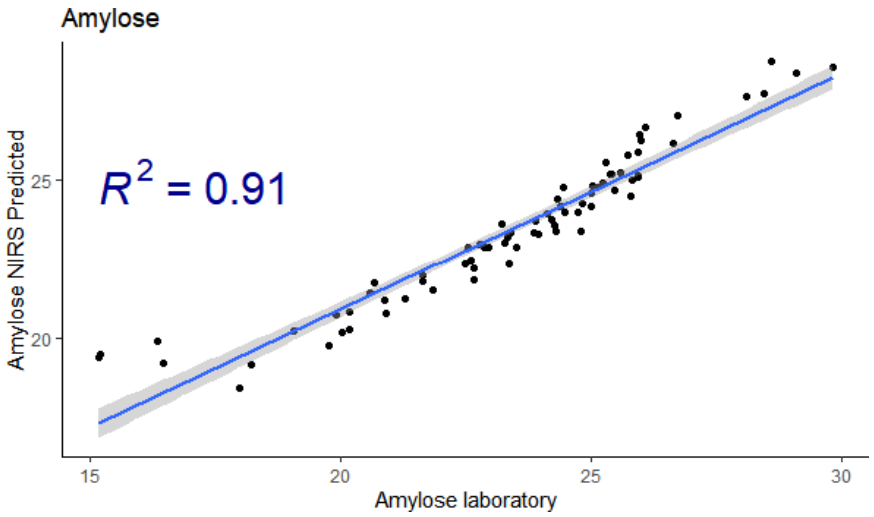

Figure 6: Scatter plot of Predicted values of amylose vs laboratory values for the test set

3 CONCLUSION

The developed NIRs model for quantification of amylose content of cassava root presents an accuracy good enough to enable cassava selection based on amylose content expressed as % of dried starch. The error of the model is SEP = 1,02% which means that a predicted value could be defined with a confidence interval of +/- 2,04 % associated with 95 % of confidence. Furthermore, the database refers to 94 different genotypes representative of the variability of the starch properties, in particular for amylose content, thus the model is robust enough for amylose quantification in fresh cassava. The procedure can be applied in cassava breeding programs based on amylose content.

Institute: Cirad – UMR QualiSud

Address: C/O Cathy Méjean, TA-B95/15 - 73 rue Jean-François Breton - 34398 Montpellier Cedex 5 - France

Tel: +33 4 67 61 44 31

Email: rtbfoodspmu@cirad.fr

Website: https://rtbfoods.cirad.fr/