# TRANSLATING CASSAVA ATTRIBUTES PREFERRED BY UGANDAN SMALLHOLDER FARMERS INTO BREEDING TARGETS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Paula Iragaba

December 2019

# TRANSLATING CASSAVA ATTRIBUTES PREFERRED BY UGANDAN SMALLHOLDER FARMERS INTO BREEDING TARGETS

Paula Iragaba, Ph.D.

Cornell University 2019

Cassava is the second most important staple food crop in Uganda. However, the adoption of improved varieties remains low, partly because some improved varieties do not have the attributes preferred by end-users. To contribute towards addressing this gap, I conducted three studies: 1) Examining cassava attributes preferred by smallholder farmers in four districts within Uganda. Results revealed yield (many roots and big root size),, early maturity, and root quality (taste, long in-ground storability of roots, softness of cooked roots, and non-bitter roots) were predominant attributes preferred by farmers; 2) Determining heritability of softness of cooked cassava roots, one of the key attributes preferred by end-users. Softness was evaluated at four cooking time intervals: 15, 30, 45, and 60 minutes using a penetrometer on 268 cassava genotypes. Estimates of broad-sense heritability ranged from 0.17 to 0.37, with the highest value observed at 45 min of cooking time. I also quantified the relationship between penetrometer and consumer-testing methods for phenotyping

softness of cooked cassava roots. The two methods were strongly correlated ($r^2$ = 0.91; *P-value* = 0.003), suggesting a penetrometer can be used for phenotyping softness; 3) Determining the genetic relationships of cassava varieties grown by farmers in Uganda, and the relationship between farmer-grown varieties and breeding lines. Genotyping-by-sequencing was used to score 287,952 single nucleotide polymorphisms in 547 samples of farmer-named cassava varieties collected by random sampling of 192 smallholder farms within four districts in Uganda, and a panel of 349 breeding lines from the cassava national breeding program in Uganda. The genetic differentiation we observed among farming districts in Uganda (mean $F_{ST}$ = 0.003) is similar to divergence observed within other countries. Despite the fact that none of the breeding lines were directly observed in farmer fields, genetic divergence between the populations was low ($F_{ST}$ = 0.020). Interestingly, we detected the presence of introgressions from the wild relative *M. glaziovii* on chromosomes one and four, which implies ancestry with cassava breeding lines. Our study highlights the importance of understanding the genetic make-up of cassava currently grown by smallholder farmers and relative to that of plant breeding germplasm.

## *BIOGRAPHICAL SKETCH*

Paula Iragaba was born on 07th September 1984 to Mr. and Mrs. Paul S. Manirakiza in a small town called Kisoro in Southwestern part of Uganda. Both parents were actively engaged in farming, and for most of the time, Paula ate the food produced locally from her parents' farm. Paula's passion for agriculture started at an early age, during her primary and secondary school, she keenly participated in school gardening in small plots apportioned to students for learning purposes under the guidance of senior teachers. Eventually, Agriculture became one of her favorite subjects during both her Ordinary and Advanced level secondary education attained from Kigezi High School (1998-2001) and St. Mary's College Senior Secondary School, Rushoroza (2004-2005) respectively. Paula performed well on her advanced level exams and was awarded a scholarship by the Government of Uganda to pursue a degree in Bachelor of Science in Agriculture at Makerere University (2006-2010). Paula's turning point happened in 2009 when she was granted an internship placement to work at Kachwekano Zonal Agricultural and Development Institute (KaZARDI) in Kabale, Uganda. During her internship, she was exposed to various agricultural technologies employed by the program to alleviate hunger and poverty among resource poor farmers in Uganda. During this same time, she got opportunities to work as an enumerator on various surveys, this helped her to directly work with farmers, listen to their challenges in agricultural production and processing and learn about their rich traditional knowledge about farming practices. Upon completing her Bachelor of Science degree in Agriculture, Paula won a competitive Alliance for Green Revolution

in Africa (AGRA) scholarship to pursue a Master of Science degree in Plant Breeding and Seed Systems at Makerere University (2011-2012). Her MSc. thesis focused on inheritance and stability of early maturity in potato (*Solanum tuberosum* L.). Upon successful completion of her MSc. training, she was recruited into NextGen Cassava breeding project in 2013 to further her career development in gender focused breeding initiatives. She pursued a Post Graduate Diploma in Gender Studies and Local Economic Development at Makerere University 2013/2014 prior to enrolling for a Ph.D. in Plant Breeding and Genetics at Cornell University during fall 2014. With all the knowledge and expertise gained over time, Paula intends to go back to sub-Saharan Africa and work as a Gender Responsive Breeder with the National Agricultural Research Organization in Uganda where she will contribute towards attaining food security and eradicating poverty.

This work is dedicated to my beloved parents Mr. and Mrs. Paul Senzoga Manirakiza

and to my precious husband Mr. Stephen Munyantwali Birikano

*ACKNOWLEDGMENTS*

First and foremost, I thank the Almighty God for giving me the gift of life and granting me the grace to successfully complete my Ph.D. study. I am grateful for the funding for this study which came from two sources: the CGIAR Program on Roots, Tubers and Bananas (RTB) and supported by CGIAR Trust Fund contributors; and the Next Generation Cassava Breeding Project through funds from the Bill and Melinda Gates Foundation (Project No. OPP1048542) and the Department for International Development of the United Kingdom (UKaid) (Project No. OPP1048542). I would like to sincerely thank my academic advisors Prof. Michael Allen Gore, Dr. Hale Ann Tufan and Prof. Rachel Nicole Bezner Kerr for their invaluable and unwavering support rendered to me from the start to this last step. This work would not have materialized without their phenomenal support. It was an enriching experience learning from them and I am forever grateful. In a similar manner, I am thankful to members of Cornell University especially Elizabeth D. Earle, Marnin Wolfe, Guillaume Bauchet, Ramu Punna, and the past and present Gore Lab members, your academic and social support has culminated into successful completion of my Ph.D. Sincere thanks also go to Dr. Robert Kawuki for mentoring and guiding me in my research activities in Uganda. I would like to acknowledge the support from NaCRRI Root Crops Program most especially Dr. Titus Alicai, Dr. Christopher Omongo, Dr. Ephraim Nuwamanya, Dr. Anthony Bua, Dr. Williams Esuma, Dr. Alfred Ozimati, Dr. Ismail Kayondo, Mr. Francis Osingada, Mr. Enoch Wembabazi, Mr. Vincent Kyaligonza and the entire team of people at NaCRRI who contributed to the success

of my work in one way or another. In a profound manner, I am grateful to the farmers from Apac, Arua, Kibaale, and Masindi who participated in this study. Special thanks to my American friends especially Jan Nye, Kelly Suzanne Frank, Leilah Krounbi, Clare Ulrich, Sarah Regenspan, Elsbeth Kane, Ellen Schilkowsky, and Adriana Hernaderz, their social support and love kept my hope alive throughout my Ph.D. journey. Finally, I would like to acknowledge support from my family: Mr. Stephen M. Birikano, Mr. and Mrs. Paul S. Manirakiza, brothers and sisters for standing with me, encouraging, and praying for me throughout my career development. My dear husband Stephen M. Birikano is further acknowledged for his kindness, understanding and allowing me to undertake my Ph.D. study. Your effort has yielded fruit. Glory be to God Now and Forever!

TABLE OF CONTENTS

*BIOGRAPHICAL SKETCH* ..... v

*ACKNOWLEDGMENTS* ..... viii

TABLE OF CONTENTS ..... x

*LIST OF FIGURES* ..... xii

*LIST OF TABLES* ..... xvi

CHAPTER 1 EXAMINING ATTRIBUTES OF CASSAVA PREFERRED BY UGANDAN MEN AND WOMEN SMALLHOLDER FARMERS: IMPLICATIONS FOR BREEDING AND ADOPTION OF NEW CASSAVA VARIETIES ..... 18

INTRODUCTION ..... 18

METHODOLOGY ..... 20

RESULTS ..... 25

DISCUSSION ..... 40

CONCLUSION ..... 43

REFERENCES ..... 44

CHAPTER 2 ESTIMATES FOR HERITABILITY AND CONSUMER-VALIDATION OF A PENETROMETER METHOD FOR PHENOTYPING SOFTNESS OF COOKED CASSAVA ROOTS ..... 54

ABSTRACT ..... 54

INTRODUCTION ..... 56

MATERIALS AND METHODS ..... 57

RESULTS ..... 68

DISCUSSION ..... 74

CONCLUSION ..... 76

REFERENCES ..... 78

CHAPTER 3 GENOMIC CHARACTERIZATION OF UGANDAN SMALLHOLDER FARMER-PREFERRED CASSAVA (*MANIHOT ESCULENTA* CRANTZ) VARIETIES ..... 84

ABSTRACT ..... 84

INTRODUCTION ..... 86

MATERIALS AND METHODS ..... 88

RESULTS ..... 93

DISCUSSION ..... 102

CONCLUSION ..... 106

REFERENCES ..... 108

CHAPTER 4 GENERAL CONCLUSIONS ..... 121

REFERENCES ..... 125

*LIST OF FIGURES*

Figure 1.1: Sampling frame used in selection of the study sites ..... 22

Figure 2.1. Scatter plot of harvest index (HI) against dry matter content (DMC) for 635 genotypes from a cassava diversity panel maintained by National Crops Resources Research Institute (NaCRRI), with the 78 genotypes (colored points) selected to sample the maximal phenotypic diversity. ..... 59

Figure 2.2. Principal Component Analysis (PCA) plot of PC1 and PC2 of the genetic diversity in the NaCRRI population, with only the 78 selected genotypes colored and labeled according to which of the four K-means clusters each genotype belongs. The PCA is based on 61,297 single-nucleotide polymorphism markers scored on 635 cassava genotypes. ..... 60

Figure 2.3. Scatter plot of harvest index (HI) against dry matter content (DMC) for the 78 selected genotypes labeled according to their K-means cluster. ..... 61

Figure 2.4. Pearson correlation coefficients (r) for best linear unbiased predictors of cassava root softness measured by a penetrometer at the end of four cooking time points: 15, 30, 45, and 60 min. All correlations were significant at $\alpha = 0.05$. ..... 70

Figure 2.5. Time series of back-transformed BLUEs for root softness measured by the penetrometer at six post-cooking time points for six cassava genotypes. ..... 72

Figure 2.6. The relationship between BLUEs of penetrometer measurements and consumer scores of softness for cassava roots of six genotypes that were cooked for 50 minutes. ..... 73

Figure 2.7. (a) Bar plot showing Tukey-Kramer HSD means separation of softness for six cassava genotypes based consumer testing evaluation; (b) Bar plot showing Tukey-Kramer HSD means separation of softness for six cassava genotypes obtained by use of the penetrometer. Means followed by the same letter are not significantly different at $\alpha = 0.05$. ..... 74

Figure 3.1. a). Dendrogram from Ward's hierarchical clustering of IBS matrix and b). Population structure (based on subpopulations; K=14) analysis on 547 cassava samples using ADMIXTURE. In the panel, each sample is indicated by a vertical bar partitioned in one or more colored segments, and the respective length of the bar represents the proportion of the individual's genome ancestry in a given subpopulation. Each color represents a different ancestry subpopulation. ..... 95

Figure 3.2. Principal components 1 and 2 based on 189,851 genome-wide SNPs scored on 119 genetically unique cassava varieties randomly selected from the clonal groups in each district after correction for multiple samples within the same clonal group (if a unique variety was represented in two or more districts, it had one entry for each of the districts) colored by district from where the samples were collected indicated minimal clustering in relation to the source of the sample. ..... 97

Figure 3.3. Correspondence between genetic identity (clonal groups labeled G_3 to G_423) (upper semi-circle) and most common variety names given by farmers (lower semi-circle). Clonal groups with more than 11 members were considered, and farmer-reported variety names that were mentioned more than 11 times in total are presented in this plot. The numbers on the axis represents the members in a given clonal group or the number of times a given variety name was mentioned during the surveys. The label for "Other_names" represent all other variety names that had a count less than 12 times. Ton represents Tonguda, Kya represents Kyawada, acol represents Gamente acol, Gam represent Gamente, NAD represents NAADs, Mab stands for Mabulu, Omo stands for Omoo, atar stands for Gamente atar, Lon stands for Longe, Buk stands for Bukalasa, Gwa stands for Gwalanda. ..... 98

Figure 3.4. Principal component analysis plots of PC1 and PC2 based on 190,556 genome-wide SNPs scored on 547 samples colored by most common a). Clonal groups from IBS-based distinctions b) Names of the cassava varieties reported by farmers. ..... 99

Figure 3.5. The principal component analysis plots of PC1 and PC2 (a), and PC3 and PC4 (b) based on 349 breeding lines and 119 unique farmers' varieties scored on 201,117 SNPs. The plots show patterns of structure between cassava varieties cultivated by smallholder farmers in four districts (Kibaale, Apac, Arua, and Masindi) in Uganda and the cassava breeding lines at NaCRRI, Uganda. Only unique varieties per district are used for the farmer varieties. ..... 100

Figure 3.6. a) Genome-wide Fst between cassava breeding lines at National Crops Resources Research Institute (NaCRRI) and farmer-varieties, b) Nucleotide diversity (pi) ratio of cassava breeding lines at NaCRRI to varieties grown by smallholder farmers in Uganda. ..... 102

*LIST OF TABLES*

Table 1.1. Predominant cassava attributes smallholders farmers from different locations prefer about cassava varieties they cultivate ..... 28

Table 1.2. Proportion (percentages) of predominant cassava attributes men and women farmers from different locations prefer about cassava varieties they cultivate ..... 29

Table 1.3. Predominant cassava attributes men and women farmers from different locations like about cassava varieties they cultivate ..... 29

Table 1.4. Predominant cassava attributes smallholder farmers disaggregated by age groups from different locations like about cassava varieties cultivated in the respective districts ..... 31

Table 1.5. Illustrative quotes for selected codes most frequently mentioned attributes about cassava by farmers who participated in semi-structured interviews ..... 32

Table 1.6. Predominant cassava attributes reported by smallholder farmers from different locations to be desirable in an ideal cassava variety suitable for fresh root consumption ..... 35

Table 1.7. Desirable attributes for an ideal cassava variety disaggregated by district and sub-county..... 36

Table 1.8. Predominant attributes associated with an ideal cassava for fresh consumption disaggregated by district and sex ..... 37

Table 1.9. Predominant attributes associated with an ideal cassava for fresh consumption disaggregated by district and age-group ..... 38

Table 1.10. Different forms in which cassava is consumed across districts ..... 39

Table 2.1. Best linear unbiased predictors and heritability estimates of cassava root softness following four different cooking durations. ..... 69

Table 2.2. F-values for fixed sources of variation from an analysis of variance for consumer scores of softness for roots from six cassava varieties that were cooked for 50 minutes ..... 71

Table 2.3. F-values for fixed sources of variation from an analysis of variance for penetrometer measurements of softness for roots from six cassava varieties. ..... 72

Table 3.1. Summary of leaf samples collected from cassava plants grown by smallholder farmers within four districts in Uganda and the number of unique varieties per household..... 93

Table 3.2. $F_{ST}$ estimates among cassava varieties cultivated by smallholder farmers in four districts of Uganda. ..... 96

Table 3.3. Average nucleotide diversity per chromosome per district..... 101

# CHAPTER 1 EXAMINING ATTRIBUTES OF CASSAVA PREFERRED BY UGANDAN MEN AND WOMEN SMALLHOLDER FARMERS: IMPLICATIONS FOR BREEDING AND ADOPTION OF NEW CASSAVA VARIETIES

## INTRODUCTION

Cassava (*Manihot esculenta* Crantz), which ranks as the sixth major crop after wheat, rice, maize, potato and barley, is an important source of food for over 800 million people globally (Lebot, 2009). In sub-Saharan Africa (SSA), it is the second most important staple food crop (Nweke et al., 2002). The primary product of cassava are the starchy roots which can be utilized into a wide range of products for food, feed and industry (El-Sharkawy, 2004; FAO, 2013a; Chiwona-Karltun et al., 2015), nevertheless, leaves are also eaten as a vegetable in some places (Kombo et al., 2012). The roots may be consumed as non-fermented or fermented products depending on the level of cyanogenic glucosides (Montagnac et al., 2009a; Bechoff et al., 2016; FAO, 2013b). "Sweet" cassava roots often have low quantities of the cyanogenic glucosides, and are the only ones that can be consumed in the fresh form, while bitter cassava roots need to be processed before they are consumed to rid the root of hazardous cyanogens (Cardoso et al., 2005; Montagnac et al., 2009).

Globally, cassava is mainly grown by smallholder farmers (FAO, 2013a) as it fits well in the farming systems of resource-poor farmers due to its inherent ability to give appreciable yields under marginal lands and limited rainfall where other crops fail (El-Sharkawy, 2004). Cassava also has flexible harvesting and planting schedules, and since it is clonally propagated, planting material does not compete with the edible part of the plant (Nweke et al., 2002). Collectively, these attributes make cassava a suitable crop for smallholder farmers who may have limited resources to invest in crop production.

Previous studies have highlighted that cassava attributes preferred by end-users influence the adoption of new varieties (Alene et al., 2013; Chiwona-Karltun et al., 2015; Bechoff et al., 2016; Teeken et al., 2018). Thus, understanding cassava attributes preferred by end-users can help determine the drivers of adoption which could in turn help plant breeders in setting their breeding priorities to be more demand driven. As an example, a study conducted in Malawi revealed that bitter cassava varieties were more preferred because they yielded more and produced better quality of food in addition to being less liable to theft, destruction by animals and unplanned harvests (Chiwona-Karltun et al., 1998). Preferences of end-users may be influenced by gender roles, norms, and relations in cassava production, processing and marketing in a given country (Masamha et al., 2018). Gender issues have been shown to impact technology adoption (Bezner Kerr, 2017) for example there were differential adoption of chemical use and improved maize varieties amongst men and women farmers which was a result of unequal access to inputs (Doss and Morris, 2001). This necessitates gender analysis of drivers of trait preferences to better design and deliver new varieties to farmers. This is particularly important for cassava production systems where women play a significant role in cassava production and processing to ensure food security and income generation for a household (Ogunlela and Mukhtar, 2009; Forsythe et al., 2016). Food security is dependent on the functioning of a sustainable agricultural production system which involves the interaction of four factors of production: land, labor, capital and entrepreneurship (Silver, 2019). For all these, gender plays a key shaping role in determining access to and control of available resources. Understanding existing gender relations could provide insights in uptake and sustainability of new technologies and addressing issues in food security. There are aspects of socio-cultural expectations, roles, rights, privileges, norms, and opportunities of men and women that can influence food security in a given community (Bezner Kerr, 2017).

In Uganda, cassava is utilized both as a food security crop and source of income for the smallholder farmers (Otim-Nape et al., 2005; UBOS, 2010; Nakabonge et al., 2018). The national root crops breeding program has made efforts to improve productivity of the cassava with the main aim of increasing high yields and disease resistance, however, some of the new varieties released lack important quality attributes desirable to farmers, contributing to low adoption (Kawuki et al., 2011). Farmers consider various attributes such as cooking quality, in-ground storability, and early maturity in deciding to conserve cassava varieties (Nakabonge et al., 2018). In order to identify the cassava attributes farmers like, there is need to work closely with them to understand their priorities. A study conducted by Turyagyenda et al. (2012) highlighted that the landraces preferred by farmers may be genetically distinct from elite materials used by breeders, thus use of the landraces with attributes preferred by farmers in breeding new varieties could potentially lead to higher chances of adopting the new varieties (Kizito et al., 2007). The overall aim of this study was to document cassava attributes preferred by farmers in Uganda with the purpose of ultimately contributing to developing varieties suitable for end-user needs and preferences. Specifically, we examined attributes of cassava preferred by men and women farmers within four districts in Uganda.

## METHODOLOGY

### *Study sites*

This study was conducted within four districts in Uganda; Arua, Apac, Masindi, and Kibaale. We selected these four districts based on the existing literature on cassava production, utilization, and disease prevalence (UBOS, 2010), and the findings from a site selection online survey that was administered to key informants (Supplementary File 1). The selected districts had high cassava production, high consumption of cassava and low disease prevalence, this selection criterion was used because we wanted areas where cassava plays a

key role in the livelihoods of people and we selected areas with low disease pressure to avoid skewing farmer preferences in one direction. After selection of the four districts, a scoping study was then conducted in each of the districts to select sub-counties and villages where the survey was to be carried out (Figure 1.1). Based on the study objectives and guiding research questions; we developed a key informant (KI) guide (Supplementary File 2) to be used in selecting sub-counties in each of the selected districts. The key informants for identifying the sub-counties were the district agricultural officers and district production officers for the respective districts which had been selected. The agricultural and production officers were requested to recommend five sub-counties within their respective districts which had high cassava production, high cassava utilization and low disease pressure. This selection criterion was used also to get study sites where cassava is the main source of livelihood for the people. Based on the five sub-counties recommended by the district officials, we randomly selected two sub-counties per district where we could do the research.

Once the sub-counties were selected, we listed all the villages in each of the selected sub-counties, the data for the list of villages per sub-county was obtained from Uganda Bureau of Statistics. For any given sub-county, we randomly selected three villages using *sample* function in *base* package in R (R Core Team, 2018). We then developed another key informant guide (Supplementary File 3) to aid in identifying which village out of the three that had been selected in a given sub-county could be our actual study site for the survey. A scoping study was conducted in each of the villages, mainly to gather information relating to population size, livelihood activities, main crops, production of cassava, utilization of cassava and average number of varieties grown by individual farmers. The information from the village scoping study was evaluated to select one village per sub-county, the villages selected had high cassava production, utilization in form of fresh root consumption and high diversity.

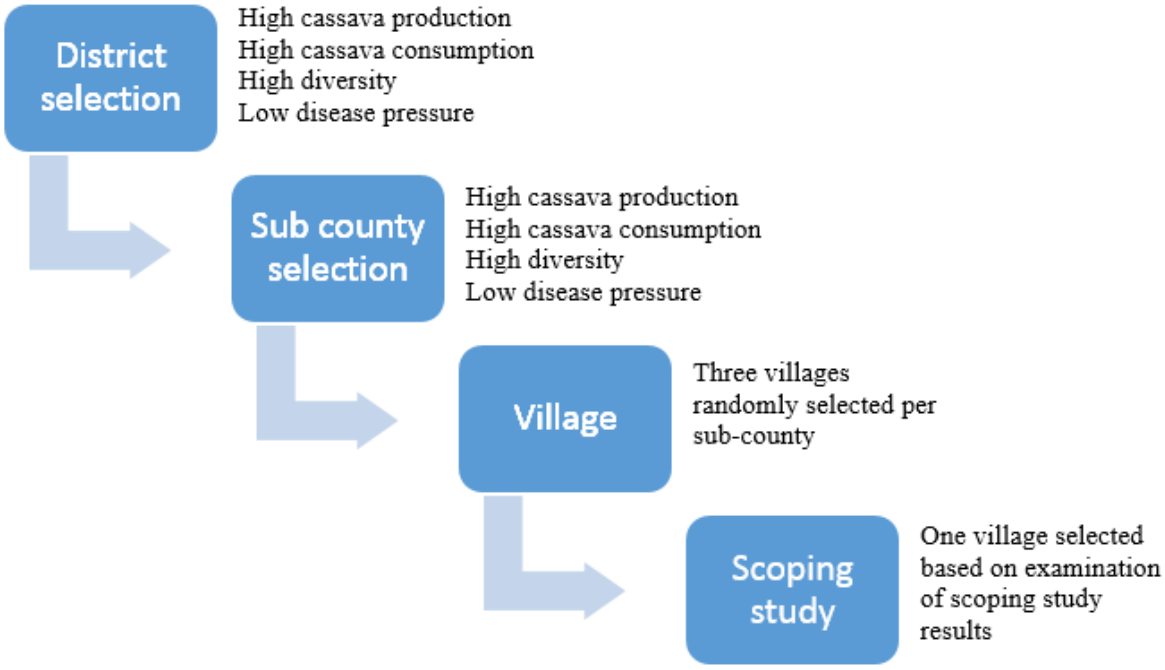



Figure 1.1: Sampling frame used in selection of the study sites

**Sampling procedures to select participants**

The plan for this study and consent forms were reviewed by the Institutional Review Board (IRB) of Cornell University and approved (IRB ID 1502005316). In each of the villages, stratified simple random sampling was used to select 64 smallholder farmers involved in cassava production and processing to participate in the study. Village rosters were used to get names of all farmers, the names were then subdivided into four strata based on age and sex i.e. younger female group (18-30 years), older female group (31-70 years), younger male group (18-30 years), and older male group (31-70 years). The age group for declaring younger and older farmers was based on the Uganda national youth policy (http://www.youthpolicy.org/national/Uganda_2001_National_Youth_Policy.pdf). After obtaining the four different strata, 16 farmers were randomly selected per stratum, of the selected farmers, 6 were randomly selected to participate in the semi-structured interviews while the remaining 10 participated in the focus group discussions.

**Training of enumerators and data collection**

Three different teams of enumerators were formed based on the most common languages spoken in the study sites: Lugbara in Arua district, Langi in Apac district and Runyoro in both Masindi and Kibaale. Each team of enumerators had high proficiency in the language spoken in a given district, and interviews were conducted in the local language. Experience collecting qualitative and quantitative data was also an important criterion in selecting enumerators.

Enumerators were trained following Johns Hopkins Bloomberg School of Public Health (JHSPH) human subjects research ethics field training guide developed by Johns Hopkins University (JHSPH, 2009). One enumerator from each team translated all questions in the tools to the local languages spoken in the respective districts where they were going to work. In order to ensure the original message was not altered, another member from the team back-translated the tool from local language to English. Then all members for each of the teams met together and went again over the tools. The tools were pre-tested by each enumerator by interviewing farmers who spoke one of the languages in the targeted study areas, necessary corrections were made accordingly.

Data collection started early June, 2015 in Masindi district followed by Arua, then Apac and data was collected last in Kibaale district by late July 2015. To collect data, we used both quantitative and qualitative methods. From the semi-structured interviews, we collected both qualitative and quantitative data (Supplementary File 4). Additional qualitative data was collected through focus group discussions (FGDs) (Supplementary File 5) with men and women smallholder farmers. The in-depth interviews enabled us to gather information on household demographics, varieties cultivated and the attributes of cassava preferred at an

individual level while the focus group discussions helped to understand the cassava attributes preferred by men and women smallholder farmers at a village level.

During the in-depth individual interviews, the enumerators read out the questions to the respondents in the local language and recorded the respondent's answers on paper. The questions were administered by enumerators to ensure inclusion of respondents that could not read and write. During the FGDs, we used two note takers: one took the notes in a note book while the other person took notes on flip-chart hanged on a wall. Taking notes by two independent people helped us to double check the quality of our data. Additionally, both the semi-structured interviews and focus group discussions were audio recorded to ensure collection of all the information provided by the farmers.

After data collection, the quantitative data in semi-structured interviews was entered in excel sheets. The qualitative data from the 192 farmers who participated in the surveys and the notes taken during the 32 different FGDs with men and women smallholder farmers were compiled into notes under common questions. Transcription of all the audio recordings to word documents was done to do a quality check of what had been written by enumerators and also extract verbatim responses to strengthen our interpretation of the collected data. Since both the interviews and FGDs were conducted in the local language, the transcribed documents were translated from the respective local languages (Runyoro, Langi and Lugbara) to English.

**Coding of the data and analysis**

A coding frame was developed for variables that had been collected using qualitative methods e.g. variables that related to different cassava varieties grown, reasons for growing cassava, cassava attributes liked or disliked by the farmer, articulation of ideal cassava varieties, different forms in which cassava was consumed (Questions 35-40, and 43-54 in the

semi-structured interview). Before developing a coding frame, all the interviews were read to capture the common themes and patterns pertaining to the attributes of cassava preferred by the respondents to get acquainted with the range of responses provided. We then developed a coding frame where we assigned a word or phrase (code) to common themes in the responses (Supplementary File 6). The coding frame table consisted of four different sections: code, definition or description of the code, relevant quotations from the interviews, and the relevant phrase in the local language respectively. After developing the coding frame, we used it to code the qualitative questions in all the 192 interviews. Upon completion of the coding process, the codes for all questions were entered in the excel sheets. The data in excel sheets was imported into SPSS and data was analyzed using descriptive statistics to obtain the frequency and percentages of the responses to each variable. A chi-square test of independence was used to discover the relationships between the given categorical variables from the semi-structured interviews.

## RESULTS

In this section, the results from the questions relating to preferences of smallholder farmers are discussed in regard to the cassava varieties they were growing, attributes of an ideal cassava variety for fresh root consumption, and best varieties to be consumed for certain food products. These results are differentiated by location, sex and age-group of the farmers, the key elements used in selecting the participants in the study.

### Preferences about the varieties cultivated

During the semi-structured individual interviews, men and women smallholder farmers were asked about the attributes of cassava varieties they were growing. The predominant preferred attributes (in terms of the percentage of a given attribute that was mentioned by respondents) were associated with yield (many roots and big root size), early

maturity, and root quality in terms of taste after cooking, long in-ground storability of cassava roots after reaching their harvest maturity, softness of cooked roots and non-bitter roots (Table 1.1). Most of these cassava attributes had a significant relationship (*P-value* < 0.05) to the district where the farmers were situated (Table 1.1). In contrast, attributes such as big root size and short cooking time were not significantly associated to the district where the farmers were situated. We observed that farmers in Arua had very low proportion of preference for tasty and soft cassava roots as opposed to farmers in other districts. Farmers in Apac district also had notably low proportion of preference for in-ground long storability of cassava roots as opposed to farmers in other districts. Farmers in all districts except for Kibaale, reported relatively low preference for disease resistance.

We also investigated whether there was a relationship between the preferred cassava attributes and the different sub-counties within a given district. Generally, there were no significant differences (*P-value* > 0.05) for cassava attributes preferred in the different sub-counties within a given district (Table 1.2). However, in some instances, the different sub-counties had significantly different proportions of preference for a given attribute, for example 35.2% of the farmers in Ogoko sub-county mentioned that they preferred big sized cassava roots, while in Logiri sub-county, it was only 5.4% of the farmers that preferred big sized roots (Table 1.2).

Further analysis of data on attributes preferred by smallholder farmers about the cassava varieties grown revealed that, there were significant differences between males and females for the proportion of at least one of the cassava attributes preferred in three of the districts (Table 1.3). For example in Apac district, the proportion of men preferring high number of roots and early maturity was significantly higher (*P-value* < 0.05) than the proportion of women. In contrast, in Arua district, women preferred bigger root size than

men, while in Kibaale district, the proportion of men who liked big root size was significantly higher (*P-value* < 0.05) than the proportion for women who wanted big root size. Additionally, in Kibaale district, sex of the farmer was associated with the preference for disease resistance, men preferred disease resistant cassava varieties more than women. In Masindi, there was no relationship (*P-value* > 0.05) between the sex of the farmer and the cassava attributes preferred (Table 1.3).

Most of the districts had significant relationship (*P-value* > 0.05) between the cassava attribute preferred and the age of the farmers (Table 1.4). For instance, in Arua and Masindi, the proportion of younger farmers preferring early maturity was significantly higher (*P-value* < 0.05) than the proportion of older farmer farmers who reported preferring the same attribute. Similarly, in Kibaale district a significantly higher (*P-value* < 0.05) proportion of younger farmers relative to older farmers preferred cassava flour that sticks together during cooking the local dish known as *Kawunga*. On the other hand, there were some cassava attributes where both younger and older farmers had similar preferences (Table 1.4).

Table 1.1. Predominant cassava attributes smallholder farmers from different locations prefer about cassava varieties they cultivate.

| | Proportion (%) of preferences per district | | | | |
|---|---|---|---|---|---|
| Cassava attribute | Apac | Arua | Kibaale | Masindi | *P-value*[*] |
| Many roots | 36.5 | 22.8 | 48.4 | 35.4 | **0.000** |
| Matures fast | 28.7 | 25.5 | 46.6 | 25.9 | **0.000** |
| Big roots | 22.6 | 20.6 | 18.6 | 17.6 | 0.774 |
| Tasty | 13.9 | 2.8 | 45.3 | 28.6 | **0.000** |
| Long storage | 7.0 | 20.0 | 24.2 | 21.1 | **0.003** |
| Soft | 20.0 | 1.4 | 29.2 | 25.2 | **0.000** |
| Not bitter | 14.8 | 20.0 | 5.0 | 15.6 | **0.001** |
| Disease resistant | 2.6 | 4.1 | 28.0 | 8.2 | **0.000** |
| Flour sticks | 1.7 | 2.8 | 11.2 | 5.4 | **0.002** |
| Long roots | 6.1 | 1.4 | 8.7 | 4.1 | **0.029** |
| Cooks fast | 5.2 | 2.8 | 5.0 | 2.7 | 0.555 |
| Non-fibrous | 7.0 | 2.1 | 5.6 | 1.4 | **0.047** |

[*]*P-value* for differences in cassava attributes between districts. The bold *P-values* were significant at $\alpha < 0.05$.

Table 1.2. Proportion (percentages) of predominant cassava attributes men and women farmers from different locations prefer about cassava varieties they cultivate.

| | Apac | | | Arua | | | Kibaale | | | Masindi | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cassava attribute | Aduku | Apac | *P-value* | Logiri | Ogoko | *P-value* | Nkooko | Nyamarunda | *P-value* | Miirya | Pakanyi | *P-value* |
| Many roots | 33.3 | 40.4 | 0.434 | 17.6 | 28.2 | 0.128 | 50.0 | 46.8 | 0.688 | 33.3 | 37.7 | 0.582 |
| Matures fast | 27.0 | 30.8 | 0.655 | 24.3 | 26.8 | 0.737 | 48.8 | 44.3 | 0.569 | 23.1 | 29.0 | 0.414 |
| Big roots | 23.8 | 21.2 | 0.735 | 5.4 | 35.2 | **0.000** | 19.5 | 17.7 | 0.771 | 9.0 | 27.5 | **0.003** |
| Tasty | 15.9 | 11.5 | 0.504 | 1.4 | 4.2 | 0.291 | 54.9 | 35.4 | **0.013** | 19.2 | 39.1 | **0.008** |
| Long storage | 4.8 | 9.6 | 0.309 | 20.3 | 19.7 | 0.934 | 31.7 | 16.5 | **0.024** | 19.2 | 23.2 | 0.557 |
| Soft | 9.5 | 32.7 | **0.002** | 2.7 | 0.0 | 0.163 | 34.1 | 24.1 | 0.159 | 33.3 | 15.9 | **0.015** |
| Not bitter | 12.7 | 17.3 | 0.488 | 21.6 | 18.3 | 0.618 | 4.9 | 5.1 | 0.957 | 17.9 | 13.0 | 0.414 |
| Disease resistant | 0.0 | 5.8 | 0.053 | 0.0 | 8.5 | **0.011** | 22.0 | 34.2 | 0.084 | 6.4 | 10.1 | 0.409 |
| Flour sticks | 0.0 | 3.8 | 0.116 | 2.7 | 2.8 | 0.967 | 12.2 | 10.1 | 0.677 | 3.8 | 7.2 | 0.364 |
| Long roots | 4.8 | 7.7 | 0.513 | 0.0 | 2.8 | 0.146 | 11.0 | 6.3 | 0.296 | 2.6 | 5.8 | 0.323 |
| Cooks fast | 0.0 | 11.5 | **0.006** | 5.4 | 0.0 | **0.047** | 2.4 | 7.6 | 0.132 | 5.1 | 0.0 | 0.056 |
| Non-fibrous | 11.1 | 1.9 | 0.054 | 2.7 | 1.4 | 0.584 | 4.9 | 6.3 | 0.689 | 1.3 | 1.4 | 0.930 |

*P-value*

Table 1.3. Predominant cassava attributes men and women farmers from different locations like about cassava varieties they cultivate.

| | Apac | | | Arua | | | Kibaale | | | Masindi | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cassava attribute | Males | Females | *P-value* | Males | Females | *P-value* | Males | Females | *P-value* | Males | Females | *P-value* |
| Many roots | 45.6 | 26.8 | **0.026** | 15.1 | 30.6 | 0.100 | 55.1 | 42.2 | 0.669 | 33.8 | 37.1 | 0.566 |
| Matures fast | 38.6 | 17.9 | **0.014** | 27.4 | 23.6 | 0.601 | 53.8 | 39.8 | 0.073 | 28.6 | 22.9 | 0.429 |
| Big roots | 24.6 | 21.4 | 0.692 | 9.6 | 30.6 | **0.002** | 26.9 | 10.8 | **0.009** | 15.6 | 20.0 | 0.483 |
| Tasty | 12.3 | 14.3 | 0.753 | 2.7 | 2.8 | 0.989 | 47.4 | 43.4 | 0.605 | 26.0 | 31.4 | 0.465 |
| Long storage | 8.8 | 5.4 | 0.479 | 13.7 | 26.4 | 0.056 | 24.4 | 24.1 | 0.969 | 23.4 | 18.6 | 0.476 |
| Soft | 17.5 | 21.4 | 0.602 | 2.7 | 0.0 | 0.157 | 34.6 | 24.1 | 0.142 | 24.7 | 25.7 | 0.885 |
| Not bitter | 14.0 | 16.1 | 0.762 | 16.4 | 23.6 | 0.280 | 5.1 | 4.8 | 0.928 | 14.3 | 17.1 | 0.634 |
| Disease resistant | 1.8 | 3.6 | 0.548 | 4.1 | 4.2 | 0.986 | 38.5 | 18.1 | **0.004** | 11.7 | 4.3 | 0.102 |
| Flour sticks | 0.0 | 3.6 | 0.150 | 4.1 | 1.4 | 0.317 | 12.8 | 9.6 | 0.522 | 2.6 | 8.6 | 0.111 |
| Long roots | 8.8 | 3.6 | 0.252 | 0.0 | 2.8 | 0.152 | 11.5 | 6.0 | 0.215 | 3.9 | 4.3 | 0.905 |
| Cooks fast | 0.0 | 8.9 | **0.021** | 1.4 | 4.2 | 0.304 | 7.7 | 2.4 | 0.123 | 2.6 | 2.9 | 0.923 |
| Non-fibrous | 5.3 | 8.9 | 0.448 | 1.4 | 2.8 | 0.522 | 7.7 | 3.6 | 0.260 | 1.3 | 1.4 | 0.946 |

Table 1.4. Predominant cassava attributes smallholder farmers disaggregated by age groups from different locations like about cassava varieties cultivated in the respective districts.

| | Apac | | | Arua | | | Kibaale | | | Masindi | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cassava attribute | Aged 18-30 years | Aged 31-70 years | *P-value* | Aged 18-30 years | Aged 31-70 years | *P-value* | Aged 18-30 years | Aged 31-70 years | *P-value* | Aged 18-30 years | Aged 31-70 years | *P-value* |
| Many roots | 39.5 | 34.3 | 0.573 | 25.0 | 21.6 | 0.651 | 39.7 | 54.1 | 0.074 | 40.5 | 33.3 | 0.413 |
| Matures fast | 32.6 | 25.7 | 0.433 | 37.5 | 19.6 | **0.020** | 49.2 | 44.9 | 0.593 | 49.2 | 44.9 | **0.032** |
| Big roots | 23.3 | 22.9 | 0.961 | 18.8 | 20.6 | 0.791 | 23.8 | 15.3 | 0.176 | 21.4 | 16.2 | 0.452 |
| Tasty | 16.3 | 11.4 | 0.461 | 2.1 | 3.1 | 0.727 | 39.7 | 49.0 | 0.247 | 38.1 | 24.8 | 0.106 |
| Long storage | 2.3 | 10.0 | 0.123 | 18.8 | 20.6 | 0.791 | 20.6 | 26.5 | 0.394 | 21.4 | 21.0 | 0.949 |
| Soft | 25.6 | 15.7 | 0.198 | 0.0 | 2.1 | 0.316 | 28.6 | 29.6 | 0.889 | 23.8 | 25.7 | 0.810 |
| Not bitter | 16.3 | 14.3 | 0.774 | 15.0 | 20.8 | 0.860 | 7.9 | 3.1 | 0.165 | 11.9 | 17.1 | 0.430 |
| Disease resistant | 4.7 | 1.4 | 0.301 | 2.1 | 5.2 | 0.382 | 33.3 | 24.5 | 0.222 | 7.1 | 8.6 | 0.775 |
| Flour sticks | 2.3 | 1.4 | 0.726 | 4.2 | 2.1 | 0.466 | 17.5 | 7.1 | **0.043** | 0.0 | 7.6 | 0.066 |
| Long roots | 7.0 | 5.7 | 0.787 | 0.0 | 2.1 | 0.316 | 6.3 | 10.2 | 0.397 | 4.8 | 3.8 | 0.792 |
| Cooks fast | 4.7 | 4.3 | 0.927 | 2.1 | 3.1 | 0.727 | 3.2 | 6.1 | 0.401 | 2.4 | 2.9 | 0.873 |
| Non-fibrous | 4.7 | 8.6 | 0.430 | 4.2 | 1.0 | 0.212 | 6.3 | 5.1 | 0.737 | 0.0 | 1.9 | 0.386 |

Illustrative quotes for most frequently mentioned attributes about cassava by farmers who participated in semi-structured interviews are presented in Table 1.5.

Table 1.5. Illustrative quotes for selected codes most frequently mentioned attributes about cassava by farmers who participated in semi-structured interviews.

| **Attributes about cassava (codes)** | **Quotations from interviews with farmers** |
|---|---|
| Soft | "Being soft properly, it doesn't disturb someone eating. It is easy to eat because it soft like sweet potatoes". |
| | "Even a young child can eat it because it is soft. Since most of us here have problems with our teeth, it becomes easy for us" |
| | "You can determine a nice cassava for eating from when it gets ready. If it's soft…!! And it is not hard after cooking and its ready, it should be soft enough." |
| Tasty | "*Timpaigwamurwaire* is very much tasty; it's our own from ages ago. |
| | When cassava is tasty, I eat and get satisfied. You cannot easily eat cassava which is not tasty and you get satisfied. Because of the great tasty it has, I said to myself, I must have this variety in my home". |
| In-ground storability | "I cultivate *Mukuma* as food security. You don't eat it very fast like this one. This one saves you from hunger very fast but it gets done very fast but *Mukuma* even if its two years when the soils are good, it will last. It is very good, to help even when you eat other food you know your granary is there". |
| | "I like it because it can stay in the garden for long without getting spoilt even if for many years just like you can decide to leave a banana to get so ready it can also stay for over three years. |
| Many roots | "You can harvest from the same plant up to three times before exhausting it, there is also another called *Bakanyisa*, I do plant it because it does produce many roots per plant". |
| Matures fast | "It matures fast like in six months, you are already eating its cassava. You can easily get money especially to those who fry it for selling. It matures within a short time. It does not take a full year. It takes between 6-7 months when you can start harvesting it if you get problems. I admired it because it matures faster and enables one get money faster." |
| | "It yielding very fast gives me the happiness, I am hopeful that after one year I start to harvest it for breakfast, relieves me from hunger." |

| | |
|---|---|
| | "If it yields quickly, you have something to feed on, it boils quickly when you don't have something (poverty)." |
| | "It should be a type which doesn't take a lot of time to mature because it's hard to get any other thing for taking tea." |
| Not bitter | "Because its sweet (not bitter at all), you can use it for making breakfast in case a visitor turns up, or at the funerals, It's sweetness is the reason why I like it, it helps when there are visitors at home, funerals, in everything it helps." |
| | "It is not like other types of cassava which are bitter which you cannot eat raw when you are hungry. For malukua you can eat it raw, or boil it for taking tea and even if you cook when it has not matured properly it doesn't become watery like other types." |
| Big roots | "It grows big roots and makes food more, you sale, me even eating I don't eat it, I cut out flour. If you are to sell it, it is in plenty." |
| | "The cassava tubers are big." |
| | "Its cassava is huge, hence much flour." |
| | "You find it has produced a great deal with heavy pieces where you'd really struggle to dig it out of the soil. That's a good thing about it." |
| | "I like the black gamente because it produces big tubers, even though it is fibrous but it has big tubers" |
| | "Going to harvest it and getting a very big tuber gives me happiness, it motivates me, it motivates me to increase on the size of land for cassava production, since variety has yielded well in my hand, I should continue with the variety and produce more next time" |

**Main attributes of an "ideal" cassava variety for fresh consumption**

In addition to asking the farmers what they prefer about the current cassava varieties they were growing, we also inquired the attributes that farmers wished an "ideal" cassava variety for fresh consumption could have. Findings from this study revealed that the predominant attributes about an "ideal" cassava variety were non-bitter roots, roots that become soft upon cooking, many roots, early maturity, disease resistance, tasty and long in-ground storage ability of the roots (Table 1.6). All these attributes are a subset of what was

listed in Table 1.1 which is about the predominant cassava attributes preferred about the various varieties that the farmers cultivate attributes. The percentage of farmers preferring a given attribute was primarily significantly different (*P-value* < 0.05) across districts (Table 1.6) although preference for some attributes such as softness and non-bitter roots was not significantly different (*P-value* > 0.05) across districts. On disaggregating the farmers by sex, apart from the percentage of farmers that preferred disease resistant varieties, which was significantly higher for men than women, the preference for the other attributes was not associated with sex (Table 1.6). Similarly, there were no differences in cassava attribute preference associated with age category except for softness attribute, the percentage of younger farmers (81.6%) preferring softness was significantly higher (*P-value* < 0.05) that of older farmers (Table 1.6).

We also investigated whether farmers in different sub-counties within a given district had different preferences for cassava attributes. There were no significant differences for the ideal attributes preferred by farmers in two sub-counties with Kibaale district. Similarly, the percentages of farmers preferring most attributes did not vary significantly for sub-counties in Apac, Arua and Masindi except for preferences for bigger root size in Apac, long in-ground storage of roots for Arua, and non-bitter roots in Masindi district (Table 1.7). Within a given district, cassava attributes preferred by men and women were not significantly different in all districts except in Apac district where only women mentioned disease resistance (Table 1.8). Similarly, age category did not significantly influence the cassava attributes preferred within a given district with an exception of preference for softness of cooked cassava roots in Kibaale district where the younger farmers preferred soft roots more than the older farmers (Table 1.9).

Table 1.6. Predominant cassava attributes reported by smallholder farmers from different locations to be desirable in an ideal cassava variety suitable for fresh root consumption.

| | District | | | | | Sex | | | Age-group | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Cassava attribute | Apac | Arua | Kibaale | Masindi | *P-value* | Men | Women | *P-value* | Aged 18-30 years | Aged 31-70 years | *P-value* |
| Not bitter | 66.7 | 68.8 | 51.1 | 46.9 | 0.167 | 56.3 | 58.9 | 0.755 | 67.3 | 51.7 | 0.066 |
| Soft | 36.4 | 46.9 | 57.4 | 40.6 | 0.254 | 43.7 | 49.3 | 0.497 | 65.5 | 34.8 | **0.000** |
| Many roots | 42.4 | 18.8 | 48.9 | 62.5 | **0.004** | 42.3 | 45.2 | 0.721 | 38.2 | 47.2 | 0.290 |
| Tasty | 9.1 | 3.1 | 70.2 | 31.3 | **0.000** | 31.0 | 34.2 | 0.677 | 41.8 | 27.0 | 0.065 |
| Matures fast | 12.1 | 18.8 | 38.3 | 43.8 | **0.009** | 29.6 | 28.8 | 0.915 | 30.9 | 28.1 | 0.718 |
| Disease resistant | 21.2 | 12.5 | 38.3 | 37.5 | **0.039** | 36.6 | 20.5 | **0.033** | 30.9 | 27.0 | 0.610 |
| Long in-ground storage | 6.1 | 21.9 | 29.8 | 37.5 | **0.020** | 28.2 | 20.5 | 0.286 | 16.4 | 29.2 | 0.081 |
| Non-fibrous roots | 12.1 | 12.5 | 42.6 | 18.8 | **0.002** | 26.8 | 20.5 | 0.380 | 25.5 | 22.5 | 0.682 |
| Big roots | 12.1 | 6.3 | 27.7 | 15.6 | **0.069** | 18.3 | 15.1 | 0.602 | 20.0 | 14.6 | 0.399 |

Table 1.7. Desirable attributes for an ideal cassava variety disaggregated by district and sub-county.

| | Apac | | | Arua | | | Kibaale | | | Masindi | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cassava attribute | Aduku | Apac | *P-value* | Logiri | Ogoko | *P-value* | Nkooko | Nyamarunda | *P-value* | Miirya | Pakanyi | *P-value* |
| Not bitter | 60.0 | 72.2 | 0.458 | 68.8 | 68.8 | 1.000 | 37.5 | 65.5 | 0.057 | 73.3 | 23.5 | **0.005** |
| Soft | 40.0 | 33.3 | 0.692 | 31.3 | 62.5 | 0.077 | 54.2 | 60.9 | 0.642 | 46.7 | 35.3 | 0.513 |
| Many roots | 33.3 | 50.0 | 0.335 | 31.3 | 6.3 | 0.070 | 50.0 | 47.8 | 0.882 | 53.3 | 70.6 | 0.314 |
| Tasty | 20.0 | 5.6 | 0.206 | 12.1 | 12.5 | 0.365 | 45.8 | 30.4 | 0.278 | 46.7 | 41.2 | 0.755 |
| Matures fast | 26.7 | 16.7 | 0.484 | 6.3 | 18.8 | 0.285 | 45.8 | 30.4 | 0.278 | 33.3 | 41.2 | 0.647 |
| Disease resistant | 6.7 | 11.1 | 0.658 | 6.3 | 0.0 | 0.310 | 75.0 | 65.2 | 0.464 | 33.3 | 29.4 | 0.811 |
| Long in-ground storage | 13.3 | 0.0 | 0.110 | 37.5 | 6.3 | **0.033** | 33.3 | 26.1 | 0.587 | 46.7 | 29.4 | 0.314 |
| Non-fibrous roots | 20.0 | 5.6 | 0.206 | 6.3 | 18.8 | 0.285 | 29.2 | 56.5 | 0.058 | 26.7 | 11.8 | 0.281 |
| Big roots | 26.7 | 0.0 | **0.019** | 6.3 | 6.3 | 1.000 | 41.7 | 13.0 | **0.028** | 13.3 | 17.6 | 0.737 |

Table 1.8. Predominant attributes associated with an ideal cassava for fresh consumption disaggregated by district and sex.

| | Apac | | | Arua | | | Kibaale | | | Masindi | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cassava attribute | Males | Females | *P-value* | Males | Females | *P-value* | Males | Females | *P-value* | Males | Females | *P-value* |
| Not bitter | 66.7 | 66.7 | 1.000 | 61.5 | 73.7 | 0.467 | 45.8 | 56.5 | 0.464 | 56.3 | 37.5 | 0.288 |
| Soft | 27.8 | 46.7 | 0.261 | 61.5 | 36.8 | 0.169 | 50.0 | 65.2 | 0.292 | 37.5 | 43.8 | 0.719 |
| Many roots | 38.9 | 46.7 | 0.653 | 15.4 | 21.1 | 0.687 | 50.0 | 47.8 | 0.882 | 56.3 | 68.8 | 0.465 |
| Tasty | 5.6 | 20.0 | 0.206 | 30.8 | 10.5 | 0.150 | 37.5 | 39.1 | 0.908 | 43.8 | 43.8 | 1.000 |
| Matures fast | 27.8 | 13.3 | 0.312 | 15.4 | 10.5 | 0.683 | 45.8 | 30.4 | 0.278 | 50.0 | 25.0 | 0.144 |
| Disease resistant | 0.0 | 20.0 | **0.047** | 0.0 | 5.3 | 0.401 | 66.7 | 73.9 | 0.587 | 37.5 | 25.0 | 0.446 |
| Long in-ground storage | 11.1 | 0.0 | 0.183 | 15.4 | 26.3 | 0.463 | 33.3 | 26.1 | 0.587 | 50.0 | 25.0 | 0.144 |
| Non-fibrous roots | 16.7 | 6.7 | 0.381 | 15.4 | 10.5 | 0.683 | 41.7 | 43.5 | 0.900 | 25.0 | 12.5 | 0.365 |
| Big roots | 22.2 | 0.0 | 0.051 | 7.7 | 5.3 | 0.780 | 25.0 | 30.4 | 0.677 | 12.5 | 18.8 | 0.626 |

Table 1.9. Predominant attributes associated with an ideal cassava for fresh consumption disaggregated by district and age-group.

| | Apac | | | Arua | | | Kibaale | | | Masindi | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cassava attribute | Aged 18-30 years | Aged 31-70 years | *P-value* | Aged 18-30 years | Aged 31-70 years | *P-value* | Aged 18-30 years | Aged 31-70 years | *P-value* | Aged 18-30 years | Aged 31-70 years | *P-value* |
| Not bitter | 84.6 | 55.0 | 0.078 | 77.8 | 65.2 | 0.491 | 61.9 | 42.3 | 0.181 | 50.0 | 45.0 | 0.784 |
| Soft | 53.8 | 25.0 | 0.092 | 66.7 | 39.1 | 0.160 | 81.6 | 38.5 | **0.003** | 50.0 | 35.0 | 0.403 |
| Many roots | 38.5 | 45.0 | 0.710 | 22.2 | 17.4 | 0.753 | 42.9 | 53.8 | 0.454 | 41.7 | 75.0 | 0.050 |
| Tasty | 23.1 | 5.0 | 0.120 | 33.3 | 13.0 | 0.186 | 23.8 | 50.0 | 0.066 | 50.0 | 40.0 | 0.581 |
| Matures fast | 15.4 | 25.0 | 0.509 | 22.2 | 8.7 | 0.298 | 38.1 | 38.5 | 0.980 | 41.7 | 35.0 | 0.706 |
| Disease resistant | 15.4 | 5.0 | 0.311 | 11.1 | 0.0 | 0.104 | 71.4 | 69.2 | 0.870 | 41.7 | 25.0 | 0.325 |
| Long in-ground storage | 0.0 | 10.0 | 0.239 | 22.2 | 21.7 | 0.976 | 19.0 | 38.5 | 0.148 | 25.0 | 45.0 | 0.258 |
| Non-fibrous roots | 15.4 | 10.0 | 0.643 | 0.0 | 17.4 | 0.181 | 52.4 | 34.6 | 0.221 | 8.3 | 25.0 | 0.242 |
| Big roots | 7.7 | 15.0 | 0.530 | 11.1 | 4.3 | 0.477 | 28.6 | 26.9 | 0.900 | 25.0 | 10.0 | 0.258 |

We also asked respondents to describe the different forms in which they consume cassava. Farmers reported consuming cassava in 15 different forms, and of these forms boiled and *Kawunga* are the most common in every district (Table 1.10). Based on the description of the form in which cassava is cooked (Supplementary Table 1), as expected there are two broad categories in which cassava is consumed: freshly harvested roots and products made from cassava flour. In all the four districts, the proportion (> 61 %) of cassava consumed from freshly harvested roots was higher than the proportion (< 39 %) of cassava from flour (Table 1.10). The key attributes about the best variety for the given food products in which they consume cassava significantly varied ($P < 0.0001$) across locations.

Table 1.10. Different forms in which cassava is consumed across districts.

| Form in which cassava is consumed | Product used | Proportion per district (%) | | | |
|---|---|---|---|---|---|
| | | Apac | Arua | Kibaale | Masindi |
| Boiled | Fresh roots | 33.3 | 34.5 | 60.8 | 51.6 |
| Roasted | Fresh roots | 15.3 | 17 | 4.4 | 11.2 |
| Raw | Fresh roots | 2.7 | 4.1 | 0.5 | 0 |
| Deep fried | Fresh roots | 11.3 | 6.2 | 8.8 | 11.2 |
| *Kahunga* | Flour | 22 | 21.1 | 21.1 | 26.1 |
| Porridge | Flour | 2.7 | 10.8 | 3.9 | 0 |
| Alcohol | Flour | 12.7 | 4.6 | 0 | 0 |
| Pancakes | Flour | 0 | 1.5 | 0.5 | 0 |

**DISCUSSION**

As breeding programs endeavor to contribute towards ensuring food security and alleviating poverty levels, especially among resource-poor farmers, it is a prerequisite for them to gain a proper understanding of the needs and preferences of men and women who are the target end-users of the products of breeding (Acheampong et al., 2018). Previous studies have noted that plant breeding is not a neutral approach, heterogeneous groups of farmers may require different types of varieties in order to meet their needs which could depend on a wide range of factors, for example availability of resources and the intended use of the varieties (Snapp & Pound, 2017). The most frequently mentioned cassava attributes preferred by smallholder farmers were yield (many roots and big root size), early maturity, and root quality (taste, long in-ground storability of roots, softness of cooked roots, and non-bitter roots) attributes. The implication of this finding is that breeders need to pay attention to these major attributes to enhance chances of adoption of the new varieties they breed for instance, the breeder may decide to attach varying weights on each of these traits and use the weights

to develop a selection index to be used during the breeding process. Although previous studies also indicated that some of the cassava attributes highlighted in this study such as early maturity, long in-ground storability of roots, and cooking quality influence selection and preservation of the cassava varieties for usage in the subsequent seasons (Tumuhimbise et al., 2012; Nakabonge et al., 2018), none of them presented their results differentiated by sex and age. There were significant differences in proportion of attributes preferred across districts, sex and age categories, indicating that gender and age can be important factors in preferences of cassava attributes, and can be geographically and culturally defined. Ideally, a variety that is liked across a wide range of locations would be desirable, but in scenarios where this is not achievable, breeders may need to develop diverse varieties to meet the diverse needs and preferences of the end-users although this approach require additional resources in terms of land, labor, time and financial resources. Our results indicate that the main cassava attributes liked in one location may not necessarily be the top preferred attributes in another location. Thus, breeders may focus on different varieties to meet the varying needs of end-users in different locations. Considering the top five attributes mentioned by farmers within each location, many number of roots and early maturity were listed in all the four districts, thus we would consider these as some of the non-negotiable attributes to consider during variety development. Softness of cooked cassava roots was also among the most predominantly attributes within all the districts except for Arua. One of this likely explanation for this kind of observation could be the differential forms in which cassava is commonly consumed within the different districts. From the focus group discussions, farmers in Arua largely mentioned that cassava is mainly consumed after processing cassava roots into flour, whereas in other districts, farmers mentioned that consumption of food made from fresh cassava roots was the most predominant form of consumption. Thus, for people who do not cook cassava roots, softness of cooked cassava

roots is not an important attribute, despite it being a key attribute for people who eat cassava in the boiled form. Interestingly, Kibaale where disease resistance appeared among the top five attributes with highest percentage preference, this could partially stem from the fact that Kibaale district had relatively high disease incidences compared to other districts. Similarly, there was noticeably lower proportion for long in-ground storage of cassava roots in Apac relative to other districts, this could probably be that most of the varieties they were growing did not have this attribute since the question asked was to name the attributes they liked about the cassava varieties they were currently growing.

There were also significant (*P-value* < 0.05) sex differentiated cassava attribute preferences observed in different districts for instance in Apac district, the proportion of liking cassava varieties with many roots and early maturity was significantly higher for men than women. This observation could probably be a result of men being more market-oriented and interested in making profits that come from extra yields and fast-maturing varieties. In most communities decisions regarding sale of the agricultural produce are determined by men. Differentiated cassava attribute preferences may stem from differentiated division of labor. Kabeer (2015) noted that social relations such as sex, class, age, ethnicity, and race can influence the norms and culture of that society. The observed differences among cassava attributes preferred based on age and sex among farmers in different districts could stem from the different social relations in the various localities, for example, we observed that some cassava attributes were associated with age-group for example in Arua district, younger farmers preferred early maturing varieties, while older farmers did less so. This observation could probably be due to the amount of resources owned by younger farmers relative to older farmers. In most cases, the younger farmers own fewer resources in terms of land, labor and capital than older farmers and thus may like more early maturing varieties in order to be able to get food for their families and earn income from the sale of cassava to be able to meet

other household needs. Given that, even within one district, we see different preferences amongst the varying social groups of smallholder farmers, there is need to provide diverse products to meet the needs and preferences of the various social categories of people.

## CONCLUSION

The work reported in this study portrays the importance of considering gender in identifying the needs of end-users during the design of breeding process. There is need to do a gender-based analysis of the needs and preferences of the target end-users. Farmers prefer multiple attributes and this may call for trait prioritization during variety development. To do this, different approaches could be used such as use of selection index, and using farmer participatory breeding approaches. Prioritizing attributes that suit the need and preferences of end-users may increase the likelihood of the new varieties being adopted by the target end-users. Our results also indicated that the majority of farmers eat cassava in fresh form, thus, there is more need to focus on the desirable attributes associated with each form in which cassava is consumed.

**REFERENCES**

Acheampong, P.P., V. Owusu, and G. Nurah. 2018. How does Farmer Preference matter in Crop variety Adoption? The case of Improved Cassava varieties' Adoption in Ghana. *Open Agriculture* 3: 466–477. doi: 10.1515/opag-2018-0052.

Acquaah, G. 2012. Principles of plant genetics and breeding. Second Edi. John Wiley and Sons, Inc., West Sussex, UK.

Alene, A.D., R. Khataza, C. Chibwana, P. Ntawuruhunga, and C. Moyo. 2013. Economic impacts of cassava research and extension in Malawi and Zambia. *Journal of Development and Agricultural Economics* 5(11): 457–469. doi: 10.5897/JDAE2013.0496.

Bechoff, A., K. Tomlins, G. Fliedel, L.A. Becerra, A. Westby, C. Hershey, D. Dufour, A. Bechoff, K. Tomlins, G. Fliedel, L.A. Becerra, C. Hershey, and D. Dufour. 2016. Cassava traits and end-user preference : Relating traits to consumer liking , sensory perception , and genetics. *Critical Reviews in Food Science and Nutrition* 0(0): 1–21. doi: 10.1080/10408398.2016.1202888.

Bezner Kerr, R. 2017. Gender and Agrarian inequalities. p. 333–370. *In* Snapp, S., Pound, B. (eds.), Agricultural Systems: Agroecology and Rural Innovations for Development. Second Edi. San Diego, CA, USA.

Cardoso, A.P., E. Mirione, M. Ernesto, F. Massaza, J. Cliff, M. Rezaul Haque, and J.H. Bradbury. 2005. Processing of cassava roots to remove cyanogens. *Journal of Food Composition and Analysis* 18: 451–460. doi: 10.1016/j.jfca.2004.04.002.

Chiwona-Karltun, L., J. Mkumbira, J. Saka, M. Bovin, N.M. Mahungu, and H. Rosling. 1998.

The importance of being bitter — a qualitative study on cassava cultivar preference in Malawi. *Ecology of Food and Nutrition* 37(3): 219–245. doi: 10.1080/03670244.1998.9991546.

Chiwona-Karltun, L., D. Nyirenda, C.N. Mwansa, J.E. Kongor, L. Brimer, S. Haggblade, and E.O. Afoakwa. 2015. Farmer preference, utilization, and biochemical composition of improved Cassava (Manihot esculenta Crantz) varieties in Southeastern Africa. *Economic Botany* 69(1): 42–56. doi: 10.1007/s12231-015-9298-7.

Doss, C.R., and M.L. Morris. 2001. How does gender affect the adoption of agricultural innovations? The case of improved maize technology in Ghana. *Agricultural Economics* 25(2001): 27–39.

El-Sharkawy, M.A. 2004. Cassava biology and physiology. *Plant Molecular Biology* 56: 481–501.

FAO. 2013a. Save and grow cassava: A guide to sustainable production intestification. Rome.

FAO. 2013b. STANDARD FOR SWEET CASSAVA: CODEX STAN 238-2003.

Forsythe, L., H. Posthumus, and A. Martin. 2016. A crop of one's own? Women's experiences of cassava commercialization in Nigeria and Malawi. *Gender, Agriculture and Food security* 1(2): 110–128.

JHSPH. 2009. Johns Hopkins School of Public Health (JHSPH) (2009) Human Subjects Research Ethics Field Training Guide.

Kawuki, R.., T. Amuge, E. Nuwamanya, G. Semakula, S. Tumwesigye, A. Bua, Y. Baguma, C. Omongo, T. Alicai, and J. Orone. 2011. A breeding scheme for local adoption of cassava (Manihot esculenta Crantz). *Plant Breeding and Crop Science* 3(7): 120–130.

Kizito, B.E., L. Chiwona-Karltun, T. Egwang, M. Fregene, and A. Wasterbergh. 2007. Genetic diversity and variety composition of cassava on small- scale farms in Uganda : an interdisciplinary study using genetic markers and farmer interviews. *Genetica* 130: 301–318. doi: 10.1007/s10709-006-9107-4.

Kombo, G.R., A. Dansi, L.Y. Loko, G.C. Orkwor, R. Vodouhè, P. Assogba, and J.M. Magema. 2012. Diversity of cassava (Manihot esculenta Crantz) cultivars and its management in the department of Bouenza in the Republic of Congo. *Genetic Resources and Crop Evolution* 59(8): 1789–1803. doi: 10.1007/s10722-012-9803-0.

Lebot, V. 2009. Tropical root and tuber crops: cassava, sweet potato, yams and aroids. Crop production science in horticulture (17). CABI, Wallingford, UK.

Masamha, B., V. Thebe, and V.N.E. Uzokwe. 2018. Mapping cassava food value chains in Tanzania's smallholder farming sector: The implications of intra-household gender dynamics. *Journal of Rural Studies* 58(2018): 82–92. doi: 10.1016/j.jrurstud.2017.12.011.

Montagnac, J.A., C.R. Davis, and S.A. Tanumihardjo. 2009. Processing techniques to reduce toxicity and antinutrients of Cassava for use as a staple food. *Comprehensive Reviews in Food Science and Food Safety* 8(1): 17–27. doi: 10.1111/j.1541-4337.2008.00064.x.

Nakabonge, G., C. Samukoya, and Y. Baguma. 2018. Local varieties of cassava: Conservation, cultivation and use in Uganda. *Environment, Development and Sustainability* 20: 2427–2445. doi: 10.1007/s10668-017-9997-6.

Nweke, I.F., S.C.D. Spencer, and K.J. Lynam. 2002. The cassava transformation: Africa's best kept secret. Michigan State University Press, East Lansing, Michigan.

Ogunlela, Y.I., and A.A. Mukhtar. 2009. Gender Issues in Agriculture and Rural Development in Nigeria : The Role of Women. *Humanity and Social Sciences* 4(1): 19–30.

Otim-Nape, G.W., A. Bua, G. Ssemakula, G. Acola, Y. Baguma, S. Ogwal, and R. Van der Grift. 2005. Cassava development in Uganda. A country case study towards a global cassava development strategy. *In* A review of cassava in Africa with country case studies in Nigeria, Ghana, the United Republic of Tanzania, Uganda and Benin. Rome.

R Core Team. 2018. A language and environment for statistical computing. R foundation for statistical computing. Vienna, Austria.

Silver, D. 2019. Alienation in a four factor world. *Journal for the Theory of Social Behaviour* 49: 84–105. doi: 10.1111/jtsb.12192.

Teeken, B., O. Olaosebikan, J. Haleegoah, E. Oladejo, T. Madu, A. Bello, E. Parkes, C. Egesi, P. Kulakow, H. Kirscht, and H.A. Tufan. 2018. Cassava Trait Preferences of Men and Women Farmers in Nigeria: Implications for Breeding. *Economic Botany*: 1–15. doi: 10.1007/s12231-018-9421-7.

Tumuhimbise, R., R. Melis, P. Shanahan, and R. Kawuki. 2012. Farmers' Perceptions on Early Storage Root Bulking in Cassava (Manihot esculenta Crantz) in East and Central Uganda and their Implication for Cassava Breeding. *World Journal of Agricultural Sciences* 8(4): 403–408. doi: 10.5829/idosi.wjas.2012.8.4.1667.

Turyagyenda, L.F., E.B. Kizito, M.. Ferguson, Y. Baguma, P. Gibson, B.W. Wanjala, and D.S.. Osiru. 2012. Genetic diversity among farmer-preferred cassava landraces in Uganda. *African crop science journal* 20(1): 15–30. doi: 10.4314/acsj.v20i1.

UBOS. 2010. Uganda census of agriculture: Crop area and production report. Kampala, Uganda.

SUPPLEMENTARY MATERIAL

**Supplementary Table 1: Description of how various food types are prepared**

| Food form | Quotations from farmers regarding food preparation |
|---|---|
| Boiling | "After harvesting the roots, you remove the hard coat (peeling), and wash it. You put it in the saucepan, add one cup of water! You look for some leaves such as banana leaves to cover it, cut it and wash it and then cover it. Then you cover it with another sauce pan or with another clean utensil. You put it on fire to boil for about 30 – 45 minutes. After boiling you can remove it from the fire and put it down. After you serve it on plates and serve people with cups of water to consume it" (FGD notes).<br><br>"You harvest your soft cassava, remove the peelings, remove some starch, cut it into small pieces, wash it, put it in water and add the beans, you put on the fire place like for thirty minutes and then remove it from the fire place. You get a mingling stick and begin pressing it until it is ready. You press it until it is ready and then serve and eat. If you have greens and bitter berries, you can press them with the cassava"<br><br>"Go get cassava, peel, scratch it well, after I wash and start chopping into small pieces. There are there like of the same small size. Have already prepared beans, but sometimes you don't prepare and make them so ready. I get my cassava that is washed very well and I add beans with salt. I put on fire, cover with a banana leaf and they get ready, but you wait until soup dries completely".<br><br>"You go and harvest it, after harvesting, you remove the hard coat, and you chop |

| | |
|---|---|
| | it into small pieces like chips wash it. You boil it for about 30 minutes, after boiling you can remove the excess water! After which you start mingling it using the mingling stick or at times you pound it using a mortar! Or you continually mingle it like as if you are mingling ordinary *Kawunga*. Or you continuously use the mortar until it becomes so soft like *Kawunga*! After pounding it you serve it on plate and serve sauce, and people consume it as food!" (FGD notes). |
| Roasting | "You harvest and put it on fire. It is first peeled, others are roasted without peeling. It takes like half an hour. After roasting you remove it from the fire then peel it with knife and put on a plate. You can eat it with salt but if you don't want it is your wish. You start to consume it with the children" (FGD notes). |
| Deep frying | "You harvest your cassava, bring it and cut off the stem, some people first wash it before peeling it if it is muddy, you remove some starch, and you cut it into pieces and get where to put it. You get a saucepan to use for frying and then begin frying. When it is ready, you get it and put in the bucket. You get to know it is ready when it has changed to yellow color" (FGD notes). |
| Cassava porridge | "You go and harvest cassava, after harvesting, you remove the hard coat, and you put it to ferment for about one to two days depending on the sunshine! Then you remove the fermented coat, after removing the coat, you crush it into smaller particles. After which you spread it on the carpet to dry in the sun! After drying, you can get in small quantity and pound using a mortar and sieve it using a sieve, or take to be milled with a milling machine to become flour! After which you put water on fire to boil and you get the flour and mix it in a clean |

| | |
|---|---|
| | utensil! After which you pour in the boiling water and stir it! If it is too thick, you add more water to make it better! After it has boiled, you add sugar in it and stir! After which you serve it in cups and serve it and put it before people!" (FGD notes). |
| | "I go to the garden and harvest the cassava, I bring it and peel it, cut it into small pieces, I take it to the rock for dying, after drying, I take it to the grinding machine or I pound using a mortar and pestle. After grinding it, I get a saucepan, put water and I boil it, after it is ready, I get the flour and mix it with cold water and then put the mixture in the hot water on the fire place and I mix. After that I bring flour ground from germinated sorghum and I add there. I then keep it for one day and then I begin taking" (FGD notes). |
| Alcohol | "You go and harvest it, after harvesting, you remove the hard coat, and you put it to ferment for about 2 – 3 days! Then you remove the fermented coat, after removing the coat, you crush it into smaller particles. After which you spread it on the carpet to dry in the sun! After drying, you go and fetch water and pour in the pot! You pour it in the pot, you mix it. You cover it with a saucepan, so that it can ferment for about 3 days! You fry it in a frying pan called "Kalaya"! After frying, you again pour it back in cold water, add millet, mill it and add to it! The millet helps it to brew for about 2 – 3 days, on the 3rd day, you drill a hole under the pot, get another big pot with water! After which you connect both pots with a pipe from one pot through the pot with cold water which is a coolant to the bottle, you add fire to the pot with brewed alcohol, so that it is distilled through the pipe into the bottle until it is filled! After which you start selling it! (FGD notes). |

| | |
|---|---|
| Raw | "You go and harvest it, after harvesting, you remove the hard coat, wash it and you chop it into small slices, put it on the plate for people to consume!" |
| Pancakes | "You go and harvest it, after harvesting, you remove the hard coat, and wash it. After which I get a piece of an iron sheet, drill some holes through it, then get the cassava that is already washed and continuously grill on the iron sheet, until it collects in the utensil. I then squeeze out the water from it and I mix it as dough with my hands! I go and purchase cooking oil and pour it in the sauce pan and put it on fire! It boils, after boiling it, I roll it into balls! I drop them in the oil to boil! It takes around 15 minutes, after changing color; I get a spoon to remove the one which are ready because I cannot dip my hands in the hot cooking oil! I pick the ready ones and put them on clean plate! I put them in the basin, I then count them! I give my wife to take it to the market!" (FGD notes). |
| *Kawunga* | "You go and harvest it, after harvesting, you remove the hard coat, and you put it to ferment for about 3 – 4 days! Then you remove the fermented coat, after removing the coat, you crush it into smaller particles. After which you spread it on the carpet to dry in the sun! After drying, you can pound it, or take to be milled with a milling machine to become flour! After which you mingle it as bread! You put water on fire! The water boils! After boiling you get some little flour and sprinkle it in the hot water to assess whether the water is ready! One the flour that you sprinkled clears, then you can pour flour into the water in large quantity, mingle it into bread!" |

**Supplementary Files**

Supplementary File 1: Site selection survey for research on gendered cassava trait preferences

Supplementary File 2: Guide for selecting the sub-county

Supplementary File 3: Scoping study: Key informant guide for village leaders

Supplementary File 4: Individual interview questionnaire

Supplementary File 5: Focus group discussion guide

Supplementary File 6: Coding frame for cassava attributes from survey in Uganda, 2015

# CHAPTER 2 ESTIMATES FOR HERITABILITY AND CONSUMER-VALIDATION OF A PENETROMETER METHOD FOR PHENOTYPING SOFTNESS OF COOKED CASSAVA ROOTS[1]


## ABSTRACT

Although breeders have made significant progress in the genetic improvement of cassava (*Manihot esculenta* Crantz) for agronomic attributes its, lack of information on heritability and limited testing of high-throughput phenotyping methods are major limitations to improving root quality attributes, such as softness after cooking, which rank high among Ugandan consumers. The objectives of this study were to determine heritability for softness of cooked cassava roots, and quantify the relationship between penetrometer and consumer testing methods for phenotyping softness of cassava roots. Softness defined as the maximum force (N) needed to penetrate cooked root samples using a penetrometer, was evaluated at four cooking time intervals: 15, 30, 45, and 60 min on 268 cassava genotypes. Estimates of broad-sense heritability (repeatability) ranged from 0.17 to 0.37, with the highest value observed at 45 min of cooking time interval. In the second study involving 135 cassava consumers from Kibaale district in Uganda, penetrometer measurements of cooked roots from six cassava varieties were found to be in strong agreement ($r^2$ = 0.91; *P-value* = 0.003) with ordinal scores of root softness from consumer testing. These results suggest that: (a) softness of cooked cassava roots is an attribute amenable for evaluation and selection; and (b) a penetrometer can readily be used for assessment of cooked root softness. These findings



[1]Iragaba, P., E. Nuwamanya, E. Wembabazi, Y. Baguma, D. Dufour, E.D. Earle, R. Bezner Kerr, H.A. Tufan, M.A. Gore, and R.S. Kawuki. 2019. Estimates for heritability and consumer-validation of a penetrometer method for phenotyping softness of cooked cassava roots. *African Crop Science Journal* 27(2): 147–163. https://doi.org/10.4314/acsj.v27i2.3.

form the basis for operationalizing the routine assessment of root softness in cassava breeding trials, an output that will enhance ongoing efforts to breed for desired end-user root quality attributes.

*Key Words*: Breeding, cooking, *Manihot esculenta*

## INTRODUCTION

Cassava (*Manihot esculenta* Crantz) is an important source of food for about 800 million people worldwide (FAO and IFAD, 2000; Lebot, 2009; FAO, 2013). With a high diversity of end-users, cassava varieties that are released without a focus on product quality may not meet end-user trait preferences and are thus poorly adopted by farmers (Chiwona-Karltun *et al.*, 2015; Bechoff *et al.*, 2018a). In a variety adoption study conducted in Malawi, only 7% of improved cassava varieties tolerant to African cassava mosaic disease were adopted, simply because the new varieties lacked consumption quality attributes desired by end-users (Alene *et al.*, 2013). In a related study conducted in the Republic of Congo, farmers abandoned varieties that did not meet both their agronomic and culinary trait preferences (Kombo *et al.*, 2012). These results highlight the critical role of end-users in variety development and adoption.

In Uganda, attributes such as high yield and cooking quality justify farmers' variety choices and hence historic cultivation of cassava varieties (Nakabonge *et al.*, 2017). End-user attributes such as softness of cooked cassava roots have been reported to be highly preferred by consumers (Ngeve, 2003; Padonou *et al.*, 2005). As a top attribute emerging from consumer surveys, we consider softness as an important breeding criterion to respond to the needs of cassava farmers and consumers in Uganda.

One major limitation to breeding end-user culinary root quality attributes such as cassava root softness is the lack of reliable, low-cost and accurate high-throughput methods for assessing root softness in clones under evaluation. Consequently, consumer acceptability testing, which involve assessing the perception of consumers regarding the new products, is used as a fallback by most breeding programs as a method to determine culinary root quality attributes (Bechoff *et al.*, 2018b). Unfortunately, consumer acceptability testing of culinary

root quality attributes is not ideal in early selection stages that are often associated with a large number of entries (IITA, 1990; Ceballos *et al.*, 2004). Moreover, consumer acceptability studies require significant amounts of resources in terms of time, funds and personnel, further complicating the selection process. It is for this reason that most culinary root quality attributes are scored later in the selection process, using a group of consumers that is tasked to evaluate only a few advanced breeding lines (Ceballos *et al.*, 2004). A major disadvantage of this approach may be the unintended loss of genetic variation for culinary root quality attributes during the early selection process (Ceballos *et al.*, 2004; Bernardo, 2010). These shortfalls justify the urgent need to develop alternative phenotyping methods which can be used on a large number of lines and are well correlated with consumer acceptance.

Most of the previous studies on softness of cooked cassava roots have focused on food quality and with few samples i.e., ($n$ <25) (Eggleston and Asiedu, 1994; Ngeve, 2003; Padonou *et al.*, 2005; Sajeev *et al.*, 2010). None of these studies estimated heritability of cassava root softness. Information on heritability estimates (in this case for root softness) helps develop and standardize methodologies to be used for routine assessment, and consequently selection; maximizing heritability readily translates to increased genetic gain (Bernardo, 2010; Acquaah, 2012). Thus, the objectives of this study were to determine broad-sense heritability (or repeatability) for softness of cooked cassava roots, and to quantify the relationship between penetrometer and consumer testing methods for phenotyping softness of cassava roots.

## MATERIALS AND METHODS

### Genetic materials

In order to examine the extent of heritability (repeatability) of root softness in cassava, we assembled 285 genotypes from two sources in Uganda, namely smallholder farmers residing

in northern, western and west Nile region, and from the cassava breeding populations developed by National Crops Resources Research Institute (NaCRRI). The names of genotypes cultivated by smallholder farmers were recorded during the end-user cassava trait preference survey.

The study plan for this survey and consent forms were reviewed by the Institutional Review Board (IRB) of Cornell University and approved (IRB ID 1502005316). After the survey, we collected stem cuttings from each of the named varieties at the sub-county level, forming a total of 76 variety names. The list of the variety names can be accessed at ftp:// ftp.cassavabase.org/manuscripts/Iragaba_et_al_2019_quality/Phenotype_infos/ (in a file named: "Farmer varieties included in the study to evaluate genetic variability of softness of cooked cassava roots.txt"). The stem cuttings were collected from three different plants for each of the 76 varieties. We treated each plant as a unique genotype, because earlier studies had shown that not all varieties with the same name are genetically identical (Rabbi *et al.*, 2015). In total, we obtained 223 genotypes from the farmers' fields.

To complement the farmer varieties, we selected 78 genotypes sampled from a diversity panel of 635 genotypes that were generated by crossing germplasm from International Institute for Tropical Agriculture (IITA) with that sourced from the International Center for Tropical Agriculture (CIAT) (Kayondo *et al.*, 2018).

To sample the extent of phenotypic variation for two critical agronomic attributes, harvest index (HI) and dry matter content (DMC), we constructed a scatter plot (HI against DMC) and divided the resultant scatter plot into four quadrants. We selected a nearly equivalent number ranging from 19 to 20 genotypes from each quadrant and thus, identified the representative 78 genotypes (Figure 2.1). The lists of the selected genotypes are stored in CassavaBase: ftp://ftp.cassavabase.org/ manuscripts/Iragaba_et_al_2019_quality/ (in a file

named: “The 78 genotypes were selected from a total of 635 genotypes that comprise the diversity panel maintained at National Crops Resources Research Institute.txt”)

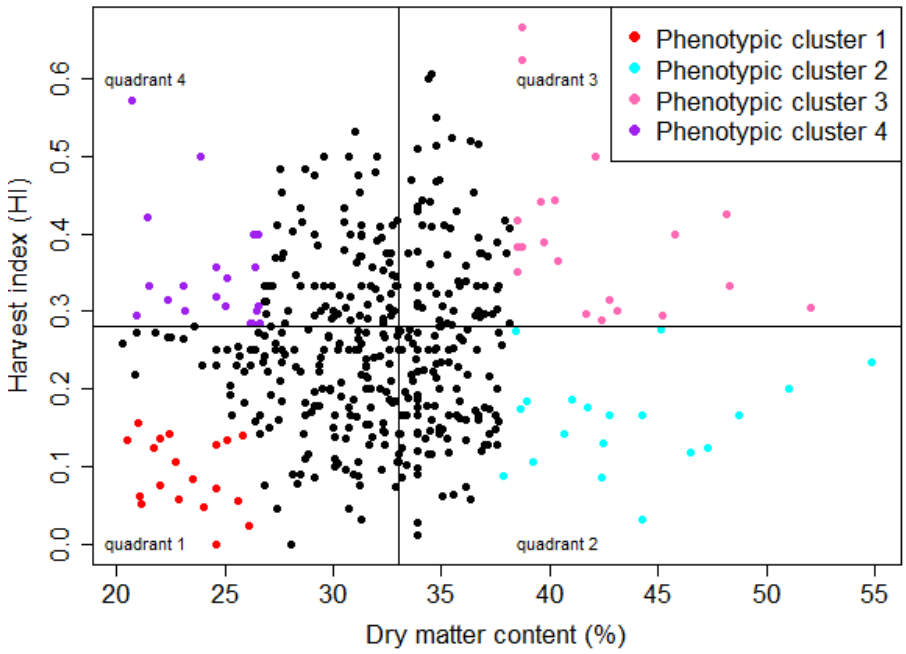



Figure 2.1. Scatter plot of harvest index (HI) against dry matter content (DMC) for 635 genotypes from a cassava diversity panel maintained by National Crops Resources Research Institute (NaCRRI), with the 78 genotypes (colored points) selected to sample the maximal phenotypic diversity.

We then ascertained whether the 78 genotypes selected for maximum variation in HI and DMC represented the extent of genotypic diversity in the panel. To achieve this, we conducted a Principal Component Analysis (PCA), with the *prcomp* function of the *stats* package in R (R Core Team, 2015), using 61,297 single-nucleotide polymorphism (SNP) markers that were previously scored on the 635 genotypes (Kayondo *et al.*, 2018). Then, *K-means* clustering was performed with the *kmeans* function of the *stats* package in R, to divide the 635 genotypes into four genetic clusters. This was followed by intersecting the PCA and *K-means* clustering results to determine which of the four clusters each of the 78 genotypes belonged (Figure 2.2). Taken together, these analyses confirmed that the 78 genotypes

selected for their phenotypic diversity also captured the major patterns of genotypic diversity present in the entire panel (Figure 2.3). However, at the time of collection of planting materials, we obtained 62 genotypes because they were the only ones with sufficient cuttings.

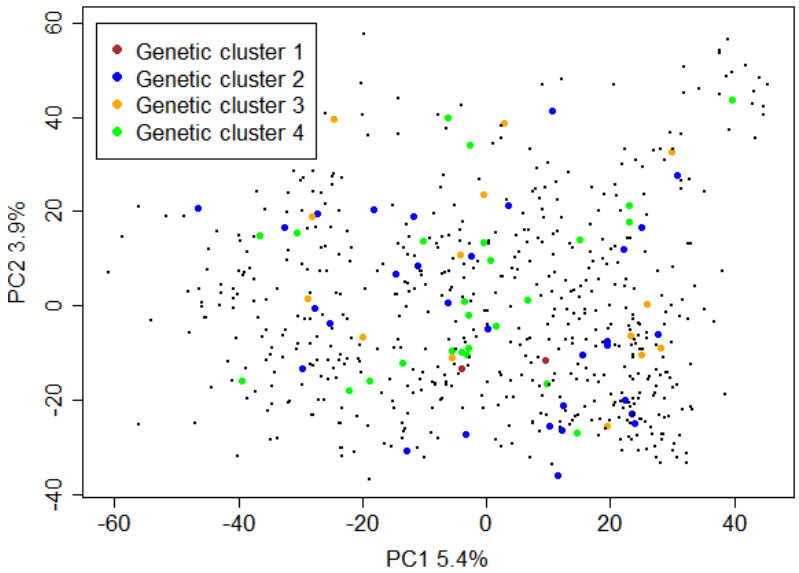



Figure 2.2. Principal Component Analysis (PCA) plot of PC1 and PC2 of the genetic diversity in the NaCRRI population, with only the 78 selected genotypes colored and labeled according to which of the four K-means clusters each genotype belongs. The PCA is based on 61,297 single-nucleotide polymorphism markers scored on 635 cassava genotypes.

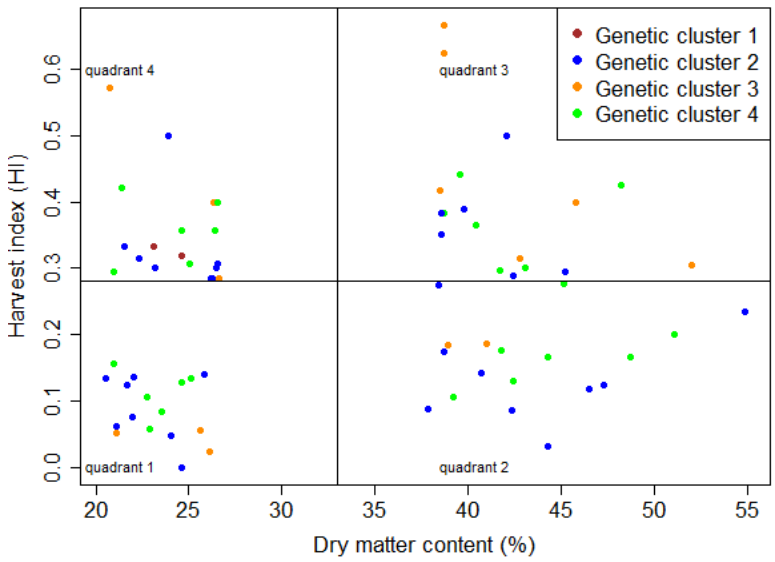


Figure 2.3. Scatter plot of harvest index (HI) against dry matter content (DMC) for the 78 selected genotypes labeled according to their K-means cluster.

**Experimental design**

During the second rains of 2015, the 285 genotypes (223 collected from farmers and 62 from breeding program) were established in the field in an augmented incomplete block design in two sites: Adravu village, Manibe sub-county, Arua district (northwestern Uganda), and Namaswata village, Kassanda sub-county, Mubende district (central Uganda). The altitude (1200 - 1400 masl), average annual rainfall (1250 mm), and average temperature (22 °C) of Arua are higher than those of Mubende (Arua District Local Government, 2012; Mubende District Local Government, 2013). Adravu village has an average temperature of 24.0 °C and annual rainfall of 1385 mm; whereas Namaswata village has an average temperature of 21.5 °C and annual rainfall of 970 mm (Fick and Hijmans, 2017). These two field sites provide contrasting environments for the phenotypic evaluation of the softness phenotype. Arua had 14 incomplete blocks, with 25 plots per incomplete block, while Mubende had 12 incomplete blocks with 30 plots per incomplete block. Each incomplete block was augmented by the addition of three check genotypes namely, NAROCASS 1, NASE 14, and TME 14 grown at the two sites. In addition, NASE 4 was used as a local check in Arua, while UGL15228 was used as the local check in Mubende. Three of these genotypes (NAROCASS 1, NASE 14 and NASE 4) have been officially released in Uganda. Within each incomplete block, the order of entries was randomized.

Experimental units were one-row plots, measuring 5 m in length; each plot was represented by five plants that were evenly spaced and had a between row spacing of 1 m. Each block was separated from the other by an alley of 2 m width.

At maturity (12 months after planting), all plants were uprooted and their roots bulked per plot. Subsequently, four to six intact, uniformly sized roots were randomly sampled from each plot, and appropriately labeled as per the field layout available at www.cassavabase.org. The sampled roots were immediately washed with tap water to remove soil debris, followed by air drying of residual moisture from the root surface for 15-30 min. To prolong the shelf life of cassava roots prior to softness assessment, we waxed the roots on the day of harvest (Lebot, 2009). The roots were waxed by immersing them for five seconds in pre-melted food-grade wax maintained at 140-160 °C, and thereafter, cooled to air temperature outside in the field from where the cassava was harvested.

Root samples could not be obtained for all 285 genotypes across the two locations because plants for some genotypes did not survive up to the time of harvest. Accordingly, data were collected from 236 and 237 genotypes from Arua and Mubende, respectively. This resulted into a total of 268 genotypes that had phenotypic data from at least one of the two environments.

**Penetrometer assessment**

The waxed roots from the two field locations were stored in a -80 °C freezer to avoid deterioration prior to softness evaluation. Freezing was done to ensure long preservation (more than one month), because waxing can only preserve roots for about one month (Lebot, 2009). In a pilot experiment that we conducted earlier with ten genotypes, root softness for frozen root samples over a duration of 30 days was strongly correlated ($r^2$ = 0.76, *P-value* = 0.0009) with softness from fresh roots. This provided confidence in the freezing process adopted to handle the relatively high number of samples analyzed in this study.

Accordingly, samples for a given block were removed from the freezer and thawed at room temperature (approximately 23 °C). After thawing, two roots per plot (genotype) were

peeled and sliced into 3 cm sections using a knife and ruler. For each genotype, four sections were randomly selected from the sliced sections and loosely wrapped in perforated aluminum foil. The selected sections were cooked in a water bath set at a constant near boiling temperature of 90 °C, and softness assessment done at four time points: 15, 30, 45 and 60 min. This was done by removing a root section at each defined time point, and softness assessment done. Thus, each root was assessed for softness at four cooking time points: 15, 30, 45 and 60 min.

Softness was defined as the maximum force used to penetrate the root section using the penetrometer. Thus, at each time point, the 7.9 mm diameter tip of a digital penetrometer (Model number: FHT-1122, Vetus Industrial Company Limited, Hefei, China) was pushed to a depth of 1 cm into each cooked root section. This was done at three different positions of the sectioned root i.e., three technical measurements taken per root section. Hence, we obtained three observations for each time point, per genotype per site, a dataset appropriate for quantifying heritability for root softness in cassava.

**Consumer acceptability testing in Kibaale**

To corroborate the quantitative softness data, particularly the softness assessment by penetrometer, we conducted on-site consumer testing in Kibaale district, one of the districts with high consumption of cassava (Nakabonge *et al*., 2017). The study plan for consumer testing and consent forms were reviewed by IRB of Cornell University and approved (IRB ID 1809008241). Accordingly, root softness for six commonly grown varieties in Kibaale (Bukalasa, Gwalanda, Kyawada, NAROCASS1, NAROCASS 2, and Matooke) were evaluated using both the consumer testing method, by 135 consumers, and by the penetrometer method. For validity of results from consumer acceptability testing, it is

recommended to involve at least 112 consumers (Hough *et al.*, 2006); thus in our study the aim was to work with a minimum of 112 consumers.

The six varieties chosen represented the greatest diversity in softness for commonly grown varieties: Bukalasa and Matooke were classified as "soft", while Gwalanda and Kyawada were described as "hard" varieties by farmers, during the survey held in 2015. We also included two recently released varieties from the NaCRRI breeding program, namely NAROCASS 1 ("soft") and NAROCASS 2 ("hard"). The consumers in this study were men and women who had historically grown and consumed cassava, and thus were considered to be experienced to provide reliable information (Safo-kantanka and Owusu-nipah, 1992; Bechoff *et al.*, 2018b).

Before starting the consumer testing activity, we read a consent statement to the entire group of participants to seek their approval to participate in the study, explained the purpose of the study, methods to be used, and let them know that their participation was voluntary and that they were free to withdraw at any time during the study. Consumers who participated in the study were disaggregated by age and sex. Of the total number of consumers, 51.5% were women (18-30 age group that constituted 25.0%, and 31-70 age group that constituted 26.5%), and 48.5% were men (18-30 age group that constituted 19.4%, and 31-70 age group that constituted 29.1%).

Freshly uprooted cassava roots of the six varieties, harvested at 12 months, were peeled and each cut into 3 cm sections to generate at least 140 root sections per variety. These root sections for each genotype were washed twice in tap water to remove debris and adhering soil particles. Thereafter, the roots were wrapped in banana (*Musa* spp.) leaves, as routinely done when preparing boiled cassava roots in the study area. The six wraps, each representing a given genotype, were carefully placed into a single saucepan with tap water at room

temperature (approximately 23 °C). The top of the saucepan was fully covered with layers of banana leaves and another saucepan was used as a lid; this was followed by cooking of roots over wood fire for 50 min.

After the cassava roots had cooked for 50 min, the roots were withdrawn from the wood fire and assessed for softness. Of the 140 root sections cooked per variety, 110 were used for assessing consumer acceptability by selected consumers, while the remaining 30 were used for measuring softness with the penetrometer. For testing of consumer acceptability, each of the 110 root sections was cut into halves, resulting in a total of 220 root sections for assessment per variety. At the time of consumer testing, each participant was served one root section per variety to score softness, using an ordinal 1-5 scale modified from ISO (1994), where 1 = extremely soft, 2 = soft, 3 = neither soft nor hard, 4 = hard, and 5 = extremely hard. The softness evaluation was based on the disintegration ability of the cooked roots in the mouth of the consumers (Padonou *et al.*, 2005). Each of the six varieties was evaluated once by each of the 135 consumers, thus 135 root sections were used. The remaining unused root sections were discarded at the end of the consumer testing process.

Penetrometer-based root softness assessment (after cooking roots for 50 min) was done on each variety at six post-cooking time points: 0, 15, 30, 45, 60 and 75 min after withdrawing the roots. The first (0 min) and last (75 min) time points corresponded to the initiation and completion of consumer testing in Kibaale, respectively. Measuring softness at the six post-cooking time points throughout the entire period of consumer evaluation allowed us to better account for the influence of cooling on root softness. At each time point, five cooked root sections per variety were randomly sampled from the bulk (in saucepan) and assessed for softness using the penetrometer, as described earlier. Assessments were done on two different

sides of each section, and thereafter, the section was discarded. In total, 360 penetrometer measurements were recorded.

**Statistical analyses**

**Softness quantitative data**

In order to determine heritability of root softness, we jointly analyzed softness data generated from Arua and Mubende trials. Initially, we processed each cooking time point (15, 30, 45 and 60 min) separately, and thus generated four datasets for root softness. For each dataset, the Box-Cox power transformation (Box and Cox, 1964) was conducted to identify the optimal transformation procedure that best corrected for unequal variances and non-normality of error terms. This analysis was conducted in the *MASS* package in R (R Core Team, 2015) by fitting a linear model where genotype, environment, incomplete block nested within environment, and plot grid row nested within environment, were fitted as fixed effects. The softness data for 30, 45, and 60 min time points did not need transformation (optimal convenient lambda = 1), whereas the square root transformation (optimal convenient lambda = 0.5) was needed for the softness data collected at 15 min.

We fitted a mixed linear model for softness at each time point, to identify and remove significant outliers from the raw data in ASReml-R version 3.0 (Gilmour *et al.,* 2009). The full model used to analyze data collected at each time point had check lines as fixed effects and all other terms from the simple linear model described above, as random effects. The Studentized deleted residuals (Neter *et al.,* 1996) were examined to identify and remove significant outliers.

After outliers were removed, an iterative mixed linear model fitting procedure of the full model was conducted in ASReml-R version 3.0 (Gilmour *et al.,* 2009). Likelihood ratio tests

were conducted to remove all random effect terms from the model that were not significant at $\alpha = 0.05$ (Littell *et al.,* 2006) to generate a final, best fitted model for softness at each time point that enabled generation of best linear unbiased predictors (BLUPs) for each genotype.

For softness at each time point, the variance components from the final model were used to estimate broad-sense heritability (or repeatability) on an entry-mean basis (Holland *et al.*, 2003). Standard errors of the heritability estimates were approximated with the delta method (Holland *et al.*, 2003). Pearson's correlation coefficients ($r$) were computed and used to assess the relationship between softness BLUPs for each pair of time points. This was done using the *cor* function of the *stats* package in R (R Core Team, 2015). The significance of the Pearson's correlation coefficient was declared at $\alpha = 0.05$ with the *cor.test* function of the *stats* package. To represent the true directionality of the relationship between softness BLUPs, the transformed BLUPs at the 15 min time point were back-transformed prior to conducting the correlation analysis.

**Analysis of consumer data**

We used a linear model to analyze the consumer evaluation scores. The Box-Cox procedure (Box and Cox, 1964) was implemented by fitting a linear model, where the fixed effects were variety, sex, and age, with consumer evaluation scores as the dependent variable in the MASS package in R. The convenient Lambda of 1.0 was obtained and thus, data were not transformed. Data were screened for outliers in ASReml-R version 3.0 (Gilmour *et al.*, 2009), by examining the Studentized deleted residuals (Neter *et al.*, 1996) obtained from a linear model fitted with variety, sex, and age as fixed effects. The resultant data were then used to refit the same model in ASReml-R, to conduct an analysis of variance (ANOVA) and estimate the best linear unbiased estimator (BLUE) for each variety. The Tukey-Kramer

honest significance (HSD) test (*P-value* < 0.05) was used to determine if varieties were significantly different from each other.

A linear model was also used to analyze penetrometer data that were collected as a comparison to the consumer evaluation scores. For the penetrometer data across all six post-cooking time points (0, 15, 30, 45, 60 and 75 min), a series of linear models with variety, time point, and their interaction were fitted as fixed effects to find the most appropriate transformation; the log transformation was made for all six time points. We screened for outliers as described earlier. We estimated a BLUE for each variety across all six time points (used for correlation analysis) and at each single time point (used for time series plot). To evaluate whether varieties were significantly different from each other across all six time points, the Tukey-Kramer HSD test (*P-value* < 0.05) was used. The relationship between BLUEs of softness from consumer testing and the penetrometer (across all six time points) was assessed as described earlier. The log transformed penetrometer BLUEs were back-transformed prior to conducting the correlation analysis.

## RESULTS

### Heritability estimates

As for the cassava roots cooked from 15 to 60 min, there was a trend of reduced force (N) required to penetrate the roots (Table 2.1). The highest average amount of force (3.33 N) required to penetrate the roots was recorded after cooking for 15 min, with genotype UG140335 requiring the highest amount of force (3.76 N); while the lowest average force (2.20 N) was recorded after the roots cooked for 60 min with genotype UG142019 requiring the least amount of force (1.89 N). The BLUPs for root softness of all the genotypes are stored in CassavaBase at ftp://ftp.cassavabase.org/ manuscripts/Iragaba_et_al_2019_quality/ Phenotype_infos/ (in a file named: "BLUPs softness manuscript.txt"). We observed the

widest range of variation (1.94 - 3.21 N) for root softness at the 45 min cooking time point; while the narrowest variation (1.89 - 2.49 N) was observed after cooking for 60 min (Table 2.1). Broad-sense heritability (or repeatability) was estimated for root softness at each of the four cooking time points; and overall moderate heritability estimates were observed (Table 2.1). Heritability estimates increased from 0.22 (after cooking for 15 min) to a maximum of 0.37 (after cooking for 45 min); and declined to 0.17 for roots cooked for 60 min.

Table 2.1. Best linear unbiased predictors and heritability estimates of cassava root softness following four different cooking durations.

| Cooking time (min)[a] | Number of cassava genotypes[b] | BLUPs[c] | | | Broad-sense heritability | |
|---|---|---|---|---|---|---|
| | | Mean (N) | SD[d] (N) | Range (N) | Estimate | SE[e] |
| 15 | 268 | 3.33 | 0.155 | 2.94 - 3.76 | 0.22 | 0.087 |
| 30 | 268 | 2.78 | 0.154 | 2.45 - 3.27 | 0.28 | 0.081 |
| 45 | 267 | 2.47 | 0.205 | 1.94 - 3.21 | 0.37 | 0.076 |
| 60 | 267 | 2.20 | 0.096 | 1.89 - 2.49 | 0.17 | 0.092 |

[a]Data based on analysis across two locations; Arua (west Nile region) and Mubende (central region); Back-transformed BLUPs are reported for the 15 min time point.

[b]Only 267 genotypes were evaluated at the 45 and 60 min cooking times, because one of the genotypes did not have enough roots for phenotypic evaluation.

[c]BLUPs, best linear unbiased predictors.

[d]SD, standard deviation of the BLUPs.

[e]SE, standard error of heritabilities.

Statistically significant (*P-value* <0.0001) correlations were detected between all pairwise comparisons of root softness (Figure 2.4). All correlations were positive, with an observed range of moderate ($r$ = 0.45) to strong ($r$ = 0.73) correlation coefficient values (Figure 2.4).

The strongest correlation was detected between trait BLUP values for the 30 min and 45 min cooking times, whereas the correlation of trait BLUP values between the 15 min and 60 min cooking times was the weakest. In general, correlations were strongest for trait BLUPs between cooking times separated by only 15 min, congruent with expectations.

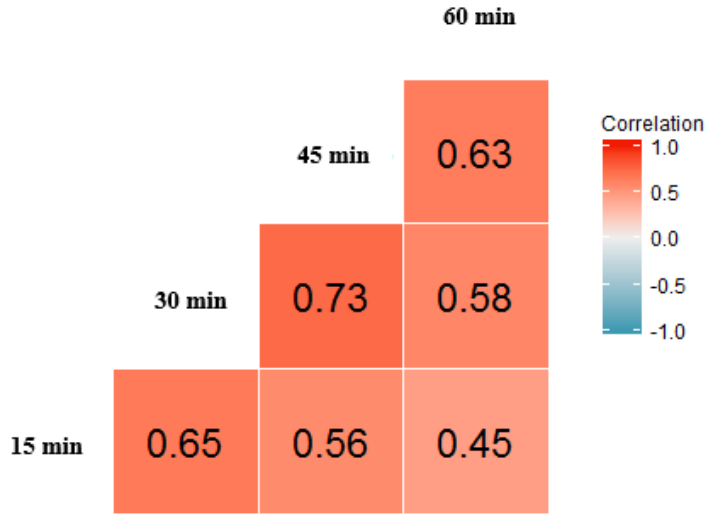



Figure 2.4. Pearson correlation coefficients (r) for best linear unbiased predictors of cassava root softness measured by a penetrometer at the end of four cooking time points: 15, 30, 45, and 60 min. All correlations were significant at $\alpha = 0.05$.

**Consumer testing *versus* penetrometer methods**

A total of 135 consumers from Kibaale district (western Uganda) participated in consumer testing of six cassava varieties, namely Bukalasa, Gwalanda, Kyawada, NAROCASS 1, NAROCASS 2, and Matooke. Unlike in the above-mentioned results where softness was assessed at four time points, for this study, softness was assessed at six post-cooking time points (0, 15, 30, 45, 60 and 75 min). Results revealed that root softness varied significantly (*P-value* < 0.0001) between varieties (Table 2.2), while sex and age of the consumers had no significant effect (*P-value* > 0.05).

Table 2.2. F-values for fixed sources of variation from an analysis of variance for consumer scores of softness for roots from six cassava varieties that were cooked for 50 minutes.

| **Source** | **DF** | **F-value** |
|---|---|---|
| Variety | 5 | $403.34^{****}$ |
| Sex | 1 | $0.29^{NS}$ |
| Age | 1 | $0.75^{NS}$ |
| Residual | 796 | |

NS, Not significant at the < 0.05 level.

**** Significant at the <0.0001 level

Given that consumer testing occurred over a period of 75 min, the penetrometer was used to collect measurements on cooked roots at six post-cooking time points (0, 15, 30, 45, 60 and 75 min). This provision allowed for taking records of cooling effects on root softness; uniquely, this provides for a fair comparison of penetrometer and consumer testing (Figure 2.5). We found statistically significant differences (*P-value* < 0.0001) for variety, time point, and their interaction (Table 2.3). Indicative of high concordance between softness assessment methods, a strong agreement ($r^2$ = 0.91) was detected between the BLUEs of penetrometer root softness and consumer testing (Figure 2.6).

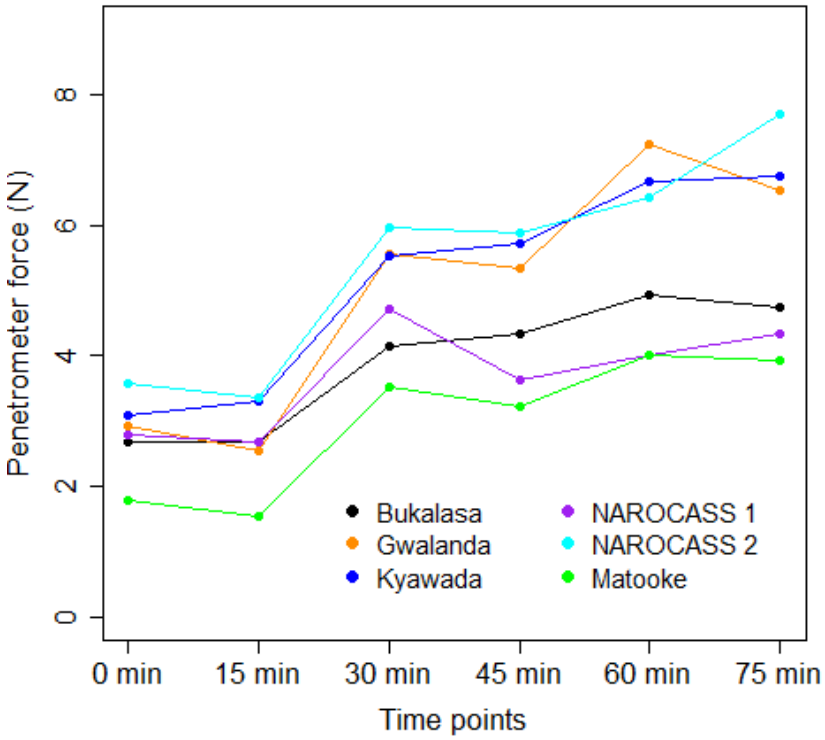



Figure 2.5. Time series of back-transformed BLUEs for root softness measured by the penetrometer at six post-cooking time points for six cassava genotypes.

Table 2.3. F-values for fixed sources of variation from an analysis of variance for penetrometer measurements of softness for roots from six cassava varieties.

| **Source** | **DF** | **F-value** |
|---|---|---|
| Variety | 5 | 228**** |
| Time point[a] | 5 | 453**** |
| Variety*Time point | 25 | 34**** |
| Residual | 143 | |

**** Significant at the <0.0001 level

[a]Roots were evaluated at six post-cooking time points (0, 15, 30, 45, 60 and 75 min) after cooking for 50 minutes

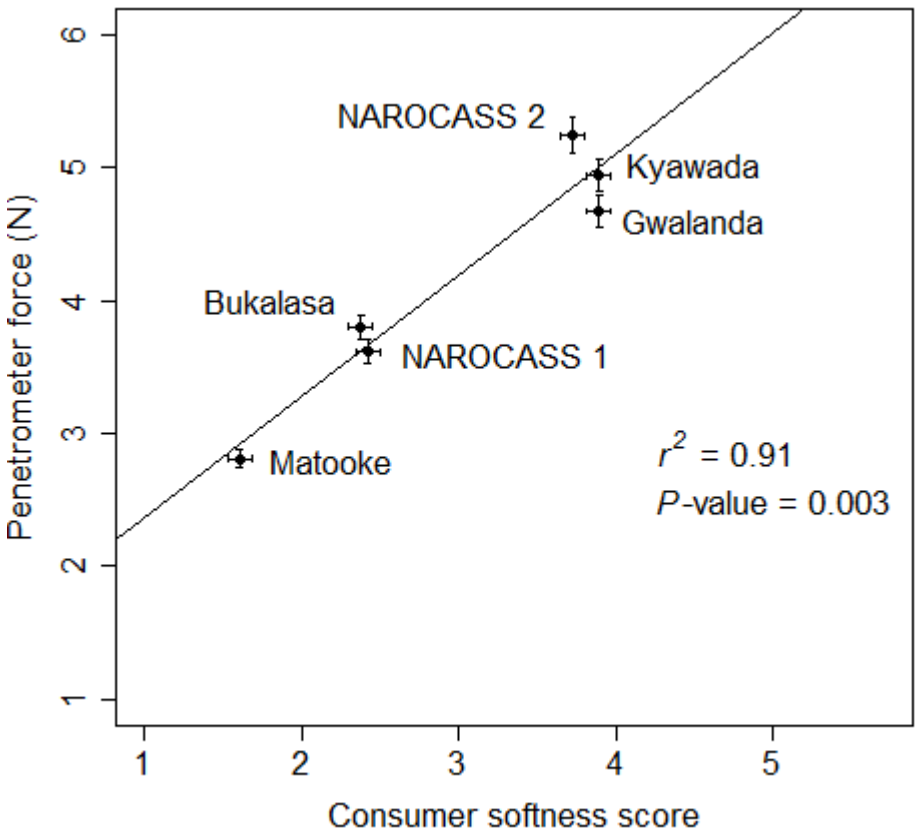



Figure 2.6. The relationship between BLUEs of penetrometer measurements and consumer scores of softness for cassava roots of six genotypes that were cooked for 50 minutes.

The Tukey-Kramer HSD test grouped the varieties from consumer testing and penetrometer methods in similar sub-groups (Figure 2.7). In both methods, Matooke, Bukalasa and NAROCASS 1 had the lowest values (softer); whereas Kyawada, Gwalanda and NAROCASS 2 had the highest values (harder). The sub-groups of the varieties were in concordance with the prior consumer classification regarding the level of softness.

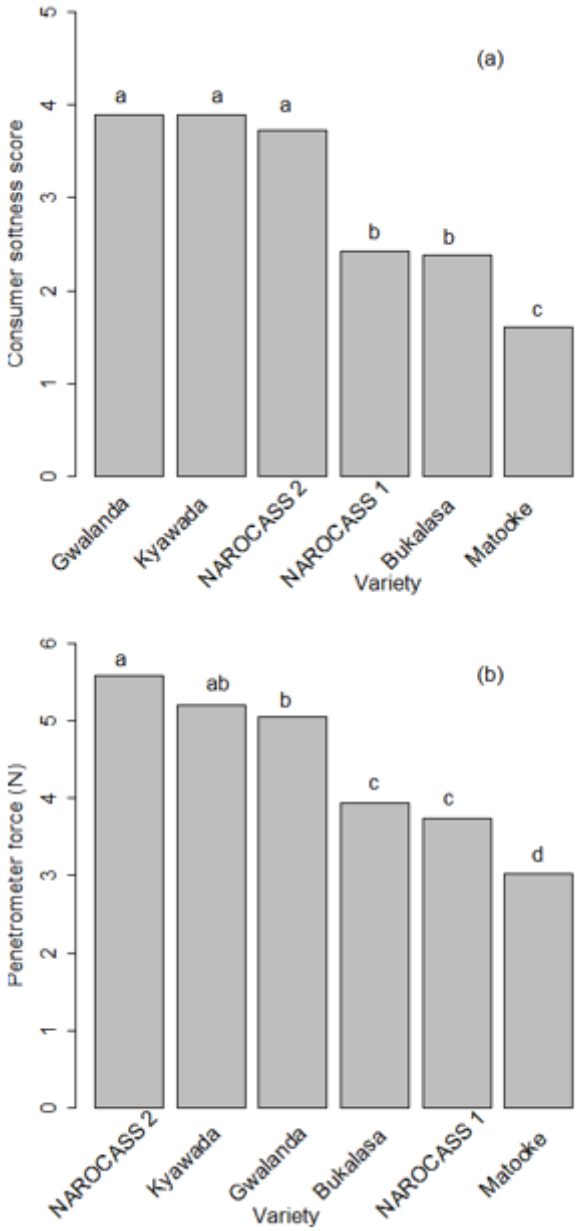



Figure 2.7. (a) Bar plot showing Tukey-Kramer HSD means separation of softness for six cassava genotypes based consumer testing evaluation; (b) Bar plot showing Tukey-Kramer HSD means separation of softness for six cassava genotypes obtained by use of the penetrometer. Means followed by the same letter are not significantly different at $\alpha = 0.05$.

## DISCUSSION

The highest broad-sense heritability (0.37) was observed among cassava genotypes from Arua and Mubende after cooking for 45 min (Table 2.1), implying that selecting for root softness at 45 min after cooking would maximize genetic gain for softness. Similarly, it was

at 45 min that we obtained the highest standard deviation for the estimates of BLUPs, indicating a considerable amount of genetic variation, and thus scope for breeding for root softness in cassava. In part, the relatively lower broad-sense heritabilities observed at 15 and 60 min could be attributed to the fact that most varieties had not yet cooked at 15 min, or had over cooked at 60 min. Previous studies, though limited by number of clones (less than 30 clones), have observed phenotypic variability for softness of cooked cassava roots (Ngeve, 2003; Beleia *et al.,* 2004a; Padonou *et al.,* 2005).

In the present study, the reported broad-sense heritability estimates provide an indication of the proportion of observed phenotypic variance of a population that is due to genetic effects. In theory, this indicates how much genetic gain can be achieved by the breeder (Holland *et al.*, 2003; Acquaah, 2012). Notably, none of the previous studies quantified heritability for softness, an important aspect in guiding selection of genotypes with favorable attributes by breeders during new variety development process. Results in the present study provide the first insights in determining the broad-sense heritability for softness.

The observed differences in softness among genotypes (Table 2.1 and Figure 2.7) could be attributed to the differences in the amount of pectin and/or intercellular cell-wall adhesions (Eggleston and Asiedu, 1994; Favaro *et al.*, 2008; Maieves *et al.*, 2012). Eggleston and Asiedu (1994) reported that the cell-walls of mealy cassava varieties were less cohesive than those of non-mealy varieties. Favaro *et al.* (2008) examined the effect of thermal treatment on softening of cassava roots, and found that longer cooking times softened tissue due to reduction in intercellular adhesive strength. Furthermore, previous studies on the softness of cooked cassava roots have indicated that age of a genotype at harvest, season and soil type affect the level of softness (Ngeve, 2003; Padonou *et al.,* 2005).

In the present study, a general trend of reduction in average force was required to penetrate the roots from 15 min to 60 min cooking times (Table 2.1). This trend conforms to what is expected since prolonged cooking should lead to softening of the roots. Several studies have reported that prolonged cooking time leads to softening of cassava roots (Lorenzi, 1994; Beleia *et al.*, 2004b, 2006).

We quantified the relationship between the two methods of assessment of root softness; the penetrometer and consumer-based evaluations. We observed a strong correlation ($r^2$ = 0.91; *P-value* = 0.003) between both methods implying that the penetrometer can be reliably used to estimate softness of cooked cassava roots as evaluated by consumers. This finding compares well with the correlation of 0.75 reported by Padonou *et al.* (2005) in determining the relationship between sensory evaluation and penetration test by an instrument in cassava.

Ideally, consumer testing would be the best direct method for evaluation of softness (Safokantanka and Owusu-nipah, 1992; Padonou *et al.*, 2005); however, it is ineffective for assessing variability for a large number of genotypes that are often encountered in early evaluation stages. Thus, the strong correlation between penetrometer and consumer testing results justify the use of the penetrometer as an indirect method to measure softness of cooked cassava roots. The advantage of the penetrometer over consumer testing would be its efficiency when dealing with large sample sizes. Furthermore, it avoids subjectivity that is commonplace when using non-expert consumer evaluation panels.

## CONCLUSION

This study has provided the first insights into heritability estimates for softness of cooked cassava roots, for which we observed moderate heritabilities ($H^2$ = 0.37) at 45 min cooking time. Evaluating genotypes at this time point will, within limits, increase selection efficiency, leading to overall increase in genetic gain. Secondly, there is a high correlation ($r^2$ = 0.91)

between penetrometer and consumer testing results for root softness. Thus, we recommend the use of the penetrometer method as a high-throughput phenotyping tool for softness of cooked cassava roots. Lastly, softness of cooked cassava root attribute was added to Cassavabase trait ontology, (https://www.cassavabase.org/cvterm/77616/view).


## ACKNOWLEDGEMENT

This research was undertaken as part of Roots, Tubers and Bananas (RTB) research, and funded by, the CGIAR Research Program on RTB and supported by CGIAR Trust Fund contributors. Additional funds were provided by the Next Generation Cassava Breeding Project through funds from the Bill and Melinda Gates Foundation (Project No. OPP1048542) and the Department for International Development of the United Kingdom (UKaid) (Project No. OPP1048542). We thank the technical staff of National Crops Resources Research Institute (NaCRRI) Root Crops Programme, who assisted with field trial establishment, maintenance, and data collection both in the field and laboratory. We also acknowledge the technical guidance from RTBfoods project staff. We give special thanks to Guillaume Bauchet for assistance in data curation.

**REFERENCES**

Acquaah, G. 2012. Principles of Plant Genetics and Breeding. Second Edi. John Wiley and Sons, Inc., West Sussex, PO19 8SQ, UK. 740pp.

Alene, A.D., Khataza, R., Chibwana, C., Ntawuruhunga, P. and Moyo, C. 2013. Economic impacts of cassava research and extension in Malawi and Zambia.

*Development and Agricultural Economics* 5(11):457–469. doi: 10.5897/JDAE 2013.0496.

Arua District Local Government. 2012. Arua District Local Government Statistical Abstract. https://www.ubos.org/online files/uploads/ubos/Arua.pdf. Accessed 25 March 2019

Bechoff, A., Tomlins, K., Fliedel, G., Becerra, L.A. Westby, A., Hershey, C. and Dufour, D. 2018a. Cassava traits and end-user preference: Relating traits to consumer liking, sensory perception, and genetics. *Critical Reviews in Food Science and Nutrition* 58(4):547–567. doi: 10.1080/ 10408398.2016.1202888.

Bechoff, A., Chijioke, U., Westby, A. and Tomlins, K.I. 2018b. Yellow is good for you: Consumer perception and acceptability of fortified and biofortified cassava products. *PLoS One* 13(9):1–23. doi: 10.1371/journal.pone.0203421.

Beleia, A., Butarelo, S.S. and Silva, R.S.F. 2006. Modeling of starch gelatinization during cooking of cassava (*Manihot esculenta* Crantz). *LWT - Food Science and Technology* 39(4): 399–404. doi: 10.1016/ j.lwt.2005.02.021.

Beleia, A., Prudencio-Ferreira, S.H., Yamashita, F., Sakamoto, T.M. and Ito,

L. 2004a. Sensory and instrumental texture analysis of cassava (*Manihot esculenta* Crantz) roots. *Journal of Texture Studies* 35: 542–553.

Beleia, A., Yamashita, F., De Moraes, S.R., Da Silveira, C.A. and Miranda, L.A. 2004b. Textural changes during cooking of cassava (*Manihot esculenta* Crantz) roots. *Journal of the Science of Food and Agriculture* 84(14): 1975–1978. doi: 10.1002/jsfa.1917.

Bernardo, R. 2010. Breeding for quantitative traits in plants. Second Edition. Stemma Press, Woodbury, MN., USA. 390pp.

Box, G.E.P., and Cox, D.R. 1964. An Analysis of Transformations. *Royal Statistical Society. Series B (Methodological)* 26(2): 211–252.

Ceballos, H., Iglesias, C.A., Perez, J.C. and Dixon, A.G.O. 2004. Cassava breeding: Opportunities and challenges: *Plant Molecular Biology* 56:503–516.

Chiwona-Karltun, L., Nyirenda, D., Mwansa, C.N., Kongor, J.E., Brimer, L., Haggblade, S. and Afoakwa, E.O. 2015. Farmer preference, utilization, and biochemical composition of improved Cassava (*Manihot esculenta* Crantz) varieties in

Southeastern Africa. *Economic Botany* 69(1): 42–56. doi: 10.1007/s12231-0159298-7.

Eggleston, G. and Asiedu, R. 1994. Effects of boiling on the texture of cassava clones: A comparison of compressive strength, intercellular adhesion and physico-chemical composition of the tuberous roots. *Tropical Science* 34: 259–273.

FAO. 2013. Save and grow cassava: A guide to Sustainable Production Intestification. Rome, Italy. 129pp. http://www.fao.org/3/a-i3278e.pdf. Accessed 20 March 2019.

FAO and IFAD. 2000. The world Cassava Economy: Facts, Trends and Outlook. Viale delle Terme di Caracalla, 00100 Rome, Italy. 46pp.

Favaro, P.S., Beleia, A., da Silva Fonseca Junior, N. and Waldron, W.K. 2008. The roles of cell wall polymers and intracellular components in the thermal softening of cassava roots. *Food Chemistry* 108:220– 227. doi: 10.1016/j.foodchem.2007.10. 070.

Fick, S. and Hijmans, R. 2017. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*. http://worldclim.org/version2.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. 2009. ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, HPI IES, UK. 372pp.

Holland, J.B., Nyquist, W.E. and CervantesMartinez, C.T. 2003. Estimating and interpreting heritability for plant breeding: An update. pp. 9–112. In: Plant Breeding Reviews. John Wiley and Sons, Inc., Hoboken, New Jersey, USA.

Hough, G., Wakeling, I., Mucci, A.,

Chambers, E., Gallardo, I.M. and Alves, L.R. 2006. Number of consumers necessary for sensory acceptability tests. *Food Quality and Preference* 17: 522–526. doi: 10.1016/j.foodqual.2005. 07.002.

International Institute of Tropical Agriculture (IITA). 1990. Cassava in tropical Africa. A reference Manual. International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria. 176pp.

ISO. 1994. Sensory analysis, Methodology, Texture profile. First Edi. ISO 11036:1994(E). 56. CH-1211, Geneva 20, Switzerland. 20pp.

Kayondo, S.I., Pino, D., Carpio, D., Lozano, R., Ozimati, A., Yona, B., Gracen, V., Offei, S., Ferguson, M. and Kawuki, R. 2018. Genome-wide association mapping and genomic

prediction for CBSD resistance in *Manihot esculenta*. *Scientific Reports* 8:1549: 1–11. doi: 10.1038/s41598018-19696-1.

Kombo, G.R., Dansi, A., Loko, L.Y., Orkwor, G.C., Vodouhè, R., Assogba, P. and Magema, J.M. 2012. Diversity of cassava (*Manihot esculenta* Crantz) cultivars and its management in the department of Bouenza in the Republic of Congo. *Genetic Resources and Crop Evolution* 59(8): 1789– 1803. doi: 10.1007/s10722-012-9803-0.

Lebot, V. 2009. Tropical root and tuber crops: Cassava, sweet potato, yams and aroids.

Crop Production Science in Horticulture (17). CABI, Wallingford, UK. 413pp.

Littell, R.C., Milliken, G.A., Stroup, W.W. and Wolfinger, R. 2006. SAS System for Mixed Models. Second Edi. SAS Publishing, Cary, NC, USA. 122pp.

Lorenzi, O. 1994. Culinary quality variation in cassava roots. *Bragantia Campinas* 53(2): 237–245.

Maieves, H.A., De Oliveira, D.C., Bernardo, C., Müller, C.M.D.O. and Amante, E.R. 2012. Microscopy and texture of raw and cooked cassava (*Manihot esculenta*

Crantz) roots. *Journal of Texture Studies* 43(2): 164–173. doi: 10.1111/j.17454603.2011.00327.x.

Mubende District Local Government. 2013. Mubende District Local Government Statistical Abstract 2012/2013. https:// www.ubos.org/onlinefiles/uploads/ubos/ Mubende.pdf. Accessed 25 March 2019.

Nakabonge, G., Samukoya, C. and Baguma, Y. 2017. Local varieties of cassava: Conservation, cultivation and use in Uganda.

*Environment, Development and Sustainability*. doi: 10.1007/s10668-0179997-6.

Neter, J., Kutner, M.H., Nachtsheim, C.J. and Wasserman, W. 1996. Applied Linear Statistical Models. Fourth edition. McGraw-Hill, Boston, USA. 1408pp.

Ngeve, J.M. 2003. Cassava root yields and culinary qualities as affected by harvest age and test environment. *Science of Food and Agriculture* 83:249–257. doi: 10.1002/ jsfa.1307.

Padonou, W., Mestres, C. and Nago, C.M. 2005. The quality of boiled cassava roots: Instrumental characterization and relationship with physicochemical properties and sensorial properties. *Food Chemistry* 89:261–270. doi: 10.1016/ j.foodchem.2004.02.033.

R Core Team. 2015. A language and environment for statistical computing. R foundation for statistical computing. Vienna, Austria. https://www.r-project.org/ Accessed 20 March 2019.

Rabbi, I.Y., Kulakow, P.A., Manu-aduening,

J.A., Dankyi, A.A., Asibuo, J.Y., Parkes, E.Y., Abdoulaye, T., Girma, G., Gedil, M.A., Ramu, P., Reyes, B. and Maredia, M.K. 2015. Tracking crop varieties using genotyping-by-sequencing markers: A case study using cassava (*Manihot esculenta* Crantz). *BMC Genetics* 16:115.

doi: 10.1186/s12863-015-0273-1.

Safo-kantanka, O. and Owusu-nipah, J. 1992. Cassava varietal screening for cooking quality: Relationship between dry matter, starch content, mealiness and certain microscopic observations of the raw and cooked tuber. *Science of Food and Agriculture* 60:99–104.

Sajeev, M.S., Sreekumar, J., Unnikrishnan, M., Moorthy, S.N. and Shanavas, S. 2010. Kinetics of thermal softening of cassava tubers and rheological modeling of the starch. *Food Science Technology* 47(5): 507–518. doi: 10.1007/s13197-010-00870.

# CHAPTER 3 GENOMIC CHARACTERIZATION OF UGANDAN SMALLHOLDER FARMER-PREFERRED CASSAVA (*MANIHOT ESCULENTA* CRANTZ) VARIETIES[2]

## ABSTRACT

Understanding the genetic relationships among farmer-preferred varieties is indispensable to cassava genetic improvement efforts. In this study, we present a genetic analysis of 547 samples of cassava grown by 192 smallholder farmers, which were sampled at random within four districts in Uganda. We genotyped these samples at 287,952 single-nucleotide polymorphisms using genotyping-by-sequencing and co-analyzed them with 349 lines from the national breeding program in Uganda. The samples collected from smallholder farmers consisted of 86 genetically unique varieties, as assessed using a genetic distance-based approach. Of these varieties, most were cultivated in only one district (30 in Kibaale, 19 in Masindi, 14 in Arua and three in Apac) and only three were cultivated across all districts. The genetic differentiation we observed among farming districts in Uganda (mean $F_{ST}$ 0.003) is similar to divergence observed within other countries. Despite the fact that none of the breeding lines were directly observed in farmer fields, genetic divergence between the populations was low ($F_{ST}$ = 0.020). Interestingly, we detected the presence of introgressions from the wild relative *M. glaziovii* on chromosomes one and four, which implies ancestry with cassava breeding lines. Given the apparently similar pool of alleles in the breeding

[2]Paula Iragaba[1], Robert S. Kawuki[2], Guillaume Bauchet[3], Punna Ramu[4], Hale A. Tufan[1,5], Elizabeth D. Earle[1], Michael A. Gore[1] and Marnin Wolfe[1]. Genomic characterization of Ugandan smallholder cassava (*Manihot esculenta* Crantz) varieties. Submitted to Crop Science Journal.

germplasm, it is likely that breeders have the raw genetic material they require to match the farmer-preferred trait combinations necessary for adoption. Our study highlights the importance of understanding the genetic make-up of cassava currently grown by smallholder farmers and relative to that of plant breeding germplasm.

**Key words**: Genotype, genotyping-by-sequencing, diversity, *Manihot esculenta,* farmer, variety

## INTRODUCTION

Cassava (*Manihot esculenta* Crantz) is an important source of food to about 800 million people globally (FAO, 2013). Although cassava was originally domesticated in Latin America (Allem, 1999), it is currently grown all over the tropics at latitudes between 30°N and 30°S (Ceballos et al., 2004). More than half of the total cassava produced in the world is grown in sub-Saharan Africa (FAO, 2016), where it ranks as the second most important staple food crop (Nweke et al., 2002). Cassava is cultivated by smallholder farmers as a reliable source of food and income crop in unstable environments: it is vegetatively propagated and tolerates marginal soils, limited rainfall, and has a flexible harvesting schedule (El-Sharkawy, 1993; FAO and IFAD, 2000; Nweke et al., 2002; Ceballos et al., 2004; Kizito et al., 2007). These attributes largely explain its wide scale adoption and cultivation across the continent.

Cassava breeding has led to significant genetic improvement for productivity traits such as fresh root yield and dry matter content (Kawano, 2003; Kawuki et al., 2011), but less so for quality traits (Ceballos et al., 2004; Lebot, 2009). Farmers in sub-Saharan Africa are the main consumers of the cassava they produce, and they often grow multiple locally adapted varieties or elite varieties that meet specific end-user traits including processing and cooking qualities (Tumuhimbise et al., 2012; Alene et al., 2013; Teeken et al., 2018). Previous studies have reported low adoption rates for varieties that do not meet the needs and preferences of end-users (Alene et al. 2013; Afolami et al. 2015). Breeding programs must therefore prioritize end-user trait preferences to increase adoption and consequently breeding impact (Bechoff et al., 2018; Nakabonge et al., 2018).

In Uganda, a census of agriculture revealed that cassava is cultivated in 96.2% of the districts in the country, and that cassava is the second most important food crop after bananas

(UBOS, 2010). A recent study conducted by Nakabonge et al. (2018) reported that Ugandan farmers grow cassava mainly for home food consumption and/or sale, and that different varieties are preferred for certain traits. Some of these important traits include cooking quality, storability in ground, texture of boiled roots, and early maturity (Tumuhimbise et al., 2012; Bechoff et al., 2018; Nakabonge et al., 2018). Despite the importance of cassava in Uganda, little is known about the genetic identity and diversity of varieties currently being grown in farmers' fields. Turyagyenda et al. (2012) studied the diversity within Ugandan farmer-preferred varieties, however, their study was limited by a small number of individuals (51 farmer-varieties and 15 elite accessions) and used only 26 simple-sequence repeat markers. Thus, there is a need to further explore the genetic diversity of the cassava varieties grown by Ugandan smallholder farmers and determine their relationships to breeding populations using a larger number of samples and dense set of genome-wide markers in order to better draw inferences about the varieties grown by farmers and the breeding germplasm.

Achieving genetic gain through artificial selection requires that: a) adequate genetic variation for the trait of interest is available; b) the trait of interest is heritable; c) the trait can be efficiently and effectively assessed to enable selection decisions (Falconer and Mackay, 2009). This requires a deep understanding of varieties that are currently being grown, an aspect that can be captured through quantitative assessment of their genotypic and phenotypic relationships and comparisons to elite breeding materials that will eventually replace what is currently grown (Acquaah, 2012; Alene et al 2013). Neither morphological descriptors nor variety names reported by farmers can provide unambiguous varietal identification (Kizito et al., 2007; Rabbi et al., 2015; de Leon et al., 2016; Nduwumuremyi et al., 2017; Nakabonge et al., 2018), a situation that can complicate excursions aimed at collecting germplasm for conservation and/or breeding purposes. In contrast, genetic markers offer robust and

objective means of variety identification, which has been demonstrated for cassava farmer-varieties in Ghana using genotyping-by-sequencing (GBS) (Rabbi et al., 2015).

In this study, we present a genetic survey of cassava cultivated in four districts within Uganda. We analyzed leaf tissue from 547 cassava plants sampled from smallholder farmers resident in above selected districts. Our first objective was to assess the structure of genetic relationships among cassava cultivated in the four major cassava growing districts. Our second objective was to compare the genetic relationships of farmer-grown varieties with 349 breeding lines sourced from the cassava breeding program at the National Crops Resources Research Institute (NaCRRI), Namulonge, Uganda.

## MATERIALS AND METHODS

### *Study sites and collection of leaf samples*

This study was conducted in four districts (Apac, Arua, Kibaale, and Masindi) in Uganda (Supplementary Figure 1). These districts were selected because they are associated with high cassava production and consumption (UBOS, 2010). Also, these districts experience low cassava mosaic and cassava brown streak disease prevalence (Alicai et al., 2007). We conducted a survey to capture cassava trait preferences within two randomly selected villages per district. Within each village, we selected 24 smallholder cassava farmers (stratified by age and sex) using simple random sampling to participate in the study. In total, 192 farmers participated in the study from the four districts. The study plan and consent forms to engage human participants in the study were reviewed and approved by the Institutional Review Board (IRB) of Cornell University (IRB ID 1502005316). The study only commenced upon farmers granting us permission. We employed an interview guide to collect data on farming practices, cassava varieties cultivated, their names, and traits liked and disliked by farmers. In addition, we sampled three to four apical leaves from each

farmer-cultivated cassava variety and preserved them in silica gel for genotyping (Girma et al., 2017). In total, this resulted in a collection of 556 samples from farmer-varieties.

*DNA extraction and genotyping*

Genomic DNA was isolated from each of the collected leaf tissue samples; extraction was undertaken at NaCRRI, Namulonge, Uganda, using the method described by Dellaporta et al. (1983). The DNA samples were shipped to the Cornell Biotechnology Resource Center, where they were analyzed using the GBS protocol of Elshire et al. (2011) with the *Ape*KI restriction enzyme (Hamblin and Rabbi, 2014). Genotype discovery and calling was done jointly on 1530 samples: in addition to the 556 new samples collected for this study, 624 samples described by Iragaba et al. (2019) were included as well as a random set of 350 samples that were selected from a diverse panel of breeding lines at NaCRRI (Kayondo et al., 2018). The single-end raw reads of 150 bp were processed through the TASSEL GBS v2 production pipeline (Glaubitz et al., 2014). Genotype calls were allowed only when a minimum of two reads were present in a given sample. This process generated 470,413 single-nucleotide polymorphisms (SNPs) on 1530 samples (Supplementary Figure 2). Sites with more than two alleles, extreme deviation from Hardy-Weinberg equilibrium (Chi-square > 20), and loci with >80% missing data were removed (Chan et al., 2016). We also removed samples that had more than 80% missing data. After this filtering, there were a total of 287,952 SNPs scored on 1519 samples. The remaining SNP loci with missing genotypes were imputed with Beagle version 4.0 (Browning and Browning, 2009). Thereafter, we obtained a subset of 968 samples from the above 1519 samples. The selected subset consisted of 547 farmer-varieties, (nine of the initial 556 samples had > 80% missing data, thus these nine were removed prior to imputation), 349 NaCRRI breeding lines, and 72 biological replicates that were the checks used in the study by Iragaba et al. (2019). These 72 biological replicates consisted of five genotypes that included: three released varieties: UG110017

(NAROCASS 1), UG110004 (NASE 4) and UG110014 (NASE 14); a common breeding line UG110015 (TME-14); and a landrace, UGL15228 (Lugigana). These five genotypes had 17, 12, 19, 18, and 6 biological replicates, respectively. The bioinformatic and statistical analysis workflow is depicted in Supplementary Figure 2 for clarity.

***Statistical analyses***

Our principal objective was to determine the number of unique varieties grown in Uganda and their relative abundances. Accordingly, using the SNP marker data, we determined a threshold of genetic similarity above which differences among samples were indistinguishable, as outlined in previous studies (Rabbi et al., 2015; Myles et al., 2011). This was done using SNP data of the five genotypes that had multiple biological replicates. We used PLINK v1.90 (Purcell et al., 2007) to compute the pairwise identity-by-state (IBS) similarities between the replicated samples. We then used the *dist* function in the *stats* R package (v3.5.1, R Core Team, 2018) to convert the IBS matrix to a dissimilarity structure. From the distance matrix, we used the *hclust* R package to conduct Ward's hierarchical clustering.

Based on the clustering results, we determined a threshold of Ward's distance that could separate biologically replicated samples into distinct clonal groups of genotypes. The chosen threshold was subsequently used in downstream analyses to declare which varieties were distinct. The selected threshold was applied to a distance matrix of a dataset including 547 samples from farmer-varieties, 349 breeding lines and 72 biological replicates using the *cutree* function in the *stats* R package. To reduce redundancy of multiple samples in the same clonal group, after clustering we subsetted each clonal group of samples such that it was represented only with a single randomly chosen sample per variety (clonal group) per district-of-sample-origin. From this point onwards, we refer to this collection of representative samples as the set of unique varieties in each district.

As a complement to our hierarchical clustering approach to identify genetically-unique varieties from the 547 samples collected from farmer-varieties, we ran the ADMIXTURE model (Alexander et al., 2009; Alexander and Lange, 2011). As recommended by Alexander et al. (2009), we first filtered our dataset to obtain a SNP marker set that was mostly in linkage equilibrium using PLINK --*indep-pairwise* with a window size of 50, step size of 10, pairwise linkage disequilibrium (LD) $r^2$ threshold of 0.3 and minor allele frequency (MAF) < 0.01. With the LD-pruned dataset (119,714 SNPs), we ran the ADMIXTURE program with the ancestral population number (*K)* varying from 1 to 18 to determine the optimal *K* based on the lowest program-reported five-fold cross-validation error rate. The ADMIXTURE results for the optimal *K* value were compared to the IBS-based set of genetically-unique varieties. We used this to verify that putative identical varieties had approximately the same ancestry proportions.

Additionally, using the IBS-derived set of clonal groups, we examined the correspondence between farmer-reported variety names and their genetic identities using a chord diagram generated with the *chordDiagram* function of the *circlize* R package. For ease of visualizing the plot, we considered only clonal groups that had more than 20 members and included farmer-reported variety names that appeared more than 11 times in our dataset.

We quantified the overall level of genetic differentiation between districts using the fixation index ($F_{ST}$) as implemented in *vcftools* (Danecek et al., 2011; Weir & Cockerham 1984). We computed between-district $F_{ST}$ using the set of samples we described above in which each clonal group (unique variety) is represented by one sample per district in which that variety was found. Prior to $F_{ST}$ computation, we removed SNPs with MAF < 0.01. We also used principal component analysis (PCA, *prcomp* function in R with center and scale set to TRUE) to reduce patterns of genetic relatedness in our dataset to a few dimensions that could be visually examined. Before any PCA analysis, we filtered SNPs with MAF < 0.01 and also

removed monomorphic SNPs. In order to observe trends in diversity across the genome, we used the *vcftools* function *--window-pi* (Danecek et al., 2011) to compute the nucleotide diversity ($\pi$) per 0.5 Mb window for the unique set of varieties per district.

Following preliminary analyses of our dataset, and given a recent study indicating the prevalence in modern cassava of large introgressions from the wild relative *Manihot glaziovii* (Wolfe et al. *in review*), from our samples, we extracted the dosage of *M. glaziovii* introgression diagnostic alleles at 31642 diagnostic markers described in Wolfe et al. *in review* (Supplementary Table 1 from that paper). We computed the proportion of *Manihot glaziovii* alleles per sample across the set of introgression diagnostic markers observed in our dataset both genome-wide and in two focal regions described in Wolfe et al. (*in review*), chromosome 1 from 25Mb to the end, and chromosome 4 from 5-25Mb.

Lastly, to explore the relationship between varieties cultivated by farmers and the NaCRRI breeding lines, we conducted another PCA, and computed $F_{ST}$ and nucleotide diversity values between the farmer-varieties and breeding lines. For these analyses, we used a random sample of unique varieties per district to represent the farmer-varieties and all the 349 breeding lines. The nucleotide diversity per 0.5 Mb window and $F_{ST}$ were computed in *vcftools* using procedures described above. Thereafter, we plotted the distribution of the ratio of nucleotide diversity per 0.5 Mb window of breeding lines to farmer-varieties.

**Data availability**

The imputed SNP genotypic data obtained from 968 samples used in this study are available on Cassavabase website https://cassavabase.org/breeders_toolbox/protocol/6 or (ftp://ftp.cassavabase.org/manuscripts/Iragaba_et_al_2019_diversity/Genotype_infos/) in a file named, “Iragaba GBS.vcf.gz”.

## RESULTS

### *Number of unique varieties grown in Uganda and their relative abundances*

In total, we collected 547 leaf samples from different cassava plants grown by 192 smallholder farmers. Collectively, this translated to 156, 139, 137 and 115 samples that were sourced from farmers' fields in Kibaale, Arua, Masindi, and Apac, respectively (Table 3.1; Supplementary Table 1). Based on varietal names assigned by farmers, we recorded mean of three varieties cultivated per farmer in Arua, Kibaale, and Masindi districts, and two in Apac. Overall, some farmers were growing as few as one and as many as six varieties.

Table 3.1. Summary of leaf samples collected from cassava plants grown by smallholder farmers within four districts in Uganda and the number of unique varieties per household.

| District | Total number of samples collected | Average number of samples per household | Total number of unique varieties[a] | Average number of unique varieties per household |
|---|---|---|---|---|
| Apac | 115 | 2.4 | 21 | 2.3 |
| Arua | 139 | 2.8 | 24 | 2.6 |
| Kibaale | 156 | 3.4 | 41 | 3.3 |
| Masindi | 137 | 2.7 | 33 | 2.6 |

[a]Unique varieties determined based on identity-by-state distinctions

Ward's distance threshold of 0.075 clearly grouped biological replicates together and distinctly separated the five known genotypes (NAROCASS 1, NASE 4, NASE 14, TME-14 and Lugigana) from each other (Supplementary Figure 3.). After applying this threshold to the 547 test samples collected from farmers, we identified a total of 86 unique varieties. Most of these unique varieties (n=65) were only found in a specific district: 30 in Kibaale, 19 in

Masindi, 13 in Arua, and three in Apac. Of the remaining 21 unique varieties, only three were present in all four districts, six were present in at least three districts, and 12 were present in at least two districts (Supplementary Table 2). Similar to farmer-reported variety names, we found an average of two genetically-unique varieties cultivated per farmer in Apac, while an average of three distinct varieties were cultivated per farmer in the other three districts (Table 3.1). Most of the identified unique varieties were highlighted less than five times ($n$=60); only 14 varieties were observed more than 10 times (Supplementary Table 3).

To complement IBS results, we used ADMIXTURE analysis on the 547 samples at $K$=14 because that had the lowest cross-validation error rate (Supplementary Figure 4). We observed that these 547 samples with the same proportion of ancestry were almost always identified to be in the same clonal group derived from the Ward's threshold (Figure 3.1; Supplementary Table 1). For example, all samples ($n$=80) in clonal group 3 had approximately 100% of their proportion derived from ancestry 11. Similarly, all samples ($n$=38) belonging to clonal group 355 were entirely derived from ancestry 8 (Supplementary Table 1).

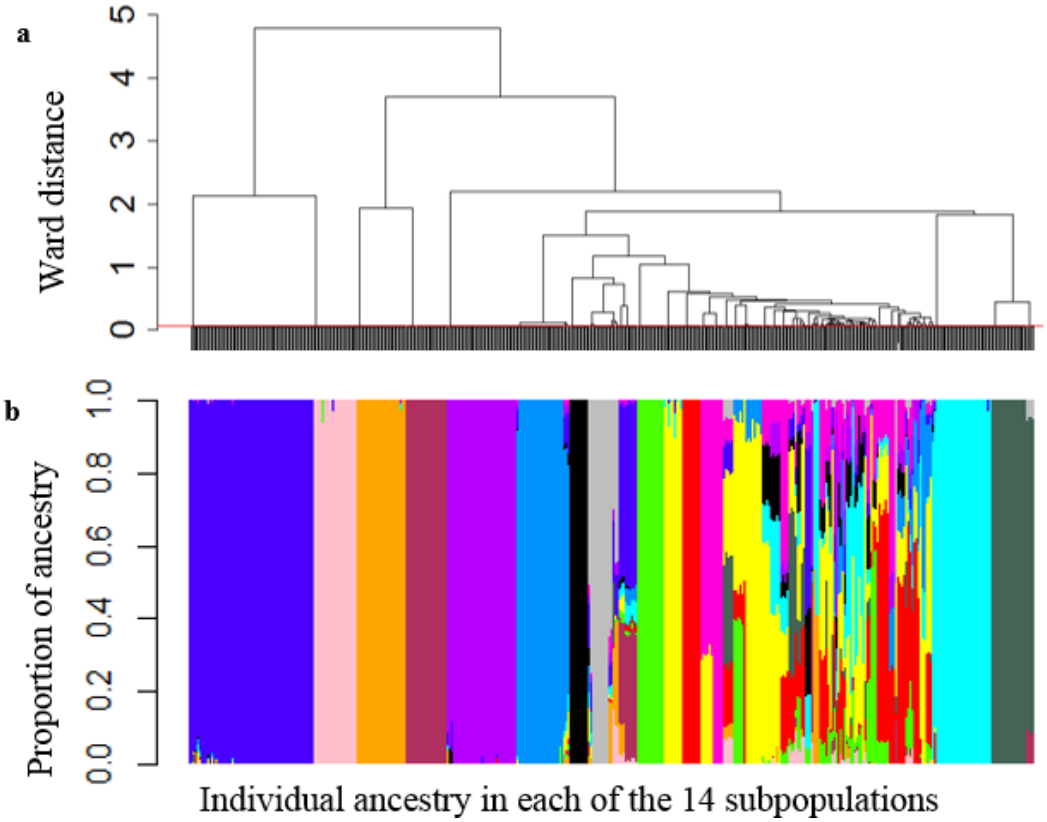



Figure 3.1. a). Dendrogram from Ward's hierarchical clustering of IBS matrix and b). Population structure (based on subpopulations; K=14) analysis on 547 cassava samples using ADMIXTURE. In the panel, each sample is indicated by a vertical bar partitioned in one or more colored segments, and the respective length of the bar represents the proportion of the individual's genome ancestry in a given subpopulation. Each color represents a different ancestry subpopulation.

***Genetic relationships of cassava varieties cultivated in different districts in Uganda***

The genetic divergence between the 86 unique cassava varieties cultivated in the four districts was low, with $F_{ST} < 0.05$ for all pairwise comparisons (Table 3.2). Additionally, results from PCA indicated no clear clustering pattern of varieties based on their location (Figure 3.2). The percentage variation explained by each of the principal components (PCs) was relatively low (Supplementary Figure 5). Furthermore, the average nucleotide diversity among farmer-varieties was $1.06 \times 10^{-4}$, $1.05 \times 10^{-4}$, $1.02 \times 10^{-4}$, and $1.01 \times 10^{-4}$ in Apac,

Arua, Kibaale, and Masindi, respectively (Table 3.3). Generally, all districts exhibited low nucleotide diversity; the least value was observed in Masindi and the highest value observed in Apac.

Table 3.2. $F_{ST}$ estimates among cassava varieties cultivated by smallholder farmers in four districts of Uganda.

| | Kibaale (n=41) | Masindi (n=33) | Arua (n=24) | Apac (n=21) |
|---|---|---|---|---|
| Kibaale | - | | | |
| Masindi | 0.002096 | - | | |
| Arua | 0.007788 | 0.006847 | - | |
| Apac | 0.002616 | 0.002748 | -0.003265 | - |

We used pruned dataset with only 119 samples representing distinct farmer-varieties randomly selected from within each district.

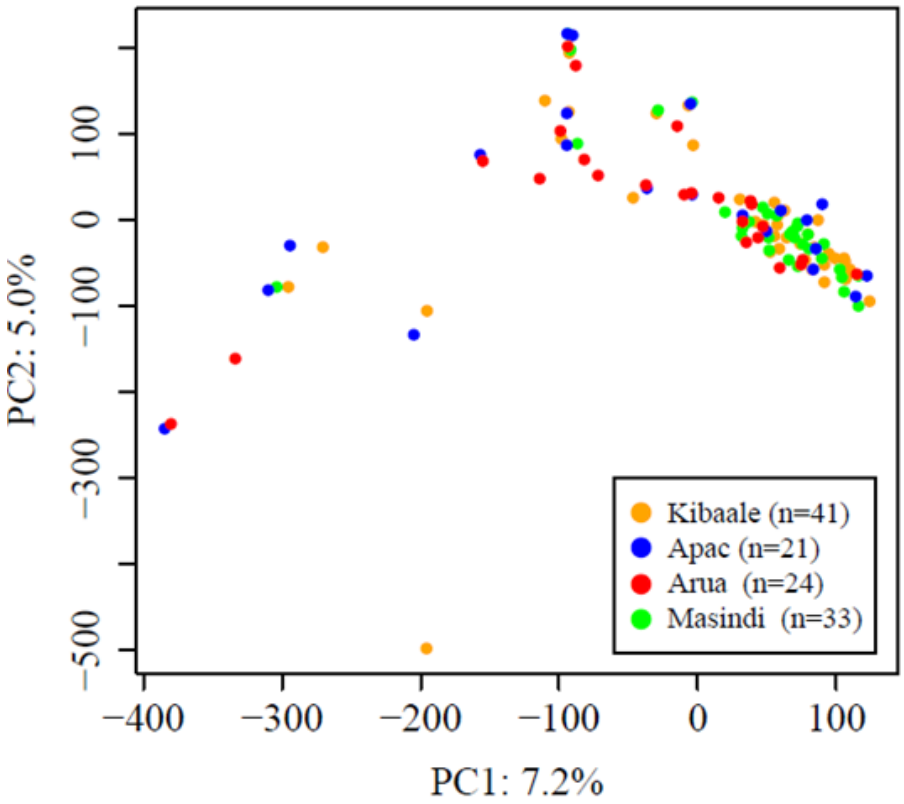



Figure 3.2. Principal components 1 and 2 based on 189,851 genome-wide SNPs scored on 119 genetically unique cassava varieties randomly selected from the clonal groups in each district after correction for multiple samples within the same clonal group (if a unique variety was represented in two or more districts, it had one entry for each of the districts) colored by district from where the samples were collected indicated minimal clustering in relation to the source of the sample.

***Correspondence between variety names reported by the farmers and their genetic identity***

Based on farmer-reported naming, 156 unique varieties were highlighted (Supplementary Table 4). Variety names ‘Gwalanda’, ‘Bukalasa’, ‘Bao’, and ‘Longe’ were the most common, accounting for 22.5% of the samples collected (Supplementary Table 4). Overall, farmer-reported variety names did not reliably correspond to genetically unique varieties as empirically revealed by SNP markers (Figure 3.3). For example, the largest clonal group (G_3), which was observed 80 times, included members that had up to 32 different

variety names assigned by farmers (Figure 3.3; Supplementary Table 3). However, there are instances when almost all farmer-reported variety name aligned within a given clonal group. For instance, 80% of the samples referred to as 'Gwalanda' had the same genetic identity (G_355) derived from IBS distinctions (Figure 3.3; Supplementary Table 3).

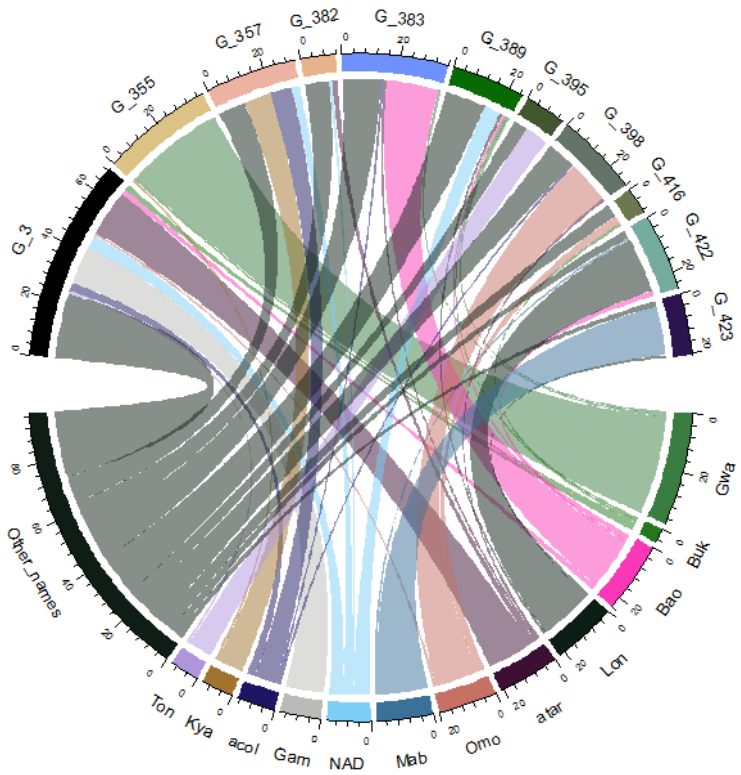



Figure 3.3. Correspondence between genetic identity (clonal groups labeled G_3 to G_423) (upper semi-circle) and most common variety names given by farmers (lower semi-circle). Clonal groups with more than 11 members were considered, and farmer-reported variety names that were mentioned more than 11 times in total are presented in this plot. The numbers on the axis represents the members in a given clonal group or the number of times a given variety name was mentioned during the surveys. The label for "Other_names" represent all other variety names that had a count less than 12 times. Ton represents Tonguda, Kya represents Kyawada, acol represents Gamente acol, Gam represent Gamente, NAD represents NAADs, Mab stands for Mabulu, Omo stands for Omoo, atar stands for Gamente atar, Lon stands for Longe, Buk stands for Bukalasa, Gwa stands for Gwalanda.

To further visualize how genetically identical varieties derived from IBS distinctions related to the respective farmer-reported variety names, the first two genetic principal components (PCs) were colored based on the most predominant (at least *n*=12) clonal groups (Figure 3.4a) and the most predominant (at least *n*=13) variety names reported by farmers during the survey (Figure 3.4b). The structuring pattern in the PC plots indicated that members of the same clonal group, grouped together as expected. However, when the same plot is colored based on variety names given by farmers, members with similar names often did not group together (Figure 3.4). Taken together, these results confirmed that a number of genetically unique varieties had multiple names reported by farmers. This phenomenon was observed both within and between districts (Supplementary Table 1).

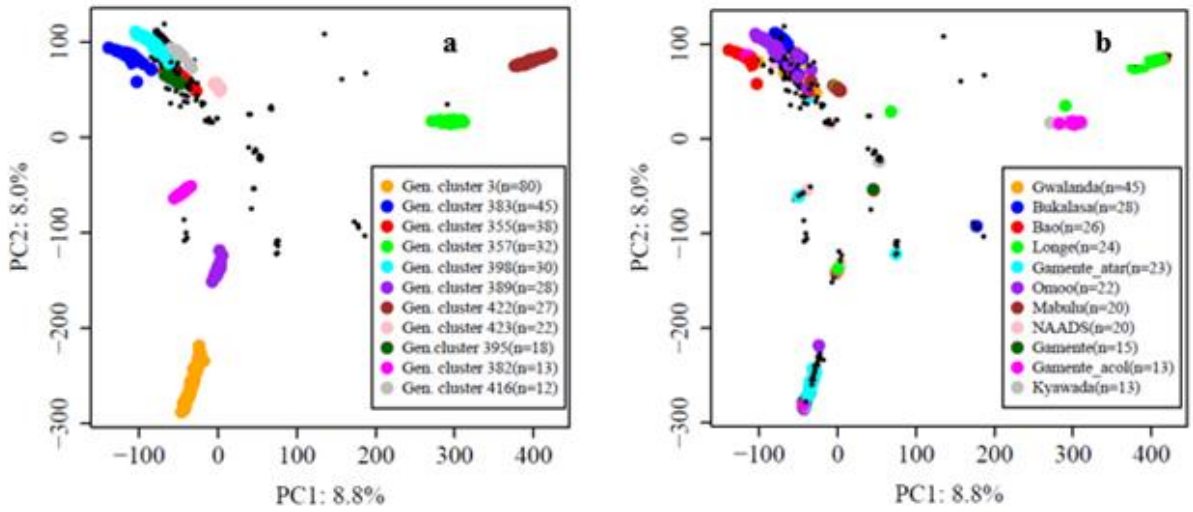



Figure 3.4. Principal component analysis plots of PC1 and PC2 based on 190,556 genome-wide SNPs scored on 547 samples colored by most common a). Clonal groups from IBS-based distinctions b) Names of the cassava varieties reported by farmers.

***Genetic relationships among farmer-grown cassava varieties and breeding lines in Uganda***

Our results based on IBS indicate that all cassava varieties cultivated by farmers in Apac, Arua, Kibaale, and Masindi districts are not clones of the elite breeding lines sourced

from NaCRRI (Supplementary Table 5). We also conducted a PCA to visualize how farmers' varieties related to breeding lines. In the genetic space described by the first four principal components, the farmer-varieties are a subset of the breeding lines (Figure 3.5). The percentage variation explained by each of the principal components (PCs) was relatively low (Supplementary Figure 5). Additionally, there was stronger clustering among farmer varieties as compared to breeding lines. It was also evident that there is more variation in breeding lines as opposed to the varieties cultivated by farmers in the four districts.

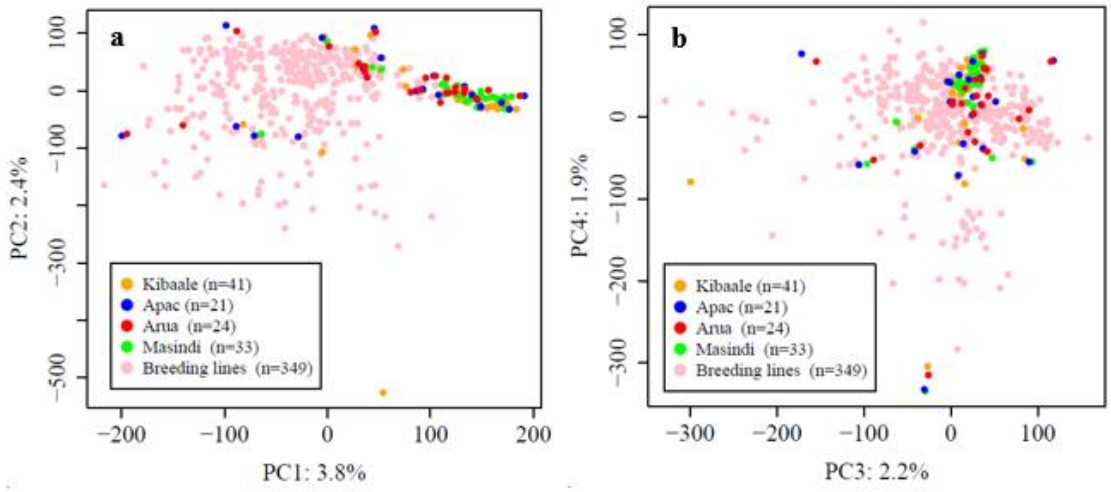



Figure 3.5. The principal component analysis plots of PC1 and PC2 (a), and PC3 and PC4 (b) based on 349 breeding lines and 119 unique farmers' varieties scored on 201,117 SNPs. The plots show patterns of structure between cassava varieties cultivated by smallholder farmers in four districts (Kibaale, Apac, Arua, and Masindi) in Uganda and the cassava breeding lines at NaCRRI, Uganda. Only unique varieties per district are used for the farmer varieties.

The $F_{ST}$ indicated low genetic differentiation (0.020) between farmers' varieties and breeding lines. The mean nucleotide diversity was generally higher among breeding lines than farmer-varieties (Figure 3.6). The highest ratio of nucleotide diversity of breeding lines to farmer-varieties was observed on chromosomes four, 18 and one respectively, while the least ratio was observed on chromosome nine.

Table 3.3. Average nucleotide diversity per chromosome for unique cassava varieties grown by smallholder farmers within four districts in Uganda.

| Chromosome | Apac_pi (n=21) | Arua_pi (n=24) | Kibaale_pi (n=41) | Masindi_pi (n=33) |
|---|---|---|---|---|
| 1 | 0.0001574 | 0.0001479 | 0.0001445 | 0.0001379 |
| 2 | 0.0001162 | 0.0001154 | 0.0001144 | 0.0001107 |
| 3 | 0.0001106 | 0.0001105 | 0.0001091 | 0.0001072 |
| 4 | 0.0001223 | 0.0001184 | 0.0001094 | 0.0001099 |
| 5 | 0.0001041 | 0.0001044 | 0.0001040 | 0.0001035 |
| 6 | 0.0001149 | 0.0001129 | 0.0001120 | 0.0001089 |
| 7 | 0.0000849 | 0.0000847 | 0.0000860 | 0.0000837 |
| 8 | 0.0000861 | 0.0000846 | 0.0000815 | 0.0000814 |
| 9 | 0.0000919 | 0.0000942 | 0.0000908 | 0.0000904 |
| 10 | 0.0001158 | 0.0001158 | 0.0001099 | 0.0001095 |
| 11 | 0.0001208 | 0.0001199 | 0.0001147 | 0.0001152 |
| 12 | 0.0000808 | 0.0000804 | 0.0000780 | 0.0000784 |
| 13 | 0.0000832 | 0.0000841 | 0.0000833 | 0.0000818 |
| 14 | 0.0001307 | 0.0001318 | 0.0001286 | 0.0001238 |
| 15 | 0.0001225 | 0.0001230 | 0.0001206 | 0.0001196 |
| 16 | 0.0000828 | 0.0000831 | 0.0000775 | 0.0000771 |
| 17 | 0.0000921 | 0.0000945 | 0.0000890 | 0.0000874 |
| 18 | 0.0000906 | 0.0000906 | 0.0000874 | 0.0000887 |
| Average | 0.0001060 | 0.0001053 | 0.0001023 | 0.0001008 |

n represent the number of unique varieties

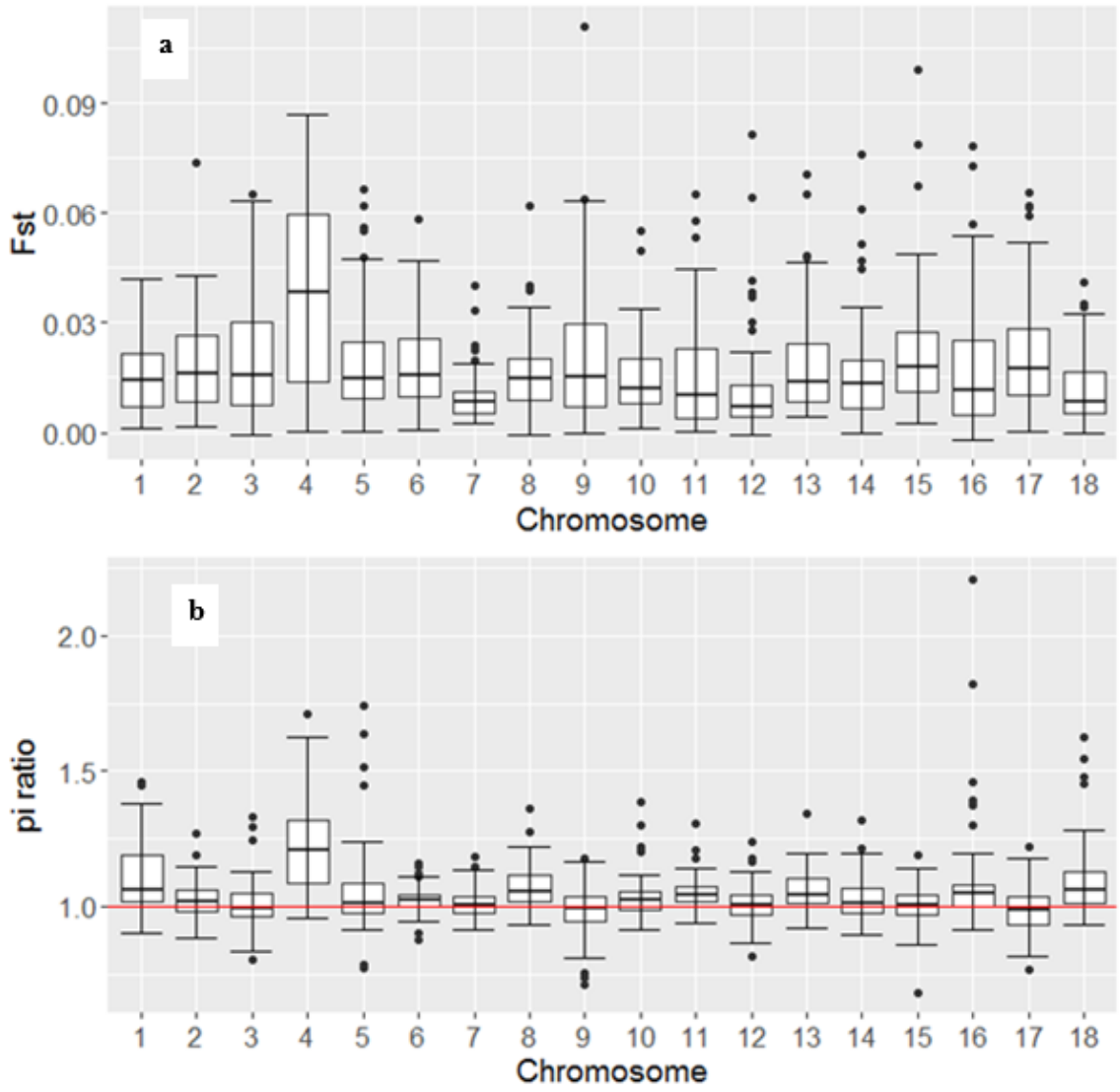



Figure 3.6. a) Genome-wide Fst between cassava breeding lines at National Crops Resources Research Institute (NaCRRI) and farmer-varieties, b) Nucleotide diversity (pi) ratio of cassava breeding lines at NaCRRI to varieties grown by smallholder farmers in Uganda.

**DISCUSSION**

Comparing genetic relationships among varieties adopted by farmers to those of breeding lines is important in developing new varieties that best meet the needs and preferences of the end-users. This study revealed that despite the low ($F_{ST}$ < 0.05) genetic differentiation among cassava varieties grown in different districts of Uganda, the varieties cultivated across different districts are often distinct genetically. Of the 547 samples collected

from Ugandan farmers, there were 86 genetically unique varieties. Of these unique varieties, most of them were cultivated in only one single district (30 in Kibaale, 19 in Masindi, 14 in Arua and three in Apac), while only three were cultivated across all the four districts. However, these unique varieties are likely to be close relatives, given the observed levels of genetic differentiation between districts.

In our study, we found that, in agreement with a recent study in Ghana (Rabbi et al., 2015), most smallholder farms cultivate two or more cassava varieties in the same field in order to meet the diverse needs of farmers *and* end-users (Nweke et al., 2002). Consequently, the different unique varieties could be serving different purposes both for the farmer (risk aversion, in case one variety or market fails) and for the consumer (processing, fresh consumption) (Nakabonge et al., 2018). For instance, during the interviews prior to leaf sample collection, some farmers mentioned that certain varieties were used as a source of food for the household members, while other varieties were largely for income generation. The genetic differentiation we observed among farming districts in Uganda (mean $F_{ST}$ = 0.003) is similar to that observed between the two breeding programs based within Nigeria (Wolfe et al. 2017, $F_{ST}$ 0.008) and lower than observed levels of differentiation between, for example, East and West Africa (Ramu et al., 2017; Wolfe et al., 2017), which range from 0.01-0.05. The levels of genetic differentiation observed in cassava populations may be due to the common practice of exchanging planting materials between neighboring farmers, friends, and relatives (Mtunguja et al., 2014). In addition, cassava is known to have a high outcrossing rate in the field (da Silva et al. 2003) and recombinant seed can establish in farmers' fields, be erroneously propagated and lead to new varieties that are closely related to what is already in production (Fregene et al., 2003; Duputie et al., 2009). Thus, continued gene flow within the continent is likely to be a significant factor in the limited population structure that has been observed.

In this study, we showed that farmer-reported variety names were not consistent with the genotype information. For example, the variety names Akena, Bao, Bukalasa, Gamente, Gotta, Kibaho, Mukuma, NAADS, and Olam that were assigned by the farmers were classified under the same genetic identity (clonal group 3). The implication of this result is that breeders should not solely rely on the farmer-given variety name in variety identification studies. This is in agreement with previous studies, which have also reported a large discrepancy between genetically unique varieties and the variety names assigned by farmers (Rabbi et al., 2015; Bredeson et al., 2016). Indeed, most farmers obtain cassava varieties from their neighboring farmers, relatives, and friends (Nweke et al., 2002; Teeken et al., 2018). The inconsistency between genotype and variety names is thus attributable to the lack of a regulated seed system with the ability to maintain genetic fidelity relative to germplasm names. A previous study revealed that naming of cassava varieties is subjective and may depend on many factors, such as the place of origin, maturity period, taste, morphology, yield, marketability, and resilience (Kizito et al 2007; Nakabonge et al., 2017). Indeed, we observed that some variety names (e.g., Bukalasa) refer to the place or source of its origin while others refer to phenotypes. For example, Gamente-acol may have been sourced from the government, as Gamente is the local language name for the government, and the last part of the name separated by a hyphen (acol) is derived from the color of stems that are mostly dark (acol means dark in the local language). Our results indicate that variety name alone is not reliable and should not be used to define unique cassava varieties in studies of adoption by local farming communities in Uganda or for the collection of farmer-varieties to be used in breeding. In a few scenarios, the samples with a similar variety name belonged to the same clonal group e.g. of the samples that were named as Gwalanda, 80% of them belonged to the same clonal group (Figure 3). All the samples named Gwalanda were collected from Kibaale

district, and one of the possibilities for the observed variation in naming pattern was due to the distinct morphological characteristics upon which the variety name was derived.

All breeding lines that we analyzed were genetically different from varieties cultivated by farmers though we did not comprehensively sample all breeding lines and those we analyzed are known not to have yet been released to the farmers. The differentiation we observed between breeding lines and farmers' varieties was similar to the level observed between breeding programs in Nigeria (Wolfe et al. 2017) and implies that both populations share a large number of alleles. In a previous study of Uganda farmer-preferred varieties, Turyagyenda et al. (2012) also found that genetic distance between landraces and breeding lines was small. Indeed, the genetic variability among breeding lines along the first four PCs (Figure 5) was greater than among the farmers' varieties, matching the observation that the breeding population is slightly more diverse (Figure 3.6, Supplementary Table 5) and similarly homozygous (Supplementary Figure 7). There was only one farmer-variety that was notably distinct (along PC2).

Chromosomes one and four appeared notably more diverse among the breeding lines compared to farmer varieties (Figure 6). Based in part on this result as well as those of Bredeson et al. (2016), we suspected that some of the farmer accessions might contain introgression segments from the wild relative *M. glaziovii*. Recently, Wolfe et al. (2019, *in review*) revealed that the introgressions on chromosomes one and four are common in breeding germplasm, and also present (but less common) in landraces. Based on introgression diagnostic markers, we found that the same was true of the difference between breeding lines and the farmer varieties we sampled in Uganda (Supplementary Table 5, Supplementary Figure 6; Wolfe et al. 2019, *in review*). Interestingly, the farmer-variety mentioned above as an outlier on PC2 (Figures 3.2 and 3.5) appears to be an $F_1$ (39.6 % introgression diagnostic alleles, mostly in the heterozygous state) hybrid between an *M. glaziovii* and an *M. esculenta*

parent. The passport data we collected from the farmer indicates that it had very bitter roots and leaves relative to other cassava varieties and that it was mainly used as a border row to deter thieves and animals from the main crop. This kind of information highlights the multiple functions of cassava varieties grown by farmers *and* the value of genetic surveys of farmer-preferred varieties.

## CONCLUSION

In this study, cassava leaf samples collected from 547 different cultivated plants grown by smallholder farmers within four districts in Uganda were genotyped with the major objective of understanding the genetic relationship among varieties grown by the farmers. We also explored the genetic relationship between these surveyed farmer varieties and breeding lines used at NaCRRI. We found that most farmers in Uganda grow two or three distinct cassava varieties and that each sampled district in Uganda contains several varieties not grown in other districts. The overall level of genetic differentiation between districts is relatively low, as is the divergence between farmer and breeding populations. Despite the fact that none of the breeding lines were directly observed in farmer fields, the presence of *M. glaziovii* introgressions on chromosomes one and four implies ancestry with cassava breeding lines. Given the apparently similar pool of alleles in the breeding germplasm, it is likely that breeders have the raw genetic material they require to match the farmer-preferred trait combinations necessary for adoption. Our study highlights the importance of understanding the genetic make-up of cassava currently grown by smallholder farmers and relative to that of plant breeding germplasm.

### *Conflict of interest*

The authors declare that they have no conflict of interest.

## *ACKNOWLEDGEMENT*

This research was undertaken as part of, and funded by, the CGIAR Research Program on Roots, Tubers and Bananas (RTB) and supported by CGIAR Trust Fund contributors. Additional funds for this work was provided by the International Programs of the College of Agriculture and Life Sciences at Cornell University through the NextGen Cassava Grant (OPP1048542) funded by the Bill and Melinda Gates Foundation and the Department for International Development of the United Kingdom (UKaid). We give special thanks to Mr. Francis Osingada and other technical staff of NaCRRI Root Crops Program who assisted in the collection of leaf tissue and genomic DNA extraction. We also acknowledge technical advice from Prof. Rachel Bezner Kerr pertaining the design of the study.

**REFERENCES**

Acheampong, P.P., V. Owusu, and G. Nurah. 2018. How does Farmer Preference matter in Crop variety Adoption? The case of Improved Cassava varieties' Adoption in Ghana. *Open Agriculture* 3: 466–477. doi: 10.1515/opag-2018-0052.

Acquaah, G. 2012. Principles of plant genetics and breeding. Second Edi. John Wiley and Sons, Inc., West Sussex, UK.

Alene, A.D., R. Khataza, C. Chibwana, P. Ntawuruhunga, and C. Moyo. 2013. Economic impacts of cassava research and extension in Malawi and Zambia. *Journal of Development and Agricultural Economics* 5(11): 457–469. doi: 10.5897/JDAE2013.0496.

Bechoff, A., K. Tomlins, G. Fliedel, L.A. Becerra, A. Westby, C. Hershey, D. Dufour, A. Bechoff, K. Tomlins, G. Fliedel, L.A. Becerra, C. Hershey, and D. Dufour. 2016. Cassava traits and end-user preference : Relating traits to consumer liking , sensory perception , and genetics. *Critical Reviews in Food Science and Nutrition* 0(0): 1–21. doi: 10.1080/10408398.2016.1202888.

Bezner Kerr, R. 2017. Gender and Agrarian inequalities. p. 333–370. *In* Snapp, S., Pound, B. (eds.), Agricultural Systems: Agroecology and Rural Innovations for Development. Second Edi. San Diego, CA, USA.

Bredeson, J. V, J.B. Lyons, S.E. Prochnik, G.A. Wu, C.M. Ha, E. Edsinger-gonzales, J. Grimwood, J. Schmutz, I.Y. Rabbi, C. Egesi, P. Nauluvula, V. Lebot, J. Ndunguru, G. Mkamilo, R.S. Bart, T.L. Setter, R.M. Gleadow, P. Kulakow, M.E. Ferguson, S. Rounsley, and D.S. Rokhsar. 2016. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. 34(5): 562–571. doi: 10.1038/nbt.3535.

Cardoso, A.P., E. Mirione, M. Ernesto, F. Massaza, J. Cliff, M. Rezaul Haque, and J.H. Bradbury. 2005. Processing of cassava roots to remove cyanogens. *Journal of Food Composition and Analysis* 18: 451–460. doi: 10.1016/j.jfca.2004.04.002.

Chiwona-Karltun, L., J. Mkumbira, J. Saka, M. Bovin, N.M. Mahungu, and H. Rosling. 1998. The importance of being bitter — a qualitative study on cassava cultivar preference in

Malawi. *Ecology of Food and Nutrition* 37(3): 219–245. doi: 10.1080/03670244.1998.9991546.

Chiwona-Karltun, L., D. Nyirenda, C.N. Mwansa, J.E. Kongor, L. Brimer, S. Haggblade, and E.O. Afoakwa. 2015. Farmer preference, utilization, and biochemical composition of improved Cassava (Manihot esculenta Crantz) varieties in Southeastern Africa. *Economic Botany* 69(1): 42–56. doi: 10.1007/s12231-015-9298-7.

Doss, C.R., and M.L. Morris. 2001. How does gender affect the adoption of agricultural innovations? The case of improved maize technology in Ghana. *Agricultural Economics* 25(2001): 27–39.

El-Sharkawy, M.A. 2004. Cassava biology and physiology. *Plant Molecular Biology* 56: 481–501.

FAO. 2013a. Save and grow cassava: A guide to sustainable production intestification. Rome.

FAO. 2013b. STANDARD FOR SWEET CASSAVA: CODEX STAN 238-2003.

Forsythe, L., H. Posthumus, and A. Martin. 2016. A crop of one's own? Women's experiences of cassava commercialization in Nigeria and Malawi. *Gender, Agriculture and Food security* 1(2): 110–128.

JHSPH. 2009. Johns Hopkins School of Public Health (JHSPH) (2009) Human Subjects Research Ethics Field Training Guide.

Kawuki, R.., T. Amuge, E. Nuwamanya, G. Semakula, S. Tumwesigye, A. Bua, Y. Baguma, C. Omongo, T. Alicai, and J. Orone. 2011. A breeding scheme for local adoption of cassava (Manihot esculenta Crantz). *Plant Breeding and Crop Science* 3(7): 120–130.

Kizito, B.E., L. Chiwona-Karltun, T. Egwang, M. Fregene, and A. Wasterbergh. 2007. Genetic diversity and variety composition of cassava on small- scale farms in Uganda : an interdisciplinary study using genetic markers and farmer interviews. *Genetica* 130: 301–318. doi: 10.1007/s10709-006-9107-4.

Kombo, G.R., A. Dansi, L.Y. Loko, G.C. Orkwor, R. Vodouhè, P. Assogba, and J.M. Magema. 2012. Diversity of cassava (Manihot esculenta Crantz) cultivars and its management in the department of Bouenza in the Republic of Congo. *Genetic Resources*

*and Crop Evolution* 59(8): 1789–1803. doi: 10.1007/s10722-012-9803-0.

Lebot, V. 2009. Tropical root and tuber crops: cassava, sweet potato, yams and aroids. Crop production science in horticulture (17). CABI, Wallingford, UK.

Masamha, B., V. Thebe, and V.N.E. Uzokwe. 2018. Mapping cassava food value chains in Tanzania's smallholder farming sector: The implications of intra-household gender dynamics. *Journal of Rural Studies* 58(2018): 82–92. doi: 10.1016/j.jrurstud.2017.12.011.

Montagnac, J.A., C.R. Davis, and S.A. Tanumihardjo. 2009. Processing techniques to reduce toxicity and antinutrients of Cassava for use as a staple food. *Comprehensive Reviews in Food Science and Food Safety* 8(1): 17–27. doi: 10.1111/j.1541-4337.2008.00064.x.

Nakabonge, G., C. Samukoya, and Y. Baguma. 2018. Local varieties of cassava: Conservation, cultivation and use in Uganda. *Environment, Development and Sustainability* 20: 2427–2445. doi: 10.1007/s10668-017-9997-6.

Nweke, I.F., S.C.D. Spencer, and K.J. Lynam. 2002. The cassava transformation: Africa's best kept secret. Michigan State University Press, East Lansing, Michigan.

Ogunlela, Y.I., and A.A. Mukhtar. 2009. Gender Issues in Agriculture and Rural Development in Nigeria : The Role of Women. *Humanity and Social Sciences* 4(1): 19–30.

Otim-Nape, G.W., A. Bua, G. Ssemakula, G. Acola, Y. Baguma, S. Ogwal, and R. Van der Grift. 2005. Cassava development in Uganda. A country case study towards a global cassava development strategy. *In* A review of cassava in Africa with country case studies in Nigeria, Ghana, the United Republic of Tanzania, Uganda and Benin. Rome.

R Core Team. 2018. A language and environment for statistical computing. R foundation for statistical computing. Vienna, Austria.

Silver, D. 2019. Alienation in a four factor world. *Journal for the Theory of Social Behaviour* 49: 84–105. doi: 10.1111/jtsb.12192.

Teeken, B., O. Olaosebikan, J. Haleegoah, E. Oladejo, T. Madu, A. Bello, E. Parkes, C. Egesi, P. Kulakow, H. Kirscht, and H.A. Tufan. 2018. Cassava Trait Preferences of Men

and Women Farmers in Nigeria: Implications for Breeding. *Economic Botany*: 1–15. doi: 10.1007/s12231-018-9421-7.

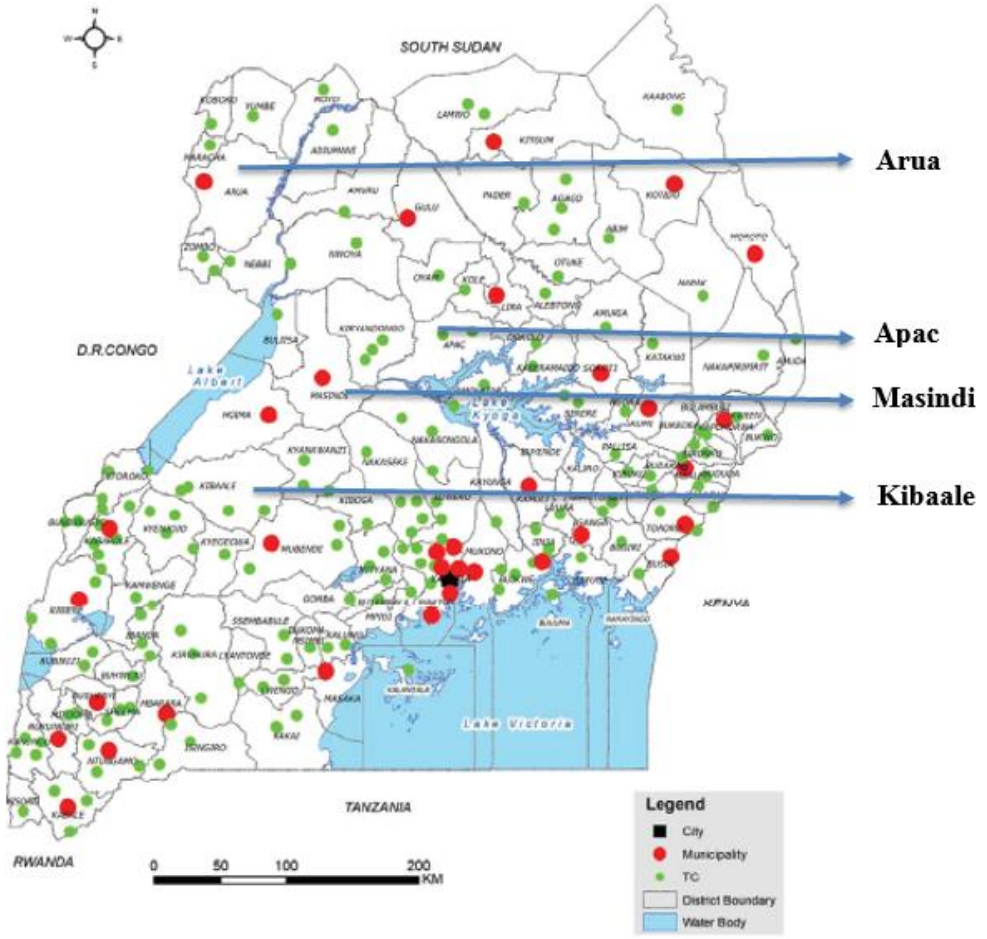
Tumuhimbise, R., R. Melis, P. Shanahan, and R. Kawuki. 2012. Farmers' Perceptions on Early Storage Root Bulking in Cassava (Manihot esculenta Crantz) in East and Central Uganda and their Implication for Cassava Breeding. *World Journal of Agricultural Sciences* 8(4): 403–408. doi: 10.5829/idosi.wjas.2012.8.4.1667.

Turyagyenda, L.F., E.B. Kizito, M.. Ferguson, Y. Baguma, P. Gibson, B.W. Wanjala, and D.S.. Osiru. 2012. Genetic diversity among farmer-preferred cassava landraces in Uganda. *African crop science journal* 20(1): 15–30. doi: 10.4314/acsj.v20i1.

UBOS. 2010. Uganda census of agriculture: Crop area and production report. Kampala, Uganda.

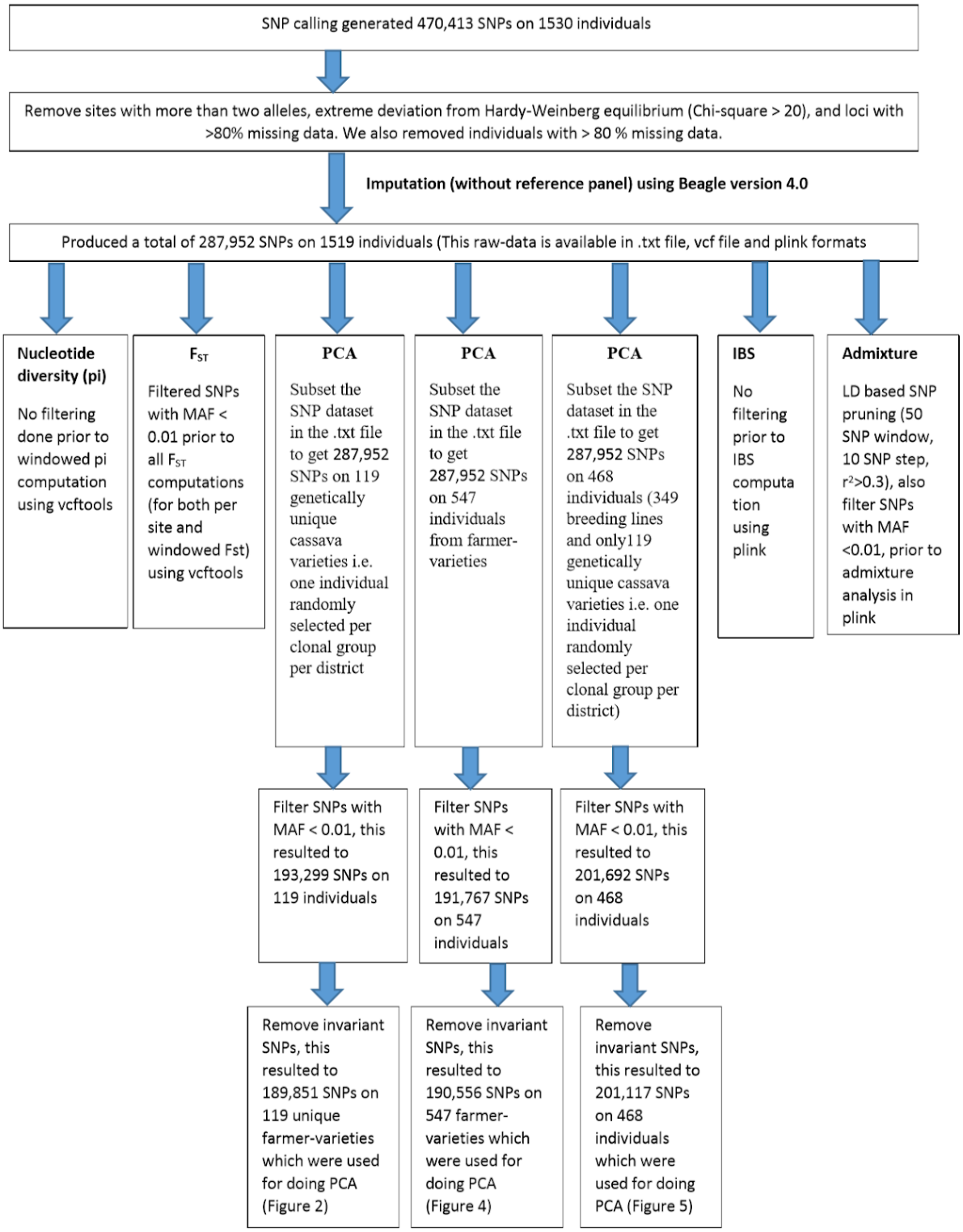
Weltzien, E., and A. Christinck. 2017. Participatory Breding: Developing Improved and Relevant Crop Varieties with Farmers. p. 259–301. *In* Snapp, S., Pound, B. (eds.), Agricultural Systems: Agroecology and Rural Innovations for Development. Second Edi. Elsevier Academic Press, San Diego, CA, USA.

**SUPPLEMENTARY MATERIAL**

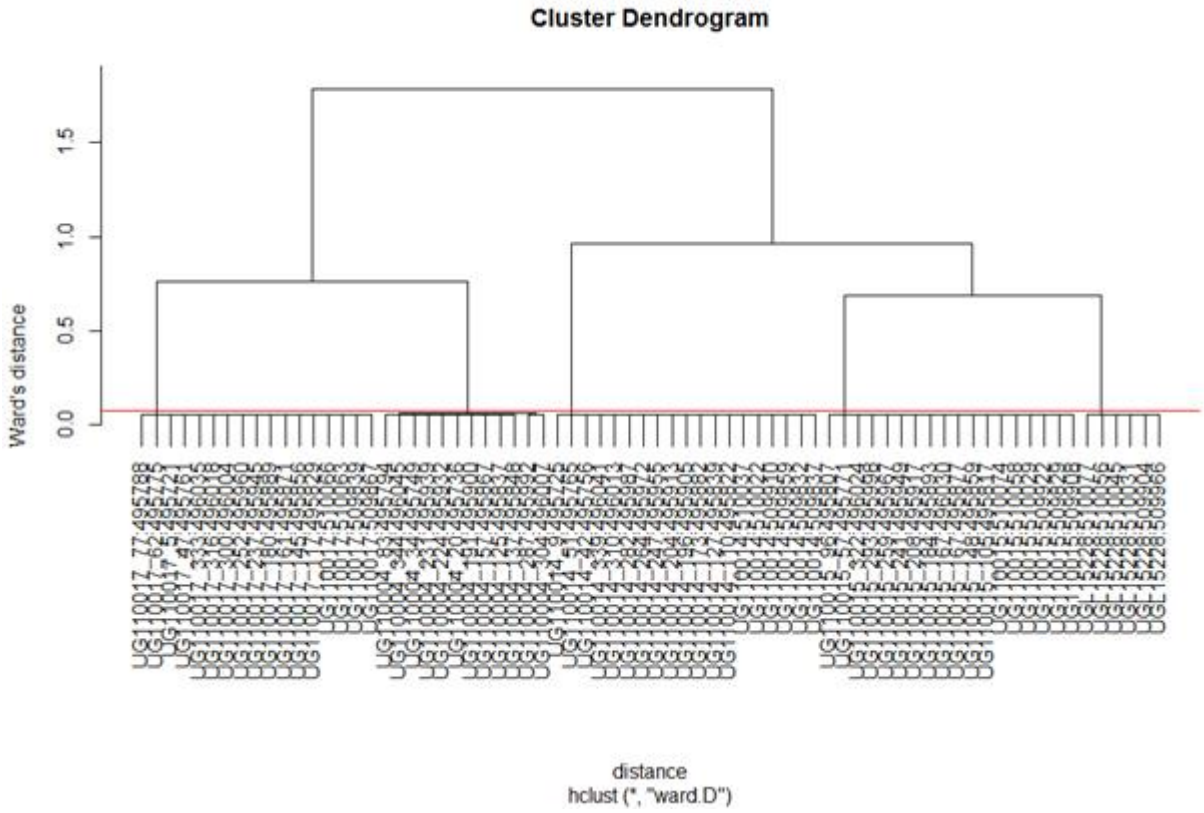



Source: UBOS (2016)

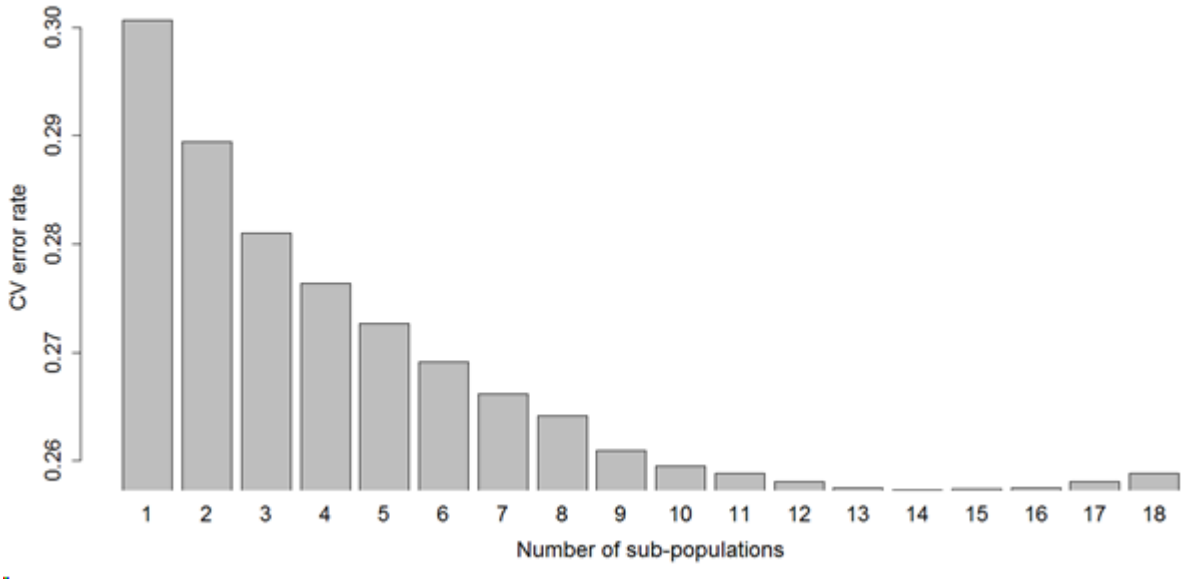
Supplementary Figure 1. Map of Uganda highlighting four districts from where cassava leaf samples were collected.



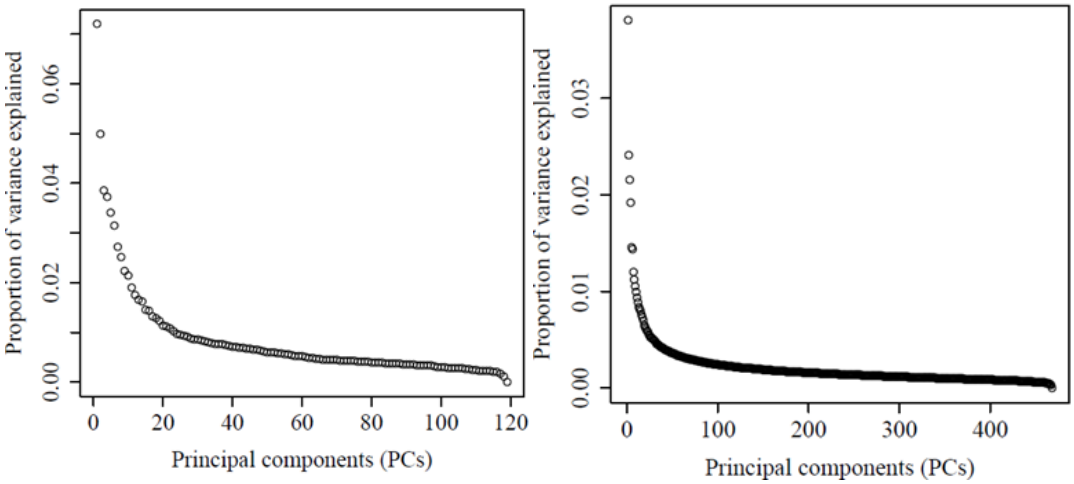

Supplementary Figure 2: Flowchart indicating the processing pipeline that was done on the genotype data used for the statistical analyses in Chapter 3.



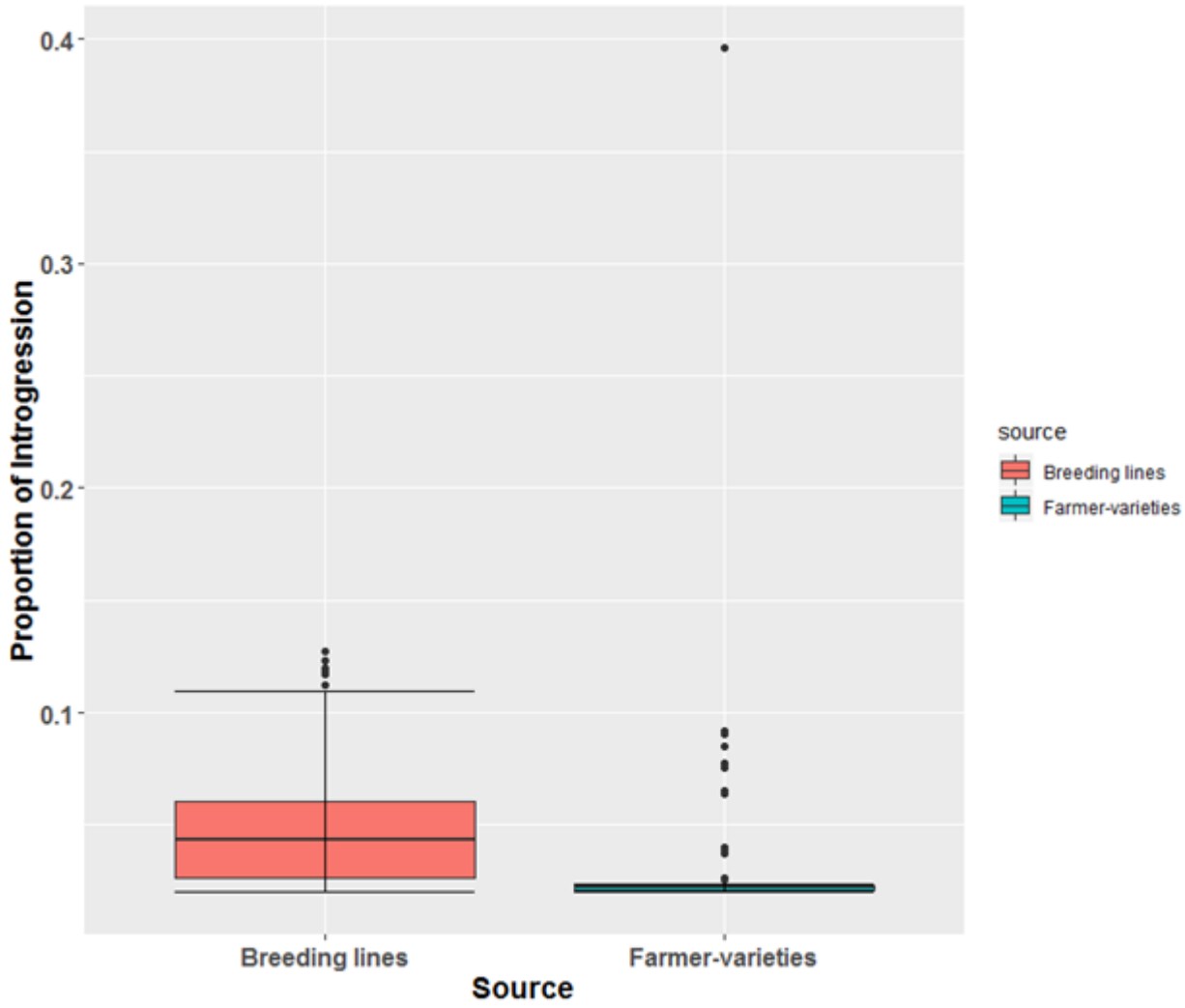

Supplementary Figure 3. Dendrogram showing clustering of biological replicated samples of five known genotypes (UG110017, UG110004, UG110014, UG110015, and UGL15228) that were used to declare the threshold for identifying unique varieties from samples collected from cassava varieties grown by smallholder farmers in Uganda.



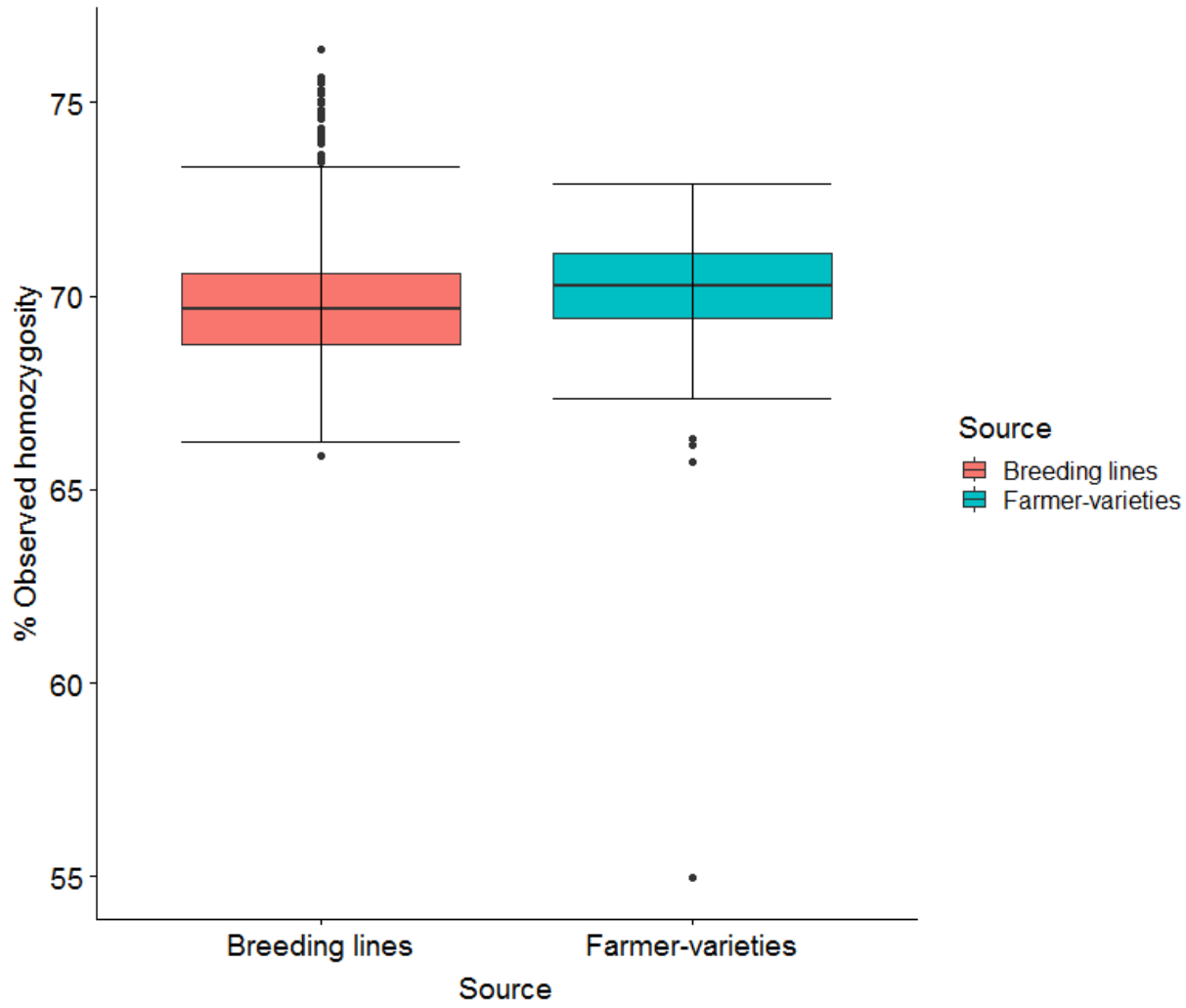

Supplementary Figure 4: Estimated five-fold cross-validation (CV) error rate of possible sub-populations (*K*) from *K* = 1 to 18 for population structure analysis using ADMIXTURE on 547 samples collected from cassava varieties grown by smallholder farmers in Uganda. The lowest CV error rate was observed at *K* =14.




Supplementary Figure 5: Scree plots indicating the proportion of variance explained by the principal components from PCA: a) Scree plot for the 119 unique farmer-varieties at district level b) Scree plot for the 119 unique farmer-varieties at district level and 349 breeding lines




Supplementary Figure 6: Boxplot indicating the overall proportion of *Manihot glaziovii* introgressed into the NaCRRI breeding lines and the farmer-varieties grown in Uganda. The farmer-varieties are subset of randomly selected unique individuals per clonal group per district.




Supplementary Figure 7: The level of observed homozygosity of the NaCRRI breeding lines and the farmer-varieties grown in Uganda. The farmer-varieties are subset of randomly selected unique individuals per clonal group per district.

SUPPLEMENTARY TABLES

Supplementary Table 1: Per sample summary of ancestries and IBS-based genetic cluster for the 547 farmer varieties

Supplementary Table 2: Unique varieties per district based on identity-by-state distinctions

Supplementary Table 3: Abundance of varieties cultivated by farmers from IBS-based genetic distinctions

Supplementary Table 4: Cassava variety names reported by farmers from four districts in Uganda

Supplementary Table 5: Unique varieties per district and from breeding lines from identity-by-state based genetic distinctions

***SUPPLEMENTARY FILES***

Supplementary File 1: Interview guide for examination of gender disaggregated cassava trait preference of smallholder farmers in Uganda

# CHAPTER 4 GENERAL CONCLUSIONS

The ultimate importance of plant breeding is crop improvement in order to satisfy the needs of the end-users (Acquaah, 2012). Despite the significant amount of contribution by plant breeders towards crop improvement, there have been incidences where adoption of the improved varieties have remained low due to lack of attributes that meet the end-user needs and preferences (Alene et al., 2013; Weltzien and Christinck, 2017). The work presented in this dissertation was done to contribute towards breeding cassava varieties that meet the needs and preferences of smallholder farmers in Uganda. Our findings from study one indicated that smallholder farmers like cassava yield, agronomic, and quality attributes. Some of the attributes had significant differences across districts, age categories, and sex. The implication of this geographically bound gender and age related differences in farmer preferences would require farmer-participatory breeding approaches in the target geographical zones in order to produce varieties that are acceptable to all the target end-users. In real world, it remains hard to produce an ideal variety that meets the needs of all end-users, thus there is need for breeding different varieties that meet the needs of varying end-users. Consequently, plant breeders need a clear understanding of the needs.

One of the novelties brought about by findings presented in this dissertation is the presentation of results disaggregated by sex and age. Previous studies on cassava trait preferences in Uganda reported their results based on different geographical locations (Kizito et al., 2007; Tumuhimbise et al., 2012; Nakabonge et al., 2018). As reported in study one, within a given locality, men and women may have different preferences, and also younger and older farmers may prefer varying attributes, hence the need to understand the needs of the target end-users in order to be able to design appropriate breeding methodologies. The attributes that had highest count based on the number of times they were mentioned to be

liked by the farmers could be used in determining the priority attributes to focus on during breeding, for example, the predominant attributes could be used to develop a selection index during the breeding pipeline. Different weights can be assigned based on the percentage of farmers that prefer given attributes. Existing gender relations in cassava production and post-harvest handling can strengthen the further understanding of the observed differences in cassava attributes preferred by men and women in various locations (Teeken et al., 2018). An example of such relations could be the division of labor between men and women, there is need to understand who does what and the likely consequences of introducing new varieties. It is also worth noting that some of the varieties liked by farmers had both positive and negative attributes, for instance most varieties mentioned to be early maturing had a negative attribute of short in-ground storage of roots. There is need to explore the genetic correlations of some of these attributes in order to be able to design appropriate techniques to break the associations between desirable and undesirable attributes.

The second major contribution from the work presented in this dissertation is the determination of broad-sense heritability estimates for softness of cooked cassava roots. Softness of cooked cassava roots is one of the key traits of the end-users who eat cassava in the boiled forms. Moderate broad-sense heritability estimate was obtained suggesting that softness of cooked cassava roots is an attribute that can be improved through breeding. Furthermore, we determined the relationship between consumer-testing and penetrometer methods for phenotyping softness of cooked cassava roots. Our results indicated a strong positive correlation between the two methods suggesting that a penetrometer could give comparable results to those of consumer-testing when used in breeding work to measure softness. This finding is good news to breeders because a penetrometer can be used to phenotype cooked cassava root softness in early breeding stages when breeders are still working with large number of samples. Contrary, consumer-testing method can only be

effective with fewer number of samples. Phenotyping for softness earlier in the breeding process will likely increase the number of elite lines with softness traits that get advanced to later generations and ultimately get released to farmers. If breeders are able to effectively breed for quality traits that meet the end-user needs and preferences, this will likely improve adoption of the new varieties. On the other hand, there is still need to focus on more advanced high throughput phenotyping technologies such as use of Near-infrared Spectroscopy (NIRS) for measuring cassava root quality traits to increase the efficiency of phenotyping. Additionally, due to limitation of resources in terms of time and funds we were not able to make crosses between varieties that have roots that get soft upon cooking and the varieties whose roots remain hard after cooking but have other desirable traits liked by farmers such as high yield and early maturity. We recommend the breeding program in Uganda undertakes such studies in an effort to develop breeding pipeline with a long term objective of improving the existing germplasm to better suit the needs of end-users.

Lastly, the work presented in this research contributes to the general understanding of genetic relations of cassava grown by the smallholder farmers in Uganda and the breeding used by national breeding program at NaCRRI. Our results revealed low genetic differentiation between cassava varieties grown by smallholder farmers in different districts in Uganda. Similar trend was found between cassava varieties grown by smallholder farmers and the breeding lines. Additionally, there was low nucleotide diversity found among both farmer-varieties and the breeding lines. Given the close relationship between the locally adopted varieties and the breeding lines, there are some chances that some of the improved varieties may have some of the attributes of the locally adopted materials, and this will likely increase the chance of being acceptable. On the other hand, for long-term genetic improvement of cassava, the breeders in Uganda need to find other sources of variation from other breeding programs to enhance the genetic base of their breeding germplasm. The

variation in nucleotide diversity of farmer-varieties and breeding lines could be due to the introgressions from wild cultivars (*Manihot glaziovii)* to cultivated cassava (*Manihot esculenta)* made on chromosome one and chromosome four (Bredeson et al., 2016). Some of the limitations of our diversity study were the smaller number of farmers that participated in the study and fewer districts where the study was conducted. It is recommended that the number of participants be increased and that sampling of other districts takes place to obtain a broader picture of the available diversity of the cassava grown in Uganda. We also recommend undertaking other genetic studies such as genome-wide association studies to understand the genetic architecture controlling the cassava traits preferred by end-users.

**REFERENCES**

Acheampong, P.P., V. Owusu, and G. Nurah. 2018. How does Farmer Preference matter in Crop variety Adoption? The case of Improved Cassava varieties' Adoption in Ghana. *Open Agriculture* 3: 466–477. doi: 10.1515/opag-2018-0052.

Acquaah, G. 2012. Principles of plant genetics and breeding. Second Edi. John Wiley and Sons, Inc., West Sussex, UK.

Alene, A.D., R. Khataza, C. Chibwana, P. Ntawuruhunga, and C. Moyo. 2013. Economic impacts of cassava research and extension in Malawi and Zambia. *Journal of Development and Agricultural Economics* 5(11): 457–469. doi: 10.5897/JDAE2013.0496.

Bechoff, A., K. Tomlins, G. Fliedel, L.A. Becerra, A. Westby, C. Hershey, D. Dufour, A. Bechoff, K. Tomlins, G. Fliedel, L.A. Becerra, C. Hershey, and D. Dufour. 2016. Cassava traits and end-user preference : Relating traits to consumer liking , sensory perception , and genetics. *Critical Reviews in Food Science and Nutrition* 0(0): 1–21. doi: 10.1080/10408398.2016.1202888.

Bezner Kerr, R. 2017. Gender and Agrarian inequalities. p. 333–370. *In* Snapp, S., Pound, B. (eds.), Agricultural Systems: Agroecology and Rural Innovations for Development. Second Edi. San Diego, CA, USA.

Bredeson, J. V, J.B. Lyons, S.E. Prochnik, G.A. Wu, C.M. Ha, E. Edsinger-gonzales, J. Grimwood, J. Schmutz, I.Y. Rabbi, C. Egesi, P. Nauluvula, V. Lebot, J. Ndunguru, G. Mkamilo, R.S. Bart, T.L. Setter, R.M. Gleadow, P. Kulakow, M.E. Ferguson, S. Rounsley, and D.S. Rokhsar. 2016. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. 34(5): 562–571. doi: 10.1038/nbt.3535.

Cardoso, A.P., E. Mirione, M. Ernesto, F. Massaza, J. Cliff, M. Rezaul Haque, and J.H. Bradbury. 2005. Processing of cassava roots to remove cyanogens. *Journal of Food Composition and Analysis* 18: 451–460. doi: 10.1016/j.jfca.2004.04.002.

Chiwona-Karltun, L., J. Mkumbira, J. Saka, M. Bovin, N.M. Mahungu, and H. Rosling. 1998. The importance of being bitter — a qualitative study on cassava cultivar preference in

Malawi. *Ecology of Food and Nutrition* 37(3): 219–245. doi: 10.1080/03670244.1998.9991546.

Chiwona-Karltun, L., D. Nyirenda, C.N. Mwansa, J.E. Kongor, L. Brimer, S. Haggblade, and E.O. Afoakwa. 2015. Farmer preference, utilization, and biochemical composition of improved Cassava (Manihot esculenta Crantz) varieties in Southeastern Africa. *Economic Botany* 69(1): 42–56. doi: 10.1007/s12231-015-9298-7.

Doss, C.R., and M.L. Morris. 2001. How does gender affect the adoption of agricultural innovations? The case of improved maize technology in Ghana. *Agricultural Economics* 25(2001): 27–39.

El-Sharkawy, M.A. 2004. Cassava biology and physiology. *Plant Molecular Biology* 56: 481–501.

FAO. 2013a. Save and grow cassava: A guide to sustainable production intestification. Rome.

FAO. 2013b. STANDARD FOR SWEET CASSAVA: CODEX STAN 238-2003.

Forsythe, L., H. Posthumus, and A. Martin. 2016. A crop of one's own? Women's experiences of cassava commercialization in Nigeria and Malawi. *Gender, Agriculture and Food security* 1(2): 110–128.

JHSPH. 2009. Johns Hopkins School of Public Health (JHSPH) (2009) Human Subjects Research Ethics Field Training Guide.

Kawuki, R.., T. Amuge, E. Nuwamanya, G. Semakula, S. Tumwesigye, A. Bua, Y. Baguma, C. Omongo, T. Alicai, and J. Orone. 2011. A breeding scheme for local adoption of cassava (Manihot esculenta Crantz). *Plant Breeding and Crop Science* 3(7): 120–130.

Kizito, B.E., L. Chiwona-Karltun, T. Egwang, M. Fregene, and A. Wasterbergh. 2007. Genetic diversity and variety composition of cassava on small- scale farms in Uganda : an interdisciplinary study using genetic markers and farmer interviews. *Genetica* 130: 301–318. doi: 10.1007/s10709-006-9107-4.

Kombo, G.R., A. Dansi, L.Y. Loko, G.C. Orkwor, R. Vodouhè, P. Assogba, and J.M. Magema. 2012. Diversity of cassava (Manihot esculenta Crantz) cultivars and its management in the department of Bouenza in the Republic of Congo. *Genetic Resources*

*and Crop Evolution* 59(8): 1789–1803. doi: 10.1007/s10722-012-9803-0.

Lebot, V. 2009. Tropical root and tuber crops: cassava, sweet potato, yams and aroids. Crop production science in horticulture (17). CABI, Wallingford, UK.

Masamha, B., V. Thebe, and V.N.E. Uzokwe. 2018. Mapping cassava food value chains in Tanzania's smallholder farming sector: The implications of intra-household gender dynamics. *Journal of Rural Studies* 58(2018): 82–92. doi: 10.1016/j.jrurstud.2017.12.011.

Montagnac, J.A., C.R. Davis, and S.A. Tanumihardjo. 2009. Processing techniques to reduce toxicity and antinutrients of Cassava for use as a staple food. *Comprehensive Reviews in Food Science and Food Safety* 8(1): 17–27. doi: 10.1111/j.1541-4337.2008.00064.x.

Nakabonge, G., C. Samukoya, and Y. Baguma. 2018. Local varieties of cassava: Conservation, cultivation and use in Uganda. *Environment, Development and Sustainability* 20: 2427–2445. doi: 10.1007/s10668-017-9997-6.

Nweke, I.F., S.C.D. Spencer, and K.J. Lynam. 2002. The cassava transformation: Africa's best kept secret. Michigan State University Press, East Lansing, Michigan.

Ogunlela, Y.I., and A.A. Mukhtar. 2009. Gender Issues in Agriculture and Rural Development in Nigeria : The Role of Women. *Humanity and Social Sciences* 4(1): 19–30.

Otim-Nape, G.W., A. Bua, G. Ssemakula, G. Acola, Y. Baguma, S. Ogwal, and R. Van der Grift. 2005. Cassava development in Uganda. A country case study towards a global cassava development strategy. *In* A review of cassava in Africa with country case studies in Nigeria, Ghana, the United Republic of Tanzania, Uganda and Benin. Rome.

R Core Team. 2018. A language and environment for statistical computing. R foundation for statistical computing. Vienna, Austria.

Silver, D. 2019. Alienation in a four factor world. *Journal for the Theory of Social Behaviour* 49: 84–105. doi: 10.1111/jtsb.12192.

Teeken, B., O. Olaosebikan, J. Haleegoah, E. Oladejo, T. Madu, A. Bello, E. Parkes, C. Egesi, P. Kulakow, H. Kirscht, and H.A. Tufan. 2018. Cassava Trait Preferences of Men

and Women Farmers in Nigeria: Implications for Breeding. *Economic Botany*: 1–15. doi: 10.1007/s12231-018-9421-7.

Tumuhimbise, R., R. Melis, P. Shanahan, and R. Kawuki. 2012. Farmers' Perceptions on Early Storage Root Bulking in Cassava (Manihot esculenta Crantz) in East and Central Uganda and their Implication for Cassava Breeding. *World Journal of Agricultural Sciences* 8(4): 403–408. doi: 10.5829/idosi.wjas.2012.8.4.1667.

Turyagyenda, L.F., E.B. Kizito, M.. Ferguson, Y. Baguma, P. Gibson, B.W. Wanjala, and D.S.. Osiru. 2012. Genetic diversity among farmer-preferred cassava landraces in Uganda. *African crop science journal* 20(1): 15–30. doi: 10.4314/acsj.v20i1.

UBOS. 2010. Uganda census of agriculture: Crop area and production report. Kampala, Uganda.

Weltzien, E., and A. Christinck. 2017. Participatory Breding: Developing Improved and Relevant Crop Varieties with Farmers. p. 259–301. *In* Snapp, S., Pound, B. (eds.), Agricultural Systems: Agroecology and Rural Innovations for Development. Second Edi. Elsevier Academic Press, San Diego, CA, USA.