



Genetic diversity of cassava landraces and documentation of farmer's knowledge in Lamu, Kenya

Christine K. Mutoni · Felister M. Nzuve · Douglas W. Miano · Benjamin M. Kivuva · Ivan J. Obare · Trushar M. Shah · Morag E. Ferguson

Received: 21 May 2023 / Accepted: 11 August 2023 / Published online: 31 October 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract To develop a strategy for improvement of cassava (*Manihot esculenta* Crantz) productivity for climate change preparedness in Lamu County, Kenya, it was important to understand which cultivars farmers are growing. Landraces should be collected and conserved prior to replacement by improved cultivars to avoid erosion of genetic resources and associated farmer knowledge. In this study, 58 samples of landraces were collected from five administrative wards in Lamu County. These were SNP genotyped using low-density Diversity Array Technologies (DArT) DaRTseq™ together with a reference set of 48 breeding lines, eight coastal and 45 inland landraces to aid identification of the collected samples. Genotyping results defined nine clades of duplicates and six

unique clones. The most representative clone was selected from each clade of duplicates and combined with the unique clones giving a set of 15 different genotypes. None of the landraces or improved cultivars used in the reference panel were found in the genotyped collection. Discriminant Analysis of Principal Components revealed four clusters; Lamu germplasm, breeding lines, inland landraces and coastal landraces with low levels of differentiation between them. According to Wright's F Statistics landraces from Lamu were most closely related to other coastal landraces ($F_{st}=0.0300$), with greater differentiation from breeding lines ($F_{st}=0.0578$) and inland landraces ($F_{st}=0.0632$). Expected heterozygosity was highest for the breeding lines (0.275) and lowest for Lamu landraces (0.240), indicating a narrow genetic base. The results provide valuable information on genetic diversity and identify germplasm for conservation and future use.

C. K. Mutoni (✉) · F. M. Nzuve · D. W. Miano
Department of Plant Science and Crop Protection,
University of Nairobi (UoN), P.O. BOX 29053-00625,
Nairobi, Kenya
e-mail: Christine.mutoni@kalro.org

C. K. Mutoni · B. M. Kivuva
Kenya Agricultural and Livestock Research Organization
(KALRO), P.O. BOX 57811-00200, City Square, Nairobi,
Kenya

C. K. Mutoni · T. M. Shah · M. E. Ferguson
International Institute of Tropical Agriculture (IITA), P.O.
Box 30709-00100, Nairobi, Kenya

I. J. Obare
Kenya Plant Health Inspectorate Service (KEPHIS), P.O.
BOX 49592-00100, Nairobi, Kenya

Keywords Breeding · Conservation · Farmer knowledge and preferences · Genetic resources

Introduction

Cassava is an important source of carbohydrates for over 800 million people globally and is ranked the fourth staple food after rice, maize and wheat. It is also a subsistence crop for 500 million people in Africa (FAO 2018). In Kenya, cassava is the second

most important root crop after potatoes mainly in coastal, eastern and western regions (MOALF 2015). Cassava is predominantly grown by smallholder farmers in marginalized agro-ecologies due to its inherent resilience to abiotic stresses, the diverse use of its starchy roots and underground storability as a food reserve allowing piecemeal harvesting in times of need (Iragaba et al. 2020). Small-scale farmers often grow traditional cultivars (landraces) that have been selected over time for their preferred taste, cooking qualities and agronomic characteristics. Landraces are typically more genetically diverse due to their large numbers, than the few modern cultivars. Farmers tend to grow a mixture of cassava landraces in a single field (Ferguson et al. 2019) to reduce risks of food insecurity and increase resilience. Farmers exchange cassava cuttings with friends and neighbours, from one village to another, thereby improving the chance of increasing diversity through hybridization between different clones, but also often changing the name of the cultivar. It is thus not reliable to identify cassava cultivars based on their given name (Ferguson et al. 2019).

Landraces are often lower yielding and more susceptible to diseases such as cassava brown streak disease (CBSD) and cassava mosaic disease (CMD) than the improved cultivars which have been specifically bred for yield and disease resistance (Munganyinka et al. 2018). It is estimated that yield losses worth more than US\$ 1 billion occur in East and Central Africa mainly attributable to CBSD and CMD (Nyirakanani et al. 2021). Since the first cassava breeding program in East Africa which started at Amani, Tanganyika (now Tanzania) in 1935, breeders have been trying to increase resistance to CBSD and CMD by screening landraces and making crosses with cultivars and even incorporating wild species (Jennings 1957). Several dual CBSD and CMD tolerant or resistant cultivars are now available (Tumwegamire et al. 2018). Replacing a large number of diverse landraces with a few improved cultivars threatens the genetic diversity upon which plant breeding depends (Gwali et al. 2011), and the associated farmer's knowledge (Ferguson et al. 2019). Understanding genetic variability present in landraces provides information on allelic richness and population structure of germplasm useful for breeders.

Genetic diversity has been previously studied using agro-morphological characters, but more

commonly now, due to their independence from environmental influence and large numbers, molecular markers are used. Single nucleotide polymorphic (SNP) markers have been used to identify duplicates in cassava germplasm repositories (Albuquerque et al. 2019; Ferguson et al. 2019); genetic diversity studies of cassava landraces from farmers' fields (Ferguson et al. 2021), genetic diversity of both local and introduced elite cassava germplasm (Pierre et al. 2022) and to track the adoption of improved cultivars (Rabbi et al. 2015).

Phenotypic and molecular characterization of landraces can help to identify duplicates prior to *in vitro* conservation, thus increasing efficiencies and reducing conservation costs (Albuquerque et al. 2019). In Kenya, although genetic diversity studies of cassava have been done in a few counties such as Kakamega, Kiboko and Katumani (Ndung'u et al. 2014) using Simple Sequence Repeat (SSR) markers, no genetic diversity study has been done in Lamu County. The choice of Lamu County in this study was due to the high reliance of the population on cassava as a staple food. Neither cassava genetic resources nor the associated farmer knowledge is represented in a major international genebank for sustainable conservation and future use in breeding programs.

In this study, landraces from farmers' fields in Lamu County were collected and associated farmer knowledge was documented. Molecular characterization using SNP markers was done on the collection to identify; (i) putative duplicates and the most representative sample from a set of duplicates, thereby identifying a set of unique cultivars for conservation and use and (ii) genetic diversity in relation to earlier landrace collections from East Africa and selected breeding lines for identification purposes. This is the first step in understanding the cassava landscape in Lamu County to design appropriate interventions to increase cassava productivity for food security and income generation, focusing on sustainable development goal (SDG) 2 on ending hunger and malnutrition by 2030. The study also responds to Agenda 21 of Earth Summit which calls for strengthening biodiversity conservation for sustainable economic development and environmental conservation.

Materials and methods

Description of the study site

The study was conducted in Lamu County, Kenya in 2020. Lamu County lies between 1° 40' and 2° 30' South and 40° 15' and 40° 38' East. It is generally flat between 0 and 50 m above sea level and covers a total land surface area of 6273 km² with a low population density of 23 persons per km². It consists of Lamu Archipelago and the mainland. It borders Garissa County to the north, Tana River County to the northwest, the Indian Ocean to the southeast and South and the Republic of Somalia to the northeast. The climate in Lamu County transitions between hot humid tropical monsoon and hot, arid steppe. The county receives a bimodal rainfall pattern with long rains (mid-April—end of June) and short rains (November–December). The annual mean temperature is 27 °C with an average humidity of 76% and annual rainfall of between 850 and 1100 mm (GoK 2013, 2019). Crop and livestock farming and fisheries are the main economic activities in the area. Cassava, maize, green grams and sorghum are the main food crops grown while sesame, cashewnuts, coconut, cotton, bixa (a shrub grown mainly for natural food colorants) and mangoes are the major cash crops. The presence of Manda airstrip and a tarmacked road connecting Lamu County to the other parts of the country presents an economic opportunity for transportation of farm produce to market.

Sample collection

A multistage sampling method was employed to select cassava farmers and fields. In the first stage Lamu West Sub-County was selected out of the two sub-counties, due to the presence of small and medium-scale cassava growers. In the second stage, five out of seven main cassava growing administrative wards were selected, based on information provided by agricultural extension officers. Third stage, farmers within the wards known to grow a large number of cassava landraces in their fields were randomly selected. The selected farmers were identified through information from key informants or referrals from previous farmers. The farmers interviewed were the head of households either men or women. The identification and collection of the cassava landraces from

farmer's fields were done jointly by the head of the household and researchers. A landrace was recorded based on the name given to them by the farmer as well as a collector identification number which was unique to each sample. The distance between the two visited farms was approximately 15 km. At each farm, three different well-structured questionnaires were administered (Cox et al. 2014); (i) key informant information to capture the farmer, their farm, the identity of the cultivars they grow, farmer and consumer preferences, use and source of their planting materials; (ii) an individual landraces form to capture the history of the landraces, farmer perceptions on the adaptability to both biotic and abiotic stresses, and agronomic practices, in addition, desirable traits associated with the landrace were documented; (iii) germplasm questionnaire to document phenotypic and morphological description of the cultivars collected, and farmers' knowledge of cassava viral diseases as documented by Fakuda et al. (2010). Cuttings of each landrace were taken from a plant selected at random, clearly labelled and tied in bundles while apical leaves from each cultivar were labelled and preserved in silica gel for subsequent genotyping. The cassava cuttings were transported to Kenya Plant Health Inspectorate Service (KEPHIS), Muguga where they were planted in pots in a screen house and from which apical meristems were sampled for *in vitro* conservation. A unique code was given to each questionnaire corresponding to a single key informant interview. This unique code and the landrace collector numbers were used to track information of each germplasm sample and associated farmer knowledge in the database.

Morphological characterization of cassava landraces

Data relating to qualitative morphological traits were collected under farmer's field conditions according to Fakuda et al. (2010). The descriptors included the colour of apical leaves, presence or absence of pubescence, shape of central leaflets, petiole colour, number of lobes, flowering, stem cortex colour, stem epidermis, growth habit, branching habit and shape of plant. Plant growth habit was recorded as either zig-zag or straight. The underground traits included the external colour of storage root, the colour of the root pulp and the root cortex which was determined by

peeling off the outer cover of the root and cutting it transversely.

Molecular characterization of cassava local landraces

Four leaf disks of approximately 5 mm diameter were collected from young fresh cassava leaf samples in farmers' fields into a 96-well plate and placed on silica gel for preservation, before being lyophilized in the lab. The cassava leaf samples were sent to Inter-tek (South Australia) for DNA extraction, and genotyping at Diversity Array Technologies (DArT) using low density DaRTSeq™ (variant DaRTSeqLDPas1). SNPs markers were scored '0' for reference allele homozygote, '1' for SNP allele homozygote and '2' for heterozygote.

To identify the collected landraces, a reference set of eight coastal and 45 inland landraces from Tanzania were included (Supplementary file S1) (Ferguson et al. 2021), and 48 breeding lines from East and West Africa. Four known biological (different plants of the same cultivar) duplicate were included to determine a suitable cut-off for defining duplicates and unique cultivars from the collected accessions. The biological replicates include (TME14-Ug, TME14, TME-14 and Abbey-Ife). Adding biological duplicates aids in determining whether two samples should be considered different cultivars or the same cultivar.

Data analysis

Morphological data from the collected landraces was coded and entered into Microsoft (MS) Excel software. It was then imported into Statistical Package for Social Sciences Statistics (SPSS) software (IBM SPSS Statistics for Windows version 25.0, IBM Corp., Chicago, IL, USA) for analysis. The analyses involved frequency, descriptive and multiple responses on household production of local landraces and conservation practices. The multiple response data generated for utilization, conservation practices and production of cassava landrace were grouped together using the multiple response command of SPSS. The results were presented as figures and tables on frequencies, percentages, means, modes and medians. Cassava landraces were also listed based on the names given by the farmers.

Allele calls from genotyped data were first converted to allele dosage format before being imported

into DartR (Gruber et al. 2018) and converted to a compact genlight format for R. All R analyses utilized version 4.2.2 (R Core Team 2018). The SNP data was filtered to select loci or individuals with high quality metrics by deleting all SNPs and genotypes with more than 5% missing data and all SNPs with minor allele frequencies of less than 0.05 including monomorphic loci (Gruber et al. 2018). A subset of 159 samples was obtained consisting of 58 farmer cultivars, a reference set of 48 breeding lines, eight coastal and 45 inland landraces (Supplementary file_S1), and four biological duplicates that were some of the controls used by Ferguson et al. (2021).

To determine the number of unique cultivars and identify duplicates in the collected samples, a distance matrix was generated using PLINK function implemented in R version (4.2.2. throughout) (R Core Team 2018), based on the similarity of four known biological duplicates (Abbey Ife, TME-14, TME14, and TME-Ug). Accessions with a genetic distance below the defined threshold were designated as putative duplicates. To reduce the redundancy of duplicate samples in the same clonal cluster, the most representative genotype within a group of putative duplicates was identified to represent the group, based on the cultivar with the highest average similarity to the rest of the members of the group. A Ward's hierarchical cluster dendrogram (Fig. 3) was constructed from identity by state (IBS) matrix using R software.

Assessment of the relationship between populations was done using discriminant analysis of principal components (DAPC) in R after eliminating duplicates. DAPC combines discriminant analysis (DA) and principal components and is suitable for populations that violate Hardy Weinberg Equation (HWE) and linkage equilibrium assumptions like the out-crossing, clonally propagated crop such as cassava. DAPC assigns individuals into groups according to membership. Seventy principal components were retained within two eigenvectors. Genetic diversity parameters and Wright's F statistics were calculated using HierFstat (ver. 0.5–11) (Goudet and Jombart (2015), by calculating observed heterozygosity (H_o) using mean "hobs" and expected heterozygosity (H_e) (Adamack and Gruber 2014).

Nei's genetic distance was calculated among sub-populations and relationships represented in weighted neighbour-joining method in R. Expected heterozygosity (H_e) is the theoretical rate of heterozygosity

assuming the population meets the Hardy–Weinberg equilibrium (HWE). HWE in every locus was calculated from the allelic frequencies according to the formula $He = 1 - \sum f_{ij}^2$ where f_{ij} is the frequency of the j th allele of i th locus. Observed heterozygosity (H_o) is the number of heterozygous loci in relation to the total number of loci in the sample. It is calculated by the genotypic frequencies from the sample at a given locus K , using the formula $HoK \sum_{ij}^{ak} = 1 P_{ij}$ where P_{ij} is the estimate of the frequency of the genotype ij at locus K and ak is the number of alleles at locus K (Yeo et al., 2021). He and H_o ranges from 0 to 1 with 0 for no heterozygosity and 1 when there are many alleles at equal frequencies.

Virus cleaning

Cassava cuttings of the unique landraces were collected from the screen-house at KEPHIS, Muguga, for virus screening and subsequent cleaning using methods described by Ferguson et al. (2020) in preparation for in vitro conservation at a designated facility.

Results

A total of 40 farmers growing cassava were interviewed in Lamu county, and cuttings and leaf samples taken from 58 cassava cultivars (Fig. 1).

Production dynamics and duration of cassava cultivation

Men made up 68.4% of the interviewed key informants while women were 31.6%. Their ages ranged from 20 to 83 years and 89.5% were the principal caretaker/ decision makers. These people made key decisions concerning cassava cultivation regarding the cultivar type and numbers of each to be grown. The majority (84.2%) of the key informants interviewed relied on farming as their main occupation and as a source of income. Others were engaged in formal employment or business. Forty-two percent (42.1%) of these had a primary level of education with 10.5% and 15.8% having secondary and tertiary levels of education respectively, whereas 31.6% had no formal education. In terms of the estimated number of years the respondent had been growing cassava, 21.1% had less than 5 years, 47.4% had between 5 and 19 years, 26.3% had between 20 and 49 years

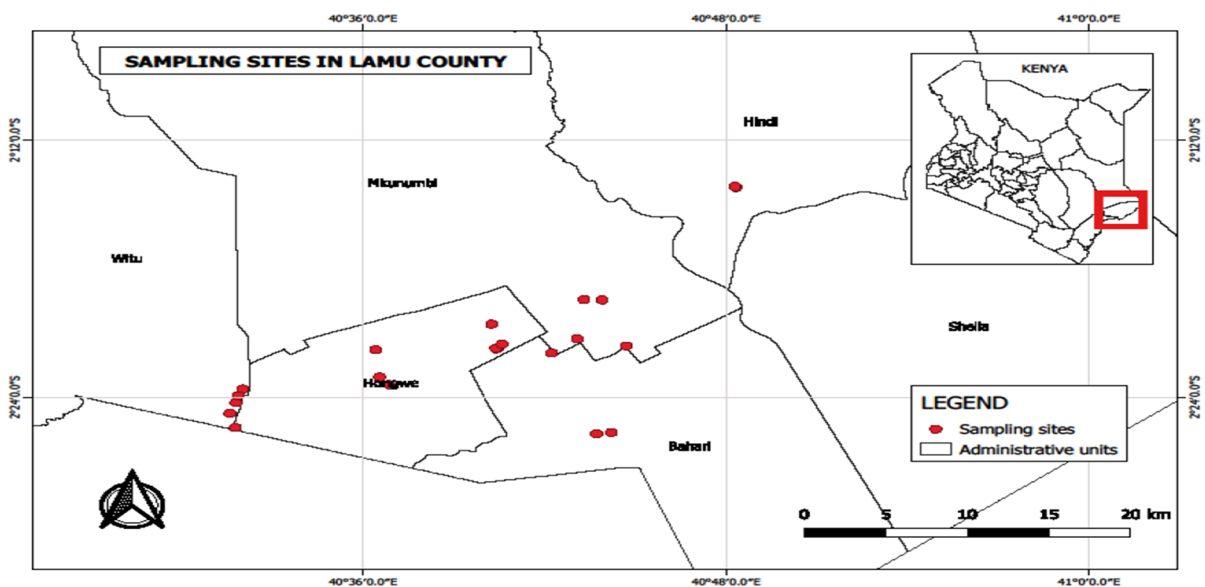


Fig. 1 Location of the sampling sites in Lamu county. Some farms were close together with more than one sample being collected from one farm leading to overlapping points

while 5.3% had 50 years or more. The majority of the farmers interviewed (57.9%) grew cassava on farms ranging from between 5 and 9.9 acres followed by 42.1% who grew cassava on more than 10 acres.

Utilization of cassava landraces

The majority of farmers (77%) grew cassava landraces for home consumption, especially the sweet cultivars. The roots were either boiled or deep fried for home use. Sweet cultivars are even chewed raw. Excess fresh roots of 15.8% are sold to the local market. Bitter cultivars (7.2%) are processed into flour for cooking stiff porridge (ugali).

The landraces

Of the 58 landraces collected, the majority had local names like Kibandameno which means ‘strong on teeth’, while others were named according to their morphological traits like Kibandameno straight, forked. The cultivars were also named based on their characteristics such as sweet taste, origin, or after the name of the person who introduced them in the area like Chengo 2, and 3, and Luvuno. In terms of the time frame when the cassava cultivars were introduced in their communities, 25.5% of interviewed farmers, estimated between 1980s and 1990s, 7.3% were introduced in 2000s, 14.4% in 2010, 10.9% in 2020s while 41.8% of the farmers were not aware of the time of introduction.

In the previous five years, 71.4% of the collected landraces had gained popularity and their growth trend increased in that area. These landraces include Kibandameno forked, kibandameno straight. The reasons cited for increased cultivation of certain landraces as per multiple responses analysis were high yields (29.0%), storability in the ground (29.1%), drought tolerance (36.0%), diseases and pest resistance (23.0%), cooking quality (34.0%), taste (34.0%) demand and marketability (29.0%). Fourteen percent of the landraces had decreased demand, while for a few, the trend was stagnant or unknown. The respondents cited susceptibility to diseases and pests and low yield as the main reason for decreased cultivation of cultivar like guzo. Besides, farmers also cited extinction of some of their preferred cultivars such as kahutele and katsunga probably due to drought.

Conservation practices undertaken by smallholders and sources of planting material

Based on collected information, farmers in Lamu County conserve and maintain their preferred landraces on the farm by growing them in multiple plots (81%), giving others to plant as backup (73.8%), use of clean planting materials (2.4%) and the remaining 2.4% of the farmers knew where to find the landraces. Seventy-three percent of farmers reported that they use their own planting materials but in case of loss or they want to try a new landrace, 40.5% obtained it as a gift from someone in the community, or outside the community (26.2%), or by purchasing within the community (7.1%), or from an outsider (2.4%) and the remaining got them from volunteer seeds in the soil (2.4%).

Morphological characterization of cassava landraces

Forty-five landraces were characterized morphologically at the time of collection with full results in Fig. 2.

Leaf traits

The majority of cassava landraces (77.8%) had purple-green apical leaves, whereas 15.6% had a purple colour while 6.7% had a light green colour. No pubescence on the apical leaflets was observed. The most common shape of the central leaf was elliptic lanceolate (95.6%) followed by obovate-lanceolate (4.4%). Most petioles were purple in colour (54%), and the majority of landraces (35%) had seven leaf lobes.

Stem traits

The stem exterior of 77.8% of landraces was silver while 8.9% was light-green, 4.4% orange, 4.4% red-green, 2.2% dark green and 2.2% brown green. Eighty-nine percent of the landraces had a straight stem growth habit while 11% had a zigzag stem growth habit. Most of the landraces (73.3%) were compact in shape while 4.4% were cylindrical, 11.1% had umbrella shape and 11.1% were open.

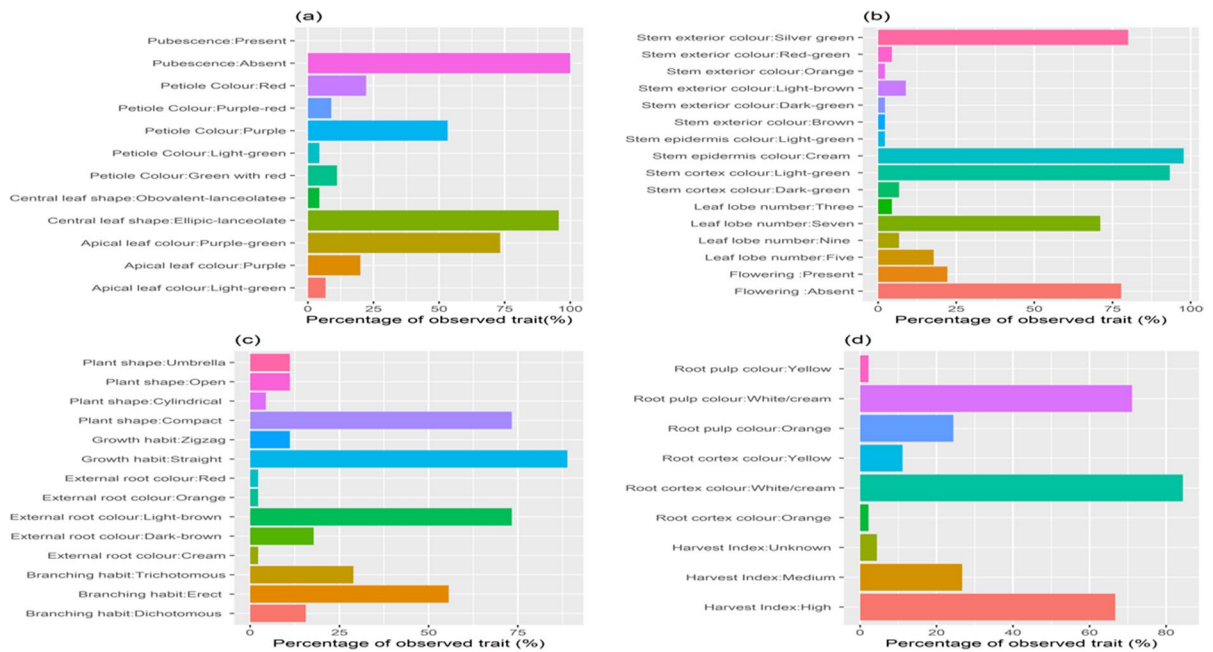


Fig. 2 Morphological descriptors for 45 genotypes evaluated according to Fukuda et al. (2010)

Root traits

External colour of storage root was light brown in 73.3% of cases, 17.8% had dark brown, 2.2% red, 2.2% orange, while 4.5% had not developed storage roots. Root pulp was cream in 75.6% of landraces, 22.2% had orange colour and 2.2% had yellow root pulp colour. From the survey conducted, thirty-one (68.9%) cassava samples were high yielding as perceived by farmers, while 26.7% cultivars had average yields, two (4.4%) of the landraces were new and young at the time of this study, farmers were yet to determine their yields.

Number of unique cultivars grown in Lamu County and their duplicates

A total of 58 leaf samples were collected and successfully genotyped using DaRTSeq™. Based on cultivar names assigned by farmers, an average of three cultivars were collected per farm, although some had as few as one and as many as four.

By using a threshold of 0.05, in the 58 accessions, six unique cultivars (CKM 05, CKM 18, CKM 32, CKM 36, CKM 50 and CKM 58) were

identified, three from Witu, two from Hindi and one from Hongwe ward (Fig. 3). From the remaining 52 farmer cultivars collected, nine clades of duplicates were defined (Supplementary file 2). One cultivar from each clonal group (clade) with the highest average similarity to the rest of the individuals within the group was selected to represent the group. From this, a total set of 15 unique cultivars were defined, virus indexed, cleaned and conserved in-vitro at the Kenya Plant Health Inspectorate Service (KEPHIS), Muguga. According to the farmer's information, all duplicates had different ethnic names making it difficult to identify cultivars based on the given name.

To explore the genetic relationships between the defined unique set of cultivars (15), reference breeding lines (48), inland (45) and coastal (eight) reference landraces, a dendrogram, based on individuals, and colour coded according to category was constructed and illustrated in Fig. 3.

The genetic relationship between these groups of cultivars is displayed on DAPC (Fig. 4).

Interestingly, the Lamu landraces clustered quite distinctly from both the coastal landraces and the inland landraces which were more closely related to each other. The breeding lines were more closely

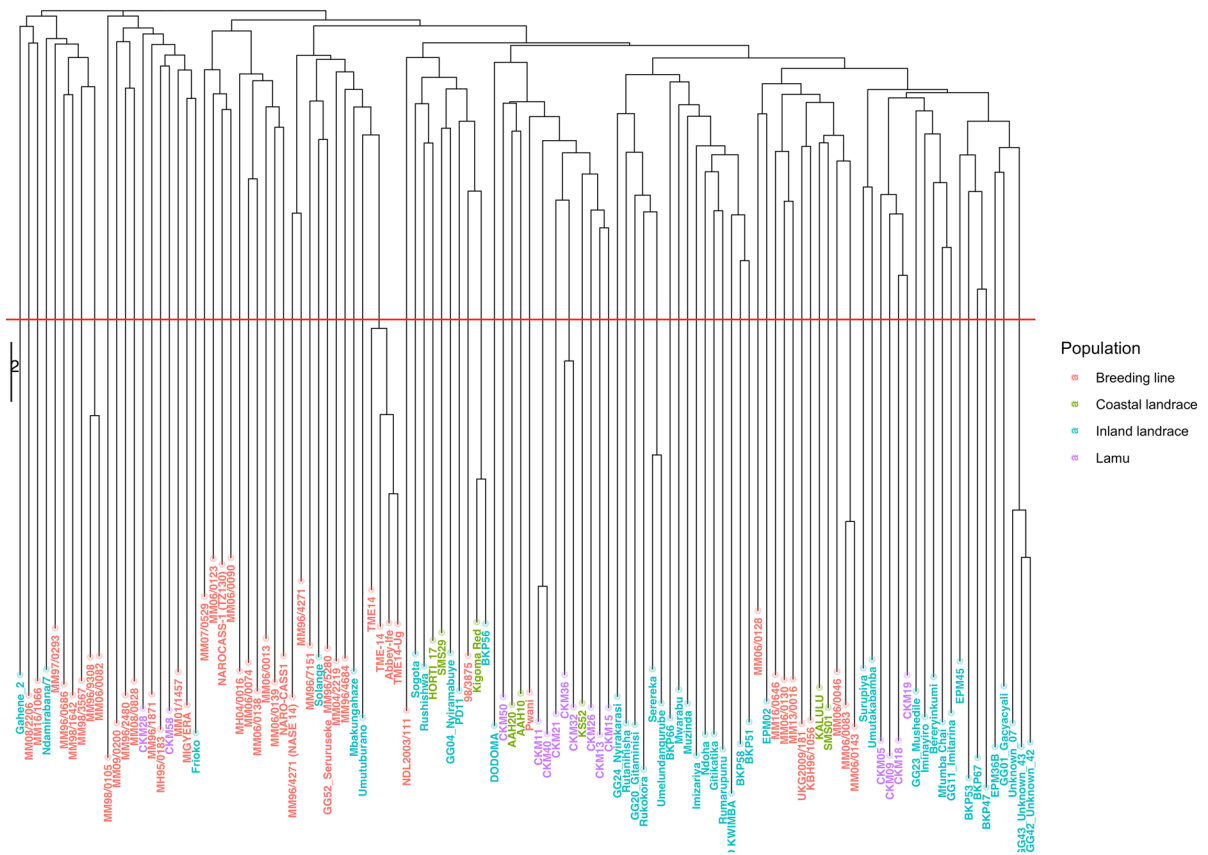


Fig. 3 Dendrogram from Ward's hierarchical clustering of Identity by State (IBS) matrix with 15 unique farmer cultivars, 48 reference breeding lines, 8 coastal, 45 inland reference landraces and four known biological replicates

related to the inland and coastal landraces than the Lamu landraces.

Determination of population statistics

As is typical of outcrossing species, expected heterozygosity (H_s) from the four data sets was lower than observed heterozygosity (H_o) with an overall H_s of 0.263 and H_o of 0.283 and negative F_{IS} value of -0.076 which is a measure of inbreeding. It was observed that Breeding lines (from East and West Africa) had slightly higher mean H_s of 0.275 followed by inland landraces with 0.265 and coastal landraces of 0.255. The Lamu landraces had H_s mean of 0.24 (Table 1).

Nei's genetic distance and pairwise fixation among populations indicated very low differentiation between populations. The most distantly related populations with the greatest differentiation (F_{st})

(Table 2), were the inland landraces and Lamu cultivars ($Nei's=0.065$, $F_{st}=0.063$), followed by breeding lines and Lamu cultivars ($Nei's=0.060$; $F_{st}=0.058$). The most closely related accessions with the least differentiation (F_{st}) were those of the coastal and inland landraces ($Nei's=0.026$, $F_{st}=0.024$), followed by breeding lines and coastal landraces ($Nei's=0.030$; $F_{st}=0.028$). The pairwise population matrix of Nei's genetic distance and F_{st} values between the four populations is given in Table 2 and illustrated in Fig. 5.

Discussion

The results from the questionnaires found that farmers use their traditional knowledge to identify and differentiate cultivars based on leaf traits (colour of apical leaves, pubescence on apical leaves, petiole

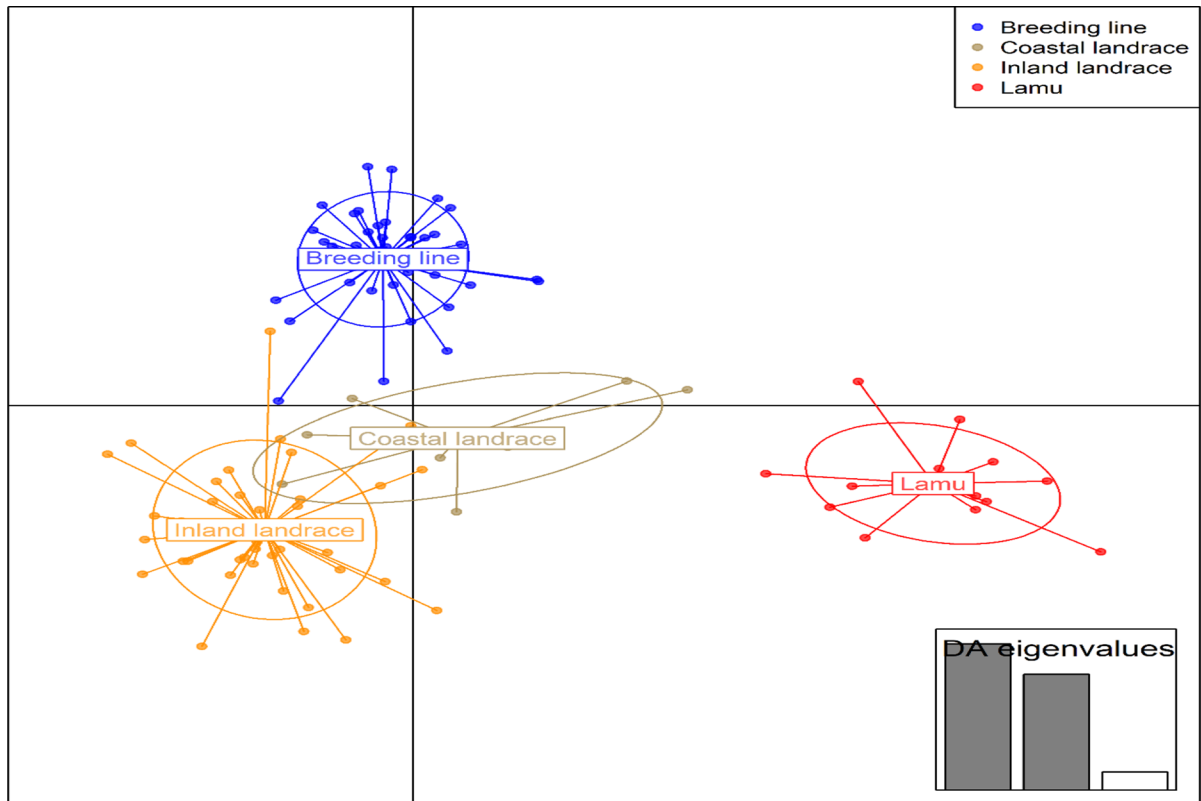


Fig. 4 Plot of DAPC from 58 local landraces collected in Lamu county, reference coastal landraces, reference inland landraces and reference breeding lines

Table 1 Genetic diversity parameters

	Ho	Hs	Ht	Fst	Fis
Overall	0.283	0.263	0.271	0.03	−0.076
Breeding lines	0.294	0.275	0.277	0.008	−0.071
Coastal landraces	0.275	0.255	0.250	0	−0.077
Inland landraces	0.278	0.265	0.270	0.016	−0.049
Lamu	0.262	0.240	0.245	0.021	−0.093

Key:

Ho = Observed heterozygosity

Hs = Expected heterozygosity; Genetic diversity within the population

Ht = overall genetic diversity

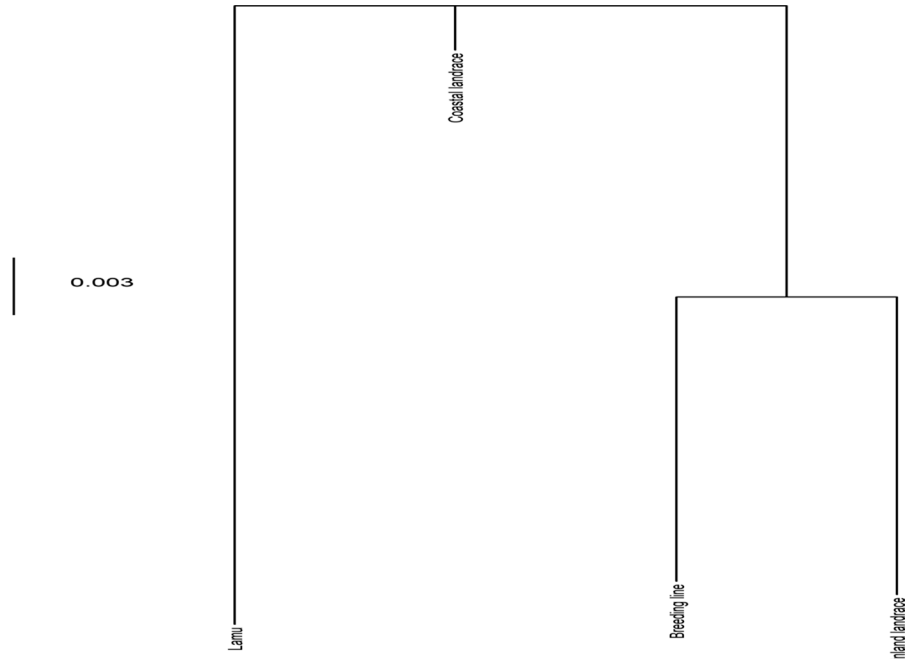
Fst (fixation index) = is a measure of the identity of individuals within sub-populations compared to individuals from other sub-populations within the total population or the total homozygosity due to the Wahlund effect

colour, number of leaf lobes), stem traits (stem cortex colour, epidermis colour, stem exterior colour, stem growth habit) and root descriptors (colour of root cortex, root pulp colour, external colour of storage root) which were found to be varied among the landraces. Similarly, Agre et al. (2017) and Asare et al. (2011) observed that farmers differentiate cassava cultivars based on leaf colour and shape of central leaf. The analysis from the study showed that the purple green was the dominant colour of apical leaf at 77.8% a similar observation made by Pierre et al. (2022) from Burundi. According to Ceballos and de la Cruz (2002), the purple colour is easily found on young shoots which later grow and change to greenish colour, agreeing with Phosaengsri et al. (2019). In cassava, root pulp is the most economically important

Table 2 Pairwise population matrix of Nei's genetic distance (He) and pairwise fixation index (Fst) in parentheses

	Breeding line	Coastal landraces	Inland landraces
Coastal landrace	0.0299 (0.0276)		
Inland landrace	0.0304 (0.0304)	0.0263 (0.0243)	
Lamu	0.0602 (0.0578)	0.0300 (0.0300)	0.0652 (0.0632)

Fig. 5 Genetic relationship between cassava populations based on Nei's genetic distance and the weighted neighbour-joining method of clustering



part. White pulp is the most preferred colour for sale of fresh roots in local markets and in flour processing (Ferguson et al. 2021). Of the 45 genotypes assessed, white/cream pulp colour was most dominant (75.6%) colour meaning they were low in carotenoids. Of the remaining cultivars ten had orange (22.2%) pulp and one had yellow (2.2%) pulp. These cultivars may be useful in breeding for high beta-carotene, a precursor of vitamin A, and should be further investigated. Breeding for enhanced beta-carotene has been a major objective in many crops, as a means of ameliorating vitamin A deficiency, including in cassava (Edoh et al. 2016; Njoku et al. 2011).

Of the farmers interviewed, 68% were men, making decisions as to which cultivar or landrace to grow and the proportion of each. Farmers reported that the selection and conservation of cultivars was based on disease tolerance, yield, cooking time, taste and early maturity. Similarly, Ferguson et al. (2021) and Mtunguja et al. (2014) both reported that the selection of

landraces was done by both men and women in Tanzania. Farmers recycle their own materials (73%) or supplement with other clones from their friends and neighbours (27%). This seed flow or exchange occurs within the village from one village to another or introductions from other communities (Mtunguja et al. 2014). Apart from farmer exchange, farmers also incorporate new cultivars from volunteer seedlings that germinate from seed in the soil seed banks (Mtunguja et al. 2014). The majority (57.9%) of respondents grew cassava on farms between 5 and 9.9 acres while 42.1% grew cassava on more than 10 acres. These are relatively substantial acreages for Kenya and demonstrate the significance of cassava in Lamu County. Farmers in Lamu County employ various practices to retain cultivars on the farm, the most common being planting in multiple plots (50.7%) and sharing with others (46.3%). In some cases, cuttings are kept viable under shade until the following season. These findings agree with other reports where

cassava is grown in East Africa (Nakabonge et al. 2018; Mtunguja et al. 2014). Despite these efforts farmers reported a loss of some older landraces due to drought and a surge in disease pressure.

SNP fingerprinting, together with the use of biological replicates to define a cut-off, was useful in identifying duplicates. Of the 58 samples collected from farmers' fields, there were six genetically distinct unique cultivars, collected only once, and nine clades of duplicates, giving a total of 15 distinct genotypes. This shows that many of the same landraces were collected under different names. This could be due to the plasticity of the morphological descriptors or farmers assigning new names when genotypes are introduced to the community as observed by Elias et al. (2001). Cultivar names have consistently been found to be unreliable in identifying clones (Iragaba et al. 2020 and Ferguson et al. 2021). Identification of duplicates is a prerequisite for efficient conservation and utilisation of cassava germplasm.

None of the genotypes collected here matched any of the improved cultivars or landraces used in the reference panel. It is possible that some of these could still be improved cultivars but samples were not included in the reference panel. The study shows the possibility of novel clones in farmers' fields and demonstrates the need for conservation and evaluation for use.

As has been found in several studies, genetic differentiation among groups of cassava cultivars in Eastern, Southern and Central (ESC) Africa is quite low i.e. $F_{ST} < 0.05$ (Wright 1978). The mean F_{ST} value of 0.03 observed between the farmer's cultivars, breeding line, coastal and inland landraces indicate a low genetic differentiation between subpopulation, despite being higher than mean F_{ST} 0.003 observed by Iragaba et al. (2020) among farmer preferred cultivars in Uganda. This may be influenced by the high genetic variability within individuals as a result of outcrossing which is indicated by negative FIS mean values in all populations.

DAPC (Fig. 4) depicted a similar situation as Nei's genetic distance (Fig. 5) with coastal landraces and inland landraces most closely related, followed by breeding lines, with the cultivars from Lamu county being most distantly related. This could be due to the exchange of planting materials in a continuum from inland landraces to coastal landraces, and subsequent outcrossing. In addition,

landraces have been used in breeding programs. The relatively distant relationship of landraces from Lamu County which lies on the coast, to largely Tanzanian coastal landraces is surprising. This can be explained by divergent evolution caused by isolation, or the 'island effect' which has been observed in cassava from the Comoros, Zanzibar and Pemba (Ferguson et al. in press). There is substantial distance and lack of movement between Lamu County and other cassava growing areas to the South. This makes their conservation more critical. As breeding lines are not closely related to Lamu landraces, the adaptation of existing breeding lines to agro-environments in Lamu county needs to be assessed, and their cooking and eating qualities need to be evaluated by consumers. There may be a need to introgress properties from local landraces into breeding lines to incorporate culinary properties and environmental adaptation.

Conclusion

Lamu farmers still grow local landraces despite the dissemination of high yielding and disease tolerant cultivars. This was confirmed by the presence of some unique cassava landraces in the region. The study also revealed levels of differentiation between Lamu clones and those from other regions in the country and neighbouring countries which indicates limited germplasm exchange between farmers in Lamu and other regions. Additionally, cassava germplasm from Lamu County has a narrow genetic base. Introductions are needed of cultivars of similar agro-ecologies to increase diversity to buffer against pests and diseases, and the uncertainties resulting from climate change.

Acknowledgements This work was funded by the World Bank through the Kenya Climate Smart Agriculture Project (KCSAP). The authors thank the Lamu West Sub-county extension staff and Lamu West Sub-county cassava farmers for providing planting materials and leaves.

Author contributions CKM conceived and implemented the study, analysed the data and wrote the first draft of the manuscript; FMN, DWM and BMK provided guidance during planning and implementation of the project as well as supervision; IJO provided support and guidance during planning;

implementation and data analysis; TMS assisted with data analysis and feedback on the manuscript and MEF provided scientific guidance, data analysis, edited the manuscript and supervision.

Funding Mutoni C.K. has received research funds from World Bank through the Kenya Climate Smart Agriculture Project (KCSAP). The authors declare that no funds, grants or other support were received during preparation of this manuscript.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

- Adamack AT, Gruber B (2014) PopGenReport: simplifying basic population genetic analyses in R. *Methods Ecol Evol* 5:384–387
- Agre AP, Bhattacharjee R, Rabbi IY, Alaba OY, Unachukwu NN, Ayanan MAT et al (2017) Classification of elite cassava cultivars (*Manihot esculenta* Crantz) cultivated in Benin Republic using farmers' knowledge, morphological traits and simple sequence repeat (SSR) markers. *Genetic Resour Crop Evol* 64:307–320
- Albuquerque HYG, Oliveira EJD, Brito AC, Andrade LRBD, Carmo CDD, Morgante CV et al (2019) Identification of duplicates in cassava germplasm banks based on single-nucleotide polymorphisms (SNPs). *Sci Agric* 2019(76):328–336
- Asare AP, Galyon IKA, Sarfo KK, Tetteh JP (2011) Morphological and molecular based diversity studies of some cassava (*Manihot esculenta* Crantz) germplasm in Ghana. *Afr J Biotechnol* 10:13900–13908
- Ceballos H, de la Cruz GA (2002) Taxonomy and morphology of cassava. OSPINA, Bernardo and CEBALLOS, Hernán. *Cassava in the Third Millennium*, pp. 16–31
- Cox TP, Nakabonge G, Ferguson M (2014) Manual for collection of cassava germplasm and associated farmer knowledge in Eastern and Southern Africa. Nigeria, Ibadan, IITA, 2014
- Edoh NL, Adiele J, Ndukwe I, Ogbokiri H, Njoku DN, Egesi CN (2016) Evaluation of high beta carotene cassava genotypes at advanced trial in Nigeria. In: *The open conference proceedings Journal 2016 Dec 13* (vol 7, No 1)
- Elias M, McKey D, Panaud O, Anstett MC, Robert T (2001) Traditional management of cassava morphological and genetic diversity by the Makushi Amerindians (Guyana, South America): perspectives for on-farm conservation of crop genetic resources. *Euphytica* 120:143–157
- FAO (2018). <http://www.fao.org/faostat/en/#data/QC>. Accessed 2nd October 2022
- Ferguson ME, Shah T, Kulakow P, Ceballos H (2019) A global overview of cassava genetic diversity. *PLoS ONE* 14(11):e0224763
- Ferguson ME, Tumwegamire S, Chidzanga C, Shah T, Mtunda K, Kulembeka H et al (2021) Collection, genotyping and virus elimination of cassava landraces from Tanzania and documentation of farmer knowledge. *PLoS ONE* 16(8):e0255326. <https://doi.org/10.1371/journal.pone.0255326>
- Ferguson M, Choti A, Munguti F (2020) Virus elimination and testing in cassava: a practical manual. Ibadan, Nigeria: IITA, pp 24
- Fukuda WMG, Guevara CL, Kawuki R, Ferguson ME (2010) Selected morphological and agronomic descriptors for the characterization of cassava. Ibadan: Nigeria IITA, 19
- GoK (Government of Kenya) (2019) 2019 Kenya Population and Housing Census: Volume I: Population by County and Sub-County; ISBN: 978-9966-102-09-6; Government Printers, Nairobi
- GoK (2013) County Development Profile (CDP) for Lamu, Ministry of Devolution and Planning, Government Printers, Nairobi, Kenya
- Goudet J, Jombart T (2015) Hierfstat: estimation and tests of hierarchical F-statistics
- Gruber B, Unmack PJ, Berry OF, Georges A (2018) Dartr: an R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol Ecol Resour* 18:691–699
- Gwali S, Okullo JBL, Eilu G, Nakabonge G, Nyeko P, Vuzi P (2011) Traditional management and conservation of Shea trees (*Vitellaria paradoxa* subspecies *nilotica*) in Uganda. *Environ Dev Sustain* 14(3):347–363
- Iragaba P, Tufan HA, Kawuki RS et al (2020) Genomic characterization of Ugandan smallholder farmer-preferred cassava cultivars. *Crop Sci* 60:1450–1461
- Jennings D (1957) Further studies in breeding cassava for virus resistance. *East Afr Agric J* 22:213–219
- MOALF (2015) Agriculture Sector Development Support Programme (ASDSP), Annual work plan and budget
- Mtunguja MK, Laswai HS, Muzanila YC, Ndunguru J (2014) Farmers knowledge on selection and conservation of cassava (*Manihot esculenta*) genetic resources in Tanzania. *J. Biol Agric Healthcare* 4:120–129
- Munganyinka E, Ateka EM, Kihurani AW, Kanyange MC, Tairo F, Sseruwagi P et al (2018) Cassava brown streak disease in Rwanda, the associated viruses and disease phenotypes. *Plant Pathol* 67:377–387
- Nakabonge G, Samukoya C, Baguma Y (2018) Local cultivars of cassava: conservation, cultivation and use in Uganda. *Environ Dev Sustain* 20:2427–2445
- Ndung'u JN, Wachira FN, Kinyua MG, Lelgut DK, Njau P, Okwaro H, Obiero H (2014) Genetic diversity study of Kenyan cassava germplasm using simple sequence repeats. *Afr J Biot* 13(8):926–935
- Njoku DN, Vernon G, Egesi CN, Asante I, Offei SK, Okogbenin E, Kulakow P, Eke-Okoro ON, Ceballos H (2011) Breeding for enhanced β -carotene content in cassava: constraints and accomplishments. *J Crop Improv* 25(5):560–571
- Nyirakanani C, Pierre BJ, Yves K et al (2021) Farmer and field survey in cassava-growing districts of Rwanda reveals key factors associated with cassava brown streak disease incidence and cassava productivity. *Front Sustain Food Syst* 5:1–14

- Phosaengsri W, Banterng P, Vorasoot N, Jogloy S, Theerakulpisut P (2019) Leaf performances of cassava genotypes in different seasons and its relationship with biomass. *Turkish J Field Crops* 24:54–64
- Pierre N, Wamalwa LN, Muiru WM, Simon B, Kanju E, Ferguson ME et al (2022) Genetic diversity of local and introduced cassava germplasm in Burundi using DArTseq molecular analyses. *PLoS ONE* 17(1):e0256002. <https://doi.org/10.1371/journal.pone.0256002>
- R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2018. <https://www.R-project.org>
- Rabbi IY, Kulakow PA, Manu-Aduening JA, Dankyi AA, Asibuo JY, Parkes EY, Abdoulaye T, Girma G, Gedil MA, Ramu P, Reyes B (2015) Tracking crop varieties using genotyping-by-sequencing markers: a case study using cassava (*Manihot esculenta* Crantz). *BMC Genet* 16(1):1–11
- Tumwegamire S, Kanju E, Legg J et al (2018) Exchanging and managing in-vitro elite germplasm to combat cassava brown streak disease (CBSD) and cassava mosaic disease (CMD) in Eastern and Southern Africa. *Food Secur* 10(2):351–368
- Wright S (1978) Evolution and the genetics of populations. Variability within and among natural populations, vol 4. University of Chicago Press, Chicago
- Yeo EF, William JL, Amoakon WJS, Pita JS, Mutuku M, N'zué B, Kouassi MK, Yao N, Otron DH, Shah T, Vanié-Leabo LPL, Kpahé KMH, Sié R, Koné FSD, N'guetta S-PA, Kouassi NK, Ferguson ME (2021) A panel of single nucleotide polymorphism (SNP) markers identifies potential duplicates2 in cassava (*Manihot esculenta* Crantz) cultivars from Côte d'Ivoire. *bioRxiv*. <https://doi.org/10.1101/2021.05.24.445412>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.