academicJournals

Vol. 12(28), pp. 4453-4464, 10 July, 2013
DOI: 10.5897/AJB2013.12348
ISSN 1684-5315 ©2013 Academic Journals
http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Simple sequence repeats (SSR) diversity of cassava in South, East and Central Africa in relation to resistance to cassava brown streak disease

A. Pariyo[1,2], P. Tukamuhabwa[1], Y. Baguma[2], R. S. Kawuki[2], T. Alicai[2], P. Gibson[1,3], E. Kanju[4], B.W. Wanjala[5], J. Harvey[5], I. Nzuki[5], I. Y. Rabbi[6] and M. Ferguson[6]**

[1]Makerere University, Department of Agricultural Production, P.O. Box 7062 Kampala, Uganda.
[2]National Crops Resources Research Institute (NaCRRI), P.O. Box 7084 Kampala, Uganda.
[3]Southern Illinois University, Department of Plants, Soils, Agricultural Systems, Carbondale, IL, USA.
[4]International Institute of Tropical Agriculture (IITA), P.O Box 34441, Dar Es Salaam, Tanzania.
[5]Biosciences eastern and central Africa (BecA) Hub ILRI, P.O. Box 30709, Nairobi 00100, Kenya.
[6]International Institute of Tropical Agriculture (IITA), c/o International Livestock Research Institute (ILRI), P.O. Box 30709, Nairobi 00100, Kenya.

Accepted 5 July, 2013

This study was conducted to: (1) determine the amount and structure of the genetic diversity of cassava in southern, eastern and central (SEC) Africa using simple sequence repeat (SSR) markers, (2) determine the frequency and distribution of alleles putatively associated with resistance to cassava brown streak disease (CBSD) derived from the variety 'Namikonga', and (3) examine the genetic relationships among some CBSD resistant genotypes. The major findings from this study were: (1) little genetic differentiation was confirmed among countries (7%) with the majority of variation distributed among individuals (93%), (2) the frequency of alleles putatively associated with resistance to CBSD was found to be generally low (0.2 to 0.5) but widely dispersed in the cassava germplasm of the SEC Africa region implying that if validated the markers will be useful for marker-assisted breeding on a broad scale, (3) several distinct putative sources of resistance to CBSD seem to be present in the cassava germplasm of SEC Africa offering the potential to pyramid genes for more effective and durable resistance and (4) it may be more informative to assess cassava diversity in the SEC Africa region based on pedigree (if known) or genetic relatedness, rather than geographical origin.

Key words: Cassava brown streak disease (CBSD), simple sequence repeat (SSR), genetic diversity, pedigree information.

INTRODUCTION

Cassava, *Manihot esculenta* Crantz, belongs to the family *Euphorbiaceae* with 98 species. All *Manihot* species examined had 2n = 36 chromosomes (Nassar, 2002). The crop which is native to South America (Allem et al., 2001) was brought to Africa during the last half of the 16th century (Purseglove, 1974). Cassava was later introduced

*Corresponding author. E-mail: m.ferguson@cgiar.org.

Abbreviations: SEC, South, East and Central; **CBSD,** cassava brown streak disease; **CMD,** cassava mosaic disease; **SSR,** simple sequence repeat; **QTLs,** quantitative trait loci; **MAS,** marker-assisted selection; **SNP,** single nucleotide polymorphism.

to Réunion and Madagascar in 1736 before it was introduced to the interior of East Africa around 1850 (Purseglove, 1974; Lebot, 2009). It now forms an important source of food, feed and industrial raw material worldwide in tropical and sub-tropical regions for over 800 million people globally and over 200 million people in sub-Saharan Africa (IITA, 2007). Africa is the largest producer of the crop in the world with the South, East and Central (SEC) Africa region constituting a major production and utilization area.

The average yield of cassava in SEC Africa is 8.25 t/ha, which is far below the genetic potential of 90 t/ha (Lebot, 2009; FAO, 2011). One major contributor to this low yield is cassava brown streak disease (CBSD). The disease was first reported at the foothills of the Eastern Usambara Mountains at Amani, Tanzania in 1936 and at that time was only known to occur at lower altitudes (<1000 masl), along the East African coastline paticularly of Kenya, Tanzania and Mozambique and the shores of Lake Malawi (Hillocks and Jennings, 2003). However, the disease has now been confirmed to occur at economic levels at latitudes above 1000 m in Tanzania and Uganda (Jennings, 1960b; Alicai et al., 2007). CBSD is now considered the most important disease of the crop in the SEC Africa region after cassava mosaic disease (CMD) (Hillocks and Jennings, 2003). Most varieties are susceptible to the disease and infection can cause 100% yield loss under severe infection (Alicai et al., 2007). CBSD has attained pandemic status severely affecting cassava in Kenya, Mozambique, Tanzania and Uganda (Mohammed et al., 2012; Alicai et al., 2007). CBSD has also been detected in the Democratic Republic of Congo (Mulimbi et al., 2012) and Burundi (Bigirimana et al., 2011) and unpublished records of the disease exist from Rwanda (Mulimbi et al., 2012) with suspected cases from Madagascar (IITA, 2011). It is feared that if the disease is not brought under control quickly, it may spread to West Africa, one of the major cassava producing regions of the world.

Breeding for host resistance is an important strategy to control the disease, given that resistance to the disease exists and is heritable (Jennings, 1957). Conventional approaches for breeding resistance to both CMD and CBSD began at Amani in Tanganyika (now Tanzania) in 1937 (Jennings, 1957). The most successful strategy involved production of interspecific hybrids using Ceara rubber, *Manihot glaziovii* Muell-Arg as the wild progenitor. Moderately resistant Ceara hybrids such as 46106/27, 37244 and 4763/16 were used in all the trials, and later 46106/27 was replaced by the resistant Ceara hybrid 5318/34. Many of the cultivars were distributed to the breeding programs in the region and eventually diffused into farmer's fields. These genotypes had the wild genome component. The identity and pedigree of most of these cultivars derived from the program has remained unclear. Knowledge on genetic relatedness based on pedigree could suggest whether the genetic basis of resistance is likely to be similar or different. Greater genetic gain in breeding for durable resistance can be achieved if parents are diverse, but pedigree information on cassava accessions is limited, so molecular diversity assessment is useful.

Several studies have used simple sequence repeat (SSR) markers in diversity assessments in cassava (Fregene et al., 2003; Kizito et al., 2005; Kawuki et al., 2013). In a recent study, SSR-defined quantitative trait loci (QTLs) putatively associated with resistance to CBSD have been mapped using a full-sib population involving *Namikonga* as the tolerant parent and *Albert* as the susceptible parent (Kulembeka, 2010). The frequency of the SSR alleles matching those present in *Namikonga* and associated with putative QTLs is not known in cassava germplasm from SEC Africa. Assuming that these alleles are genuinely associated with resistance, knowledge of their frequency in regional germplasm is important in assessing the usefulness of these alleles in marker-assisted selection (MAS).

The objectives of this study were to: (1) determine the amount and structure of the genetic diversity of cassava in SEC Africa using SSR markers, (2) determine the frequency and geographical distribution of alleles putatively associated with *Namikonga* derived resistance to CBSD in order to assess their potential utility in MAS, and (3) examine the genetic relationships among some CBSD resistant genotypes in terms of alleles defining putative QTLs for CBSD tolerance from *Namikonga*.

MATERIALS AND METHODS

Plant material

A total of 192 cassava DNA samples were selected from seven countries in SEC Africa. This comprised 151 landraces and 41 cultivars. Landraces are farmer-varieties and cultivars are varieties derived from formal breeding programs. Kenya was represented by 17 landraces and six cultivars, Tanzania by 27 landraces and five cultivars, DRC by 19 landraces and five cultivars, Madagascar by 17 landraces and eight cultivars, Mozambique by 23 landraces and six cultivars, Rwanda by 15 landraces and five cultivars and Uganda by 33 landraces and six cultivars. Twenty-two (22) of these genotypes had known reaction to CBSD and were previously used in classical genetic studies (Table 1) (Kulembeka and Masumba per. com; Kulembeka, 2010; Zacarias, 2008; Munga, 2008). *Namikonga* and *Albert,* parents of the CBSD mapping population from which QTLs were derived (Kulembeka, 2010) were also included as resistant and susceptible controls, respectively, in this study. In describing plant virus resistance, breeders emphasize the effect on yield and quality in contrast to pathologists who consider the fate of the virus in the plant (Lapidot and Friedmann, 2002). Thresh et al. (1998) states that, 'truly resistant varieties are not readily infected, even when exposed to large amounts of vector-borne inoculum; when infected they develop inconspicuous symptoms that are not associated with obvious deleterious effects on growth and yield and support low virus content and thus likely to be poor source of inoculum from which further spread can occur'. Here we use the term 'resistance' to imply 'field resistance'. A field resistant cultivar may get infected but root symptoms (necrosis) are very mild (class 2 or below on a subjective scale of 1 to 5, where 1

Table 1. Variety type and reaction to CBSD of some selected cassava genotypes.

Genotype	Origin	Type	Field resistance
Albert[3]	Mtwara, Tanzania	Landrace	Susceptible
Amani/46106/27[3&6]	Amani hybrid	BC_3 *M.glaziovii*	Resistant
Bamunanika[5]	Bamunanika, Uganda	Landrace	Susceptible
Chigoma Mafia[1&6]	Mozambique	Landrace	Resistant
IMM30025[1]	IITA, Mozambique	Elite	Susceptible
Kalolo[3]	Bagamoyo, Tanzania	Landrace	Resistant
Kalulu[6]	Tanzania	Landrace	Resistant
Kibaha[3]	Kibaha, Tanzania	Elite	Susceptible
Kibandameno[2&6]	Kenya	Landrace	Susceptible
Kigoma Red[3&6]	Kigoma, Tanzania	Landrace	Resistant
Kiroba[3&6],	Kibaha, Tanzania	Landrace	Resistant
Lihipukalye[4]	Tanzania	Landrace	Resistant
Macia1[1]	Mozambique	Landrace	Resistant
Mfaransa[4&6]	Tanzania	Landrace	Resistant
Mulaleia[1&6],	Mozambique	Landrace	Resistant
Muzege[6]	Tanzania	Landrace	Resistant
MZ89186[1]	IITA, Mozambique	Elite	Susceptible
Nachinyaya[3&6],	Mtwara, Tanzania	Landrace	Resistant
Namikonga[3&6]	Mtwara, Tanzania	Landrace	Resistant
NASE4[5]	Uganda	Serere half sib	Susceptible
NDL90/034[3]	NARI, Tanzania	Kibaha half sib	Resistant
TZ/06/140[5]	Introduction, Uganda	Kibaha half sib	Susceptible
UKG93/041[3]	Tanzania	Kibaha half sib	Resistant

Source: Reports on reaction of genotypes to CBSD used in the present study were from several sources marked in superscript. 1 = Zacarias (2008); 2 = Munga (2008); 3 = Kulembeka (2010); 4 = Kulembeka and Musumba pers. comm.; 5 = Field observations by the first author and colleagues at Namulonge Uganda; 6 = Kanju et al. (2010).

implies no visible symptoms and 5 = very severe symptoms) and reduction in growth and yield is inconspicuous.

SSR selection and PCR analysis

Thirteen (13) SSR primer pairs used in this study were previously used in QTL analysis by Kulembeka (2010). Of the 13 SSR loci analyzed, 10 were associated with the putative CBSD resistance in 'Namikonga x Albert' and 'Namikonga x Namikonga' mapping populations (SSRY4, ESSRY96, NS945, SSRY195, NS667, NS78, NS235, SSRY19, SSRY295 and NS347). The additional three loci were SSRY71, SSRY179 and SSRY242. These markers exceeded the LOD threshold in at least one location in one season (Kulembeka, 2010). These QTLs were identified from a small mapping population consisting of 60 F_1s and 51 S_1s which were used in the final linkage analysis, with phenotypic data from a single season and site in some cases. For these reasons, the accuracy of the effect of these QTL is not high.

Amplification using fluorescently labeled primers was performed in 10 µl reaction volume consisting of 1x PCR buffer (10mM Tris HCl at pH 8.3 and 50 mM KCl), 2 mM $MgCl_2$, 0.2 mM dNTPs, 0.08 pmoles each of the forward and reverse primers each, 0.375 U Taq polymerase and 50 ng of DNA as a template. The PCR profile consisted of 95°C for 2 min, followed by 35 cycles of 95°C for 30 s, 57°C for 1 min and 72°C for 1 min and a final extension at 72°C for 30 min. For M13 primers, amplifications was done in 10 µl volume consisting of 1x PCR buffer, 2 mM $MgCl_2$, 0.2 Mm dNTPs, 0.08 pmoles of the forward primer and 0.32 pmoles of reverse primer, 0.375U Taq Polymerase, 0.32 pmoles of universal tail and 50 ng of DNA as a template. The touchdown PCR was used: an initial denaturing step of 5 min at 95°C followed by 35 cycles of 95°C for 1 min, annealing profile as described below and extension at 72°C for 1 min, respectively. The annealing temperature was decreased by 1°C per cycle, from 57°C in the first cycle to 52°C after the 5th cycle, and then kept constant for the remaining 30 cycles. After 35 cycles, a final extension step was performed at 72°C for 20 min.

Data analysis

Genetic diversity parameters by country of origin

Allele frequencies were calculated and private alleles were identified using 'CONVERT' (Glaubitz, 2004). This software was also used to produce data format for different software packages. Further, calculation of diversity assessment parameters including number of effective alleles, $Ne = 1 / \Sigma p_i^2$; Shannon information index, a measure of amount of diversity in a population $(I) = -1*\Sigma (p_i * Ln (p_i))$; observed heterozygosity, H_o = No. of Hets /n; expected heterozygosity, $H_e = 1 - \Sigma p_i^2$; fixation index, $F = 1 - (H_o / H_e)$ under random mating (where, p_i is the frequency of the ith allele for the population, and Σp_i^2 is the sum of the squared population allele frequencies) were computed using the GenAlex 6.1 computer

Table 2. Distribution and frequency by country of 10 SSR alleles from eight loci which are putatively associated with resistance to CBSD.

Marker	Allele size (bp)	Allele frequency								
		DRC	Kenya	Ma'scar	Moza'	RWAD	TZ	UG	Av	Var
ESSRY96*	167	0.48	0.31	0.56	0.45	0.56	0.47	0.64	0.50	0.0105
NS78	383	0.29	0.30	0.16	0.40	0.43	0.37	0.49	0.36	0.0103
SSRY4	266	0.25	0.22	0.22	0.19	0.18	0.16	0.15	0.19	0.0011
ESSRY96*	179	0.16	0.31	0.24	0.19	0.16	0.16	0.07	0.17	0.0050
NS235*	173	0.09	0.11	0.08	0.17	0.20	0.19	0.18	0.15	0.0022
NS235*	195	0.09	0.17	0.06	0.07	0.18	0.21	0.17	0.14	0.0032
SSRY19	198	0.02	0.02	0.08	0.05	0.03	0.18	0.10	0.08	0.0030
SSRY295	185	0.12	0.02	0.00	0.16	0.08	0.10	0.01	0.07	0.0032
NS667	206	0.10	0.09	0.00	0.05	0.00	0.08	0.04	0.06	0.0017
NS945	375	0.00	0.00	0.00	0.00	0.03	0.11	0.00	0.02	0.0015

(*) = SSR marker exhibiting more than one polymorphic allele.

software (Peakall and Smouse, 2006). Arlequin ver3.5 was used to compute population fixation indices and analysis of molecular variance (AMOVA) to assess the genetic differentiation among countries (Excoffier and Schneider, 2005).

Frequency of alleles putatively associated with CBSD resistance

The frequency of alleles putatively associated with CBSD resistance was examined in the germplasm from SEC Africa to determine whether they occurred at low enough frequency to be useful in marker-assisted breeding if the association of that marker with CBSD resistance is confirmed. In the study, frequency distribution was categorized thus: very low = ≤ 5%, moderately low >5% to ≤10%, high = ≥10% to ≤50% and very high = >50%. Data was also examined for private alleles which are defined as those that occur in a single country or germplasm source only.

Genetic relationships among individual genotypes

Simple-matching Euclidean distance was used to compute a distance matrix among all individuals using PowerMarker version 3.25 (Liu and Muse, 2005). Clustering, illustrated in the form of a dendrogram, was constructed using the weighted neighbor-joining algorithm in DARwin5.0 (Perrier and Jackquemound-Collet, 2006). Those clusters containing genotypes with known field response to CBSD infection were subjected to further analysis and interpretation.

RESULTS

Genetic diversity by locus

All the 13 SSR primers amplified were polymorphic across the 192 samples assayed and the summary of 10 alleles associated with resistance to CBSD is presented in Table 2. A total of 98 alleles were detected with an average of seven alleles per locus with the lowest (4) for SSRY192 and the highest (11) for NS347 with an average of five alleles per country. The mean PIC value was 0.62 with a range of 0.49 to 0.79. The mean value for He and Ho were 0.63 and 0.69, respectively (Table 3). Ho was higher than He depicting an excess of heterozygotes at most loci examined except SSRY19.

Partitioning of variation

Results of AMOVA presented in Table 5 indicate little genetic differentiation among countries (7%), with the vast majority of variation due to heterozygosity within individuals (93%). This is supported by an F_{IS} of -0.090 which indicate a slight excess of heterozygotes compared to expectations from random-mating. In general, if F_{ST} is 0.05 to 0.15 then differentiation in allele frequencies between populations is considered moderate. Here, an F_{ST} of 0.027 indicates little differentiation.

Genetic diversity by country of origin

Levels of genetic diversity among countries was similar with a mean Shannon information index of I=1.23, and SE=0.03 in range of 1.17 to 1.33 (Table 3). A total of 14 genotypes (7.3%) exhibited 17 unique alleles occurring in five different countries with the highest number of alleles (10) detected in Tanzanian landraces with an average of three to four unique alleles per genotype at five loci (Data not presented). Malagasy germplasm exhibited the second highest number of unique alleles (nine) at three loci, all in elite genotypes. Low numbers of unique alleles were recorded in Kenyan (1), Ugandan (2) and Mozambican (2) germplasm. No unique alleles were detected in populations from Rwanda and DRC. Unique alleles here are defined as alleles only found in a particular country / population. The three cultivars from Tanzania that exhibited unique alleles also possess some of the following alleles putatively associated with CBSD resistance: Mkukumkuku NS667 (206 bp), NS945 (375 bp), Muzege NS945 (375 bp), ESSRY96 (167 bp), Mfaransa NS667 (206 bp), NS945 (375 bp), NS235 (123 bp) and ESSRY (176 bp).

Table 3. Mean and SE over loci for each population for level of heterozygosity, F-statistics and polymorphism.

Population	Mean/SE	Na	Ne	I	Ho	He	F
DRC	Mean	5.08	3.01	1.27	0.69	0.65	-0.05
	SE	0.35	0.18	0.05	0.04	0.02	0.05
Kenya	Mean	4.77	2.92	1.19	0.67	0.62	-0.07
	SE	0.32	0.27	0.08	0.05	0.03	0.04
Madagascar	Mean	5.00	2.72	1.17	0.66	0.60	-0.11
	SE	0.45	0.22	0.08	0.04	0.03	0.04
Mozambique	Mean	5.23	3.28	1.33	0.74	0.67	-0.10
	SE	0.38	0.25	0.08	0.03	0.02	0.03
Rwanda	Mean	5.08	2.81	1.20	0.65	0.61	-0.04
	SE	0.38	0.26	0.09	0.06	0.04	0.07
Tanzania	Mean	5.23	3.34	1.27	0.72	0.64	-0.11
	SE	0.41	0.42	0.11	0.06	0.04	0.03
Uganda	Mean	4.92	2.98	1.21	0.69	0.62	-0.11
	SE	0.31	0.36	0.09	0.04	0.03	0.03
Total	Mean	5.04	3.01	1.23	0.69	0.63	-0.08
	SE	0.14	0.11	0.03	0.02	0.01	0.02

Na, Number of different alleles; Ne, number of effective alleles; I, Shannon's information index; Ho, observed heterozygosity; He, = expected heterozygosity / gene diversity; F = fixation index.

Frequency of alleles putatively associated with CBSD resistance

Ten alleles were putatively associated with CBSD resistance (Kulembeka, 2010) however in this study, alleles at ESSRY96 and NS235 were found to be identical in *'Namikonga'* and '*Albert'*. The frequency of the alleles from the remaining eight loci putatively associated with resistance to CBSD and derived from the CBSD resistant parent, *Namikonga*, were examined in the 192 varieties from SEC Africa (Table 2). These markers exhibited a total of 10 alleles and occurred at frequencies from 2.0% for NS945-375 bp to over 50% for ESSSRY-167 bp (Table 2). Other alleles (NS667-206bp, SSRY295-185bp and SSRY19-198bp) occurred at moderately low frequency (<10%). The variance in allele frequency across countries was low and ranged from 0.0011 to 0.0103. All alleles were detected in each country except SSRY295-185 bp which was not detected in Madagascar; NS667-206 bp was not detected in Madagascar and Rwanda and NS945-375 bp which was only present in Rwanda at allele frequency of 0.03 and Tanzania at 0.11. On average, the greatest number of alleles with high frequencies was recorded in Tanzania and of low frequency alleles was recorded in Madagascar.

Genetic relationships among individual genotypes

The dendrogram in Figure 1 revealed 12 clusters with cluster 1 and 5 having three and two sub-clusters, respectively. For the purpose of this study, clusters marked with asterisk in Figure 1 were subjected to further analysis and interpretation on the basis of the eight SSR markers putatively associated with resistance to CBSD. The selected clusters are named according to one principal genotype with known record for field resistance to CBSD and the most common allele(s) that characterizes the cluster. Table 4 provides a summary of selected genotypes by cluster. The detailed results of this analysis are presented below.

Cluster 1 contains three subclusters, however subcluster 1a does not contain any genotype with known resistance or susceptiblity and therefore was not subjected to further analysis. Subclusters 1b and 1c were further analysed as follows:

The *Nachinyaya* cluster (1b) is comprised of 10 genotypes from four countries of DRC, Mozambique, Tanzania and Uganda. *Nachinyaya* is a CBSD resistant cultivar from Tanzania (Kanju et al., 2010) with ESSRY96-167bp allele. In the same cluster, two of the *Namikonga*-identified putative alleles for resistance, NS667-206bp and NS945-375bp, were identified in cv. *Mkukumkuku* only. This cluster is characterized by allele SSRY179-184bp as the most frequent but not associated with resis-tance. *Nachinyaya* genotype does not carry any of the putative CBSD resistance alleles identified in *Namikonga*. All the genotypes in this cluster are of landrace origin.

The *Kalulu* cluster (1c) is characterized by the allele SSRY179-198bp which was not found associated with Namikonga–identified resistance in the present study. It includes many genotypes from Tanzania, including *Kalulu* which is a CBSD tolerant cultivar (Kanju et al., 2010).

However, cv. *Kalulu* does not carry any of the Namikonga

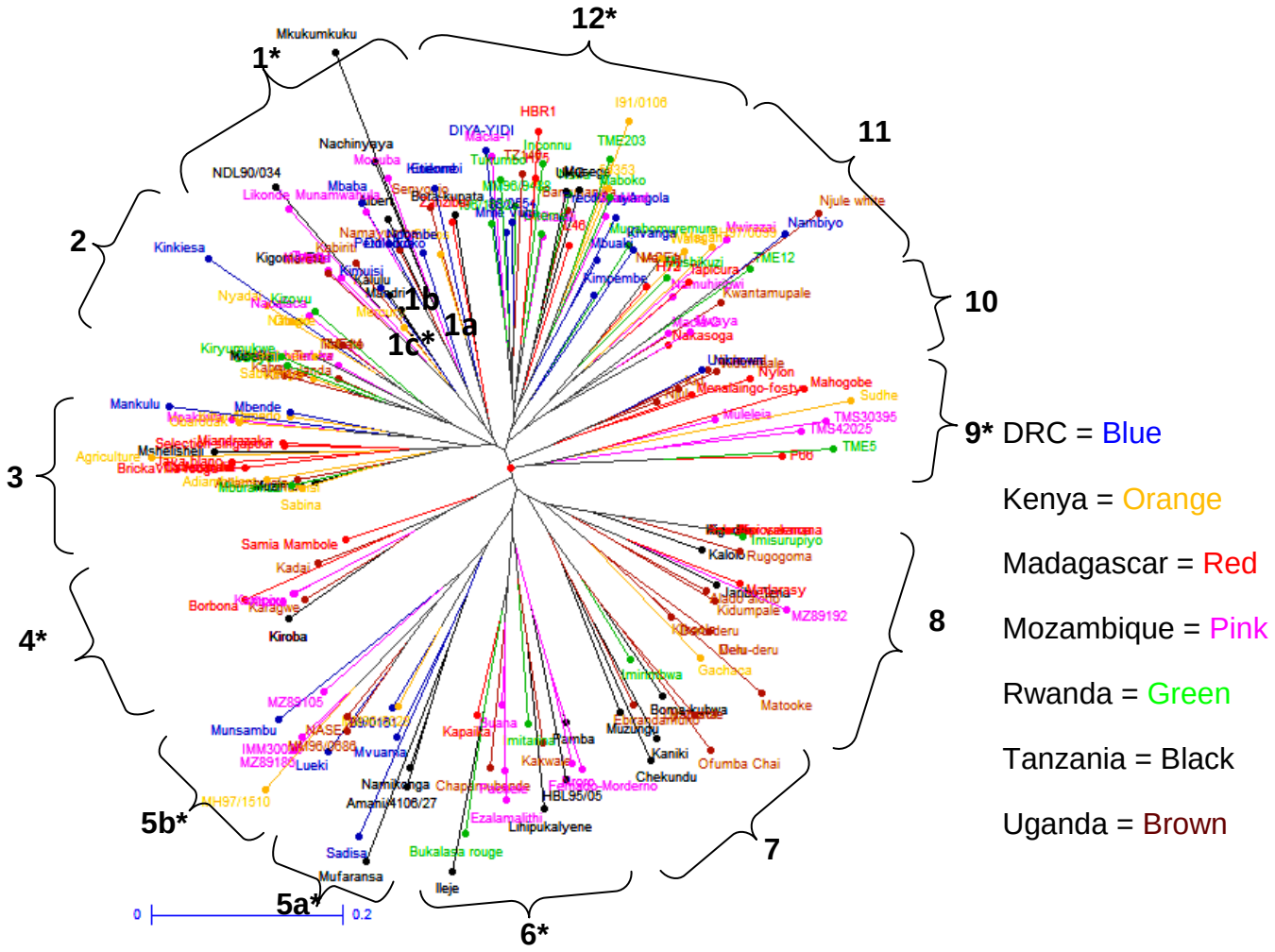

Figure 1. Clusters of cassava genotypes produced by the weighted neighbor-Joining method. Genotype names are colored by country. Clusters are discussed in the text in relation to pedigree and reaction to CBSD. *Clusters 1, 4, 5a, 5b, 6, 9 and 12 were subjected to further analysis.

CBSD associated alleles. The cluster also contains Kigoma Red, a known source of resistance to the disease. Most genotypes in this cluster are landraces except NDL90/034 which is an officially released cultivar (Naliendele) in Tanzania. NDL90/034 was derived from a half-sib family of cv. Kibaha. Cluster 1c is closely related to cluster 1b and may contain a source of resistance that might be a close variant of Nachinyaya resistance and might have a shared recent common pedigree.

The *Kiroba* cluster (4), named after a CBSD resistant cultivar from Tanzania (Kanju et al., 2010), includes only landraces from Madagascar, Mozambique, Tanzania and Uganda. This group does not possess alleles with low frequency such as NS667-206bp and NS945-375bp. The cluster is characterized by allele SSRY179-182 bp which is not associated with the resistance to CBSD and occurs in the majority of the genotypes including Kiroba. This cluster also contains cv. *Chigoma mafia* which is a landrace from Mozambique. It is suspected that cv. *Chigoma mafia* is an Amani derivative.

Namikonga cluster (5) is composed of two sub-clusters (a) and (b), and is characterized by allele 206 bp (NS667) which occurs in all the genotypes with the majority of those genotypes coming from Tanzania, DRC and Mozambique. This allele is putatively associated with resistance to CBSD. The sub-cluster (a) consists of genotypes mainly from DRC and Tanzania. The cluster is named after *Namikonga*, which is a resistant cultivar in Tanzania (Kanju et al., 2010). This sub-cluster uniquely contains the informative NS945-375bp allele in all the genotypes from Tanzania and also the 399bp allele of the same locus in six genotypes. Some of genotypes in the cluster carry the allele of SSRY179-204bp. *Namikonga* and the Amani hybrid 46106/27 fell in this sub-cluster and both carry all the *Namikonga*-identified alleles for resistance to CBSD. Both cultivars exhibit resistance to CBSD in Tanzania. The clustering of the same cultivars is in agreement with earlier observations that this Amani

Table 4. Clusters of cassava genotypes according to the SSR markers associated with resistance to cassava brown streak disease in SEC region of Africa.

Cluster name	Genotype	Description	Country of Origin
Nachinyaya cluster (1b)	Albert	Landrace	Tanzania
	Kimpembe	Landrace	DRC
	Mbaba	Landrace	DRC
	Mkukumkuku	Landrace	Tanzania
	Mocuba	Landrace	Mozambique
	Munamwahula	Landrace	Mozambique
	Nachinyaya	Landrace	Tanzania
	Namayumba	Landrace	Uganda
	Petit Nkoko	Landrace	DRC
	Senyonjo	Landrace	Uganda
Kalulu (1c)	Likonde	Landrace	Mozambique
	Kabiriti	Landrace	Uganda
	Kalulu	Landrace	Tanzania
	Kigoma Red	Landrace	Tanzania
	Mandri	Landrace	Tanzania
	NDL90/034	Elite	Tanzania
Kiroba cluster (4)	Borbona	Landrace	Madagascar
	Chigoma Mafia	Landrace	Mozambique
	Kadai	Landrace	Uganda
	Karagwe	Landrace	Uganda
	Kiroba	Landrace	Tanzania
	Samia Mambole	Landrace	Madagascar
	Xitaxe	Landrace	Mozambique
Namikonga sub-cluster (5a)	99/0161	Elite	DRC
	Amani 46106/27	Elite	Tanzania
	MM96/5028	Elite	Kenya
	Mfaransa	Landrace	Tanzania
	Mvuama	Landrace	DRC
	Namikonga	Landrace	Tanzania
	Sadisa	Elite	DRC
Namikonga sub-cluster (5b)	IMM30025	Elite	Mozambique
	Lueki	Elite	DRC
	MH97/1510	Elite	Kenya
	MM96/0686	Elite	Uganda
	Munsambu	Landrace	DRC
	MZ89105	Elite	Mozambique
	MZ89186	Elite	Mozambique
	NASE4	Elite	Uganda
Lihipukalyene cluster (6)	Buana	Landrace	Mozambique
	Bukalasa rouge	Landrace	Rwanda
	Chapamubende	Landrace	Uganda
	Fernado Morderno	Landrace	Mozambique
	HBL95/05	Elite	Tanzania
	Imitarina	Landrace	Rwanda

Table 4. Contd.

	Kakwale	Landrace	Uganda
	Lihipukalyene	Landrace	Tanzania
	Lleje	Landrace	Tanzania
	Pachele	Landrace	Mozambique
Mulaleia cluster (9)	Mahogobe	Landrace	Madagascar
	Mulaleia	Landrace	Mozambique
	Pamba	Landrace	Tanzania
	TME5	Elite	Rwanda
	TMS30395	Elite	Mozambique
	TMS42025	Elite	Mozambique
Macia cluster (12)	88/0554	Elite	DRC
	Bamunanika	Landrace	Uganda
	Diya Yidi	Landrace	DRC
	HBR1	Elite	Madagascar
	I96/1632	Elite	Rwanda
	Inconnu	Landrace	Rwanda
	L46	Elite	Madagascar
	Maboko	Landrace	Rwanda
	Macia 1	Landrace	Mozambique
	Mbuaki	Landrace	DRC
	MM96/9488	Elite	Rwanda
	Muzege	Landrace	Tanzania
	Precoce d' Angola	Elite	DRC
	TME203	Elite	Rwanda
	Tukumbo	Landrace	Rwanda
	UKG93/041	Elite	Tanzania

Table 5. Analysis of molecular variance (AMOVA).

Source	Degrees of freedom	Sum of Squares	Mean Square	Estimated variance	% Variance
Among populations	6	142.19	23.7	0.573	7
Within populations	185	1502.52	8.12	8.122	93
Total	191	1644.71		8.694	100

hybrid could be the same genotype as *Kaleso* and *Namikonga* (Kanju et al., 2010, Hillocks and Jennings 2003). The cluster consists mainly of elite clones apart from *Mfaransa*, and *Mvuama* which are referred to as landraces from Tanzania and DRC, respectively. *Mfaransa* is known to be tolerant to CBSD (Kanju et al., 2010).

Sub-cluster 5a includes two elite cultivars from Mozambique, cv. IMM30025 and MZ89186 which are all CBSD susceptible. This cluster is referred to as the '*Namikonga* late pedigree' cluster because of its close genetic distance with the *Namikonga* cluster and therefore might have shared pedigree. In sub-cluster 5b, the majority of genotypes are from Mozambique and are uniquely characterized by the NS945-389bp allele. This allele is not associated with *Namikonga* - identified resistance and these genotypes may be CBSD susceptibility variants of *Namikonga*. In this sub-cluster (5b), most of the genotypes are of elite germplasm except the *Musambu* landrace from DRC.

Lihipukalyene cluster (6) consists of a majority of genotypes from Mozambique and Tanzania and is characterized by cv. *Lihipukalyene* which is a CBSD resistant genotype from Tanzania (Kulembeka and Masumba per. com). The cluster does not contain any of the low frequency alleles putatively associated with CBSD. The

allele SSRY179-182bp occurred in relatively equal proportions in all the countries in the cluster and allele SSRY179-204bp (not informative for CBSD resistance) only occurred in two genotypes. Most genotypes in this cluster are landraces except HBL95/05 from Tanzania.

The *Mulaleia* cluster (9) is composed of genotypes mainly from Mozambique and Madagascar. *Mulaleia* is a resistant landrace from Mozambique. This cluster is mainly characterized by allele SSRY179-202bp and only one allele 182bp from the same loci was detected in two individuals; one from Madagascar and the other from Rwanda. *Mulaleia* genotype does not carry any of the low frequency alleles putatively associated with resistance to CBSD.

The *Macia* cluster (12) is the most diverse in terms of geographical distribution, and is characterized by *Macia1,* which is a resistant landrace from Mozambique. Most of the genotypes in this cluster are from DRC and Rwanda. All the four alleles putatively associated with resistance to CBSD were detected in the cluster in seven genotypes. However, *Macia1* genotype does not carry any of the low frequency alleles informative for resistance to CBSD except SSRY295-185 bp. The allele that characterizes this cluster and occurred in high frequency was NS945-381bp and is not informative for resistance to CBSD. This cluster contains two other popular resistant cultivars UKG93/041 and Muzege both from Tanzania.

DISCUSSION

Of the initial ten SSRs putatively associated with CBSD resistance in an F_1 population between *Namikonga* and *Albert*, only eight loci were found to be variable between the two. This suggests variation within the farmer-variety or landrace *Albert*. Similar observations were made during linkage disequilibrium mapping studies (Kulembeka, 2010).

Genetic diversity

Shannon's information index revealed low levels of genetic variation within cassava germplasm from SEC Africa. Although there were a relatively large number of alleles, they occurred at low frequency and thus do not contribute much to diversity. This is indicative of a genetic bottleneck in the region and consistent with the findings of Kawuki et al. (2013). Greater levels of diversity have been reported from different regions, countries or collections (Montero-Rojas et al., 2011; Fregene et al., 2003). Cassava is a clonally propagated and predominantly an out-crossing crop with new variations being generated from natural crossings and the traditional regular maintenance of volunteers particularly in small-holder farming systems (Silva et al., 2003; Doyle et al., 2001).

Genetic structure

The limited genetic differentiation in cassava germplasm among countries from SEC Africa found in this study confirms the earlier report by Kawuki et al. (2013) who found some sub-structure with landraces from Tanzania occupying a pivotal position, and those from Mozambique and DRC being more distant. In that study, cultivars from Tanzania and Madagascar were somewhat distanced from those from other countries. This limited differentiation could be attributed to: (1) dispersion by farmers who traditionally exchange cassava stakes even between countries since phyto-sanitary regulations restricting the movement of cassava germplasm across international borders have only come into effect relatively recently, (2) similar selection pressures by farmers across countries, and (3) movement of improved varieties among breeding programs which have then been disseminated to farmer's fields. The lack of substantial differentiation implies that analyzing genetic diversity in terms of pedigree or genetic relatedness of individual clones may be more informative than on geographical origin alone.

Unique alleles

The high number of unique alleles in Malagasy and Tanzanian germplasm is likely to reflect germplasm introductions and inter-specific breeding work in these countries to mitigate CMD epidemics as well as breed for CBSD resistance. Madagascar experienced a devastating epidemic of CMD in 1934-36 which led to the almost total elimination of local varieties (Cours et al., 1997). During this time, a number of germplasm lines were introduced, including Javanese varieties, and intra-specific crosses were undertaken. In addition, inter-specific crosses were undertaken that included ceara rubber (*M. glaziovii*). Inter-specific crosses were also undertaken from the 1930s at Amani, Tanzania (Nichols, 1947). High levels of resistance to CMD were obtained from *M. glaziovii x M. esculenta* crosses (Jennings, 1994). Before the Amani breeding program collapsed, the valuable germplasm was distributed to various breeding programs in Africa (Beck, 1982). The initiatives in Madagascar and Tanzania generated new genetic variability and could explain the relatively large number of unique alleles in this germplasm. In addition, the high number of unique alleles in Madagascar could be attributed to isolation from mainland Africa which restricted germplasm exchange both informally and formally and limited gene flow between the two populations (Lebot, 2009). The restriction might have prevented some alleles that were introduced directly from South America to Madagascar from crossing into the mainland of East Africa.

Distribution of alleles putatively associated with resistance to CBSD

The frequency of alleles associated with resistance to CBSD was found to be generally low but widely dispersed in the cassava germplasm of the SEC Africa region. This implies that if the association between these

alleles and resistance to CBSD in *Namikonga* is validated, then these alleles are likely to be useful for MAS in the majority of germplasm used in breeding programs in SEC Africa. For markers to be useful there must be polymorphism between the donor allele(s) of *Namikonga* and the recipient allele(s) at the locus associated with resistance.

One of the most disease resistant clones resulting from the Amani program was Amani 46106/27 which is grown in Kenya under the name *Kaleso* (Hillocks and Jennings, 2003), and in Tanzania as *Namikonga*. Kanju et al. (2010) could not differentiate between *Namikonga* and *Kaleso* using over 500 single nucleotide polymorphism (SNP) markers. The Amani germplasm was widely distributed after the end of the program. For example, Amani germplasm was introduced to Uganda at the government Experimental station at Bukalasa in 1934 (Nicols, 1950) and Moore Estate in Nigeria where it was utilized by IITA cassava breeding program.

Interestingly, the same genotypes in Tanzania that carry the unique alleles also carry the alleles putatively associated with the QTLs that define the *Namikonga*-derived resistance to CBSD. This suggests that the unique alleles may be derived from *M. glaziovii*. The alleles associated with resistance to CBSD were subsequently widely distributed through all countries in the region during the early dissemination of the Amani-derived cassava germplasm although in varying proportions.

The current study shows that not all the CBSD resistant varieties do cluster together. Detailed analysis of clusters containing varieties with known CBSD resistance suggests different sources of resistance with possibilities of different genetic control mechanisms. This in turn implies the opportunity for combining these sources of resistance for more effective and durable resistance to CBSD.

It is possible, based on genetic relationships in Figure 1 and Table 4 that some of the key genotypes that exhibit field resistance to CBSD such as *Lihipukalyene, Nachinyaya*, *Kalulu* from Tanzania, *and Muleleia* and *Macia1* from Mozambique are carrying different resistance genes than those derived from *Namikonga.* This evidence supports an earlier hypothesis by Jennings (1957) that, 'probably, in most cases, a combination of genes from both sources was present; different genes were probably present in the three different lines of hybrids descended from the three first backcross clones used as parents for the second backcross and also in the sub-lines descended from the several second backcross clones later used as parents.

Furthermore, since cassava was recently introduced to Africa and given the fact that CBSD is only found in Africa, it is likely that the resistance in cassava to CBSD co-evolved with the increased disease pressure (Rausher, 2001). Such information on genetic relationships among cultivars with respect to their reaction to CBSD in the region could provide an indication of where novel sources of CBSD resistance may be found for future resistance breeding. However, the above results are based on 13 SSR markers and therefore, there is need for further validation with additional markers. In addition, analysis for detection of QTL in a range of other varieties would provide greater insights into the genetic basis of CBSD resistance.

Implications for breeding resistance to CBSD

The low frequency in SEC Africa of alleles putatively associated with resistance to CBSD implies that if validated they may be useful in MAS. In addition, the apparently diverse sources of CBSD resistance offer potential for breeding for more effective and durable resistance by pyramiding different resistance genes or through deployment strategies that avoid imposing a strong selection pressure.

Although no definitive answers on the genetic basis of CBSD can be given here due to non-validation of markers and the small population size used for mapping, this study does provide some initial data and a framework for assessing variation for disease resistance. Before more concrete direction can be drawn from this study, it is important that QTL are validated on a larger population size. Subsequent to this study, QTL analysis has been performed using a larger number of SNP markers which provide greater resolution to QTL. These markers will be more appropriate for MAS for CBSD resistance derived from *Namikonga*. These SNP markers are currently being validated in Uganda in collaboration with IITA-Nairobi.

There is an urgent need to screen additional germplasm for resistance to CBSD. The germplasm used in this type of study should also be expanded to include additional sources of resistance. In addition, several new markers could be added (SNPs or SSRs) from the scaffolds of the cassava genome sequence on which the QTLs lie.

Conclusions

The following conclusions can be drawn from this study: (1) the frequency of alleles putatively associated with resistance to CBSD was found to be generally low but widely dispersed in the cassava germplasm of the SEC region, (2) little genetic differentiation was confirmed among countries and relatively low genetic diversity with the majority of variation distributed among individuals, (3) several distinct putative sources of resistance to CBSD seem to be present in the cassava germplasm of the SEC region; and (4) it may be more informative to assess cassava diversity in the SEC Africa region based on pedigree (if known) or genetic relatedness, rather than geographical origin. Once markers are validated, the information on relatedness provided here can provide a basis for further investigation on genetic basis of CBSD resistance to identify different sources of resistance that may be combined for more durable resistance.

ACKNOWLEDGEMENTS

We thank the World Bank through its project coded: Millennium Science Initiative through the government of Uganda represented by National Council of Science and Technology for funding this study. We thank the International Livestock Research Institute (ILRI) - BioSciences eastern and central Africa (BecA) hub for permitting us to carry out the experiment using their laboratory facilities. We are indebted to NARS of the regional for providing germplasm for the study. We are grateful to Ms Esther Masumba and Dr. Heneriko Kulembeka of National Cassava Breeding program, Tanzania, and Dr. Edward Kanju (IITA) for providing information on field reaction of some key genotypes to CBSD in Tanzania.

REFERENCES

Alicai T, Omongo CA, Maruthi M, Hillocks RJ, Baguma Y, Kawuki R, Bua A, Otim-Nape GW, Colvin J (2007). Re-emergence of cassava brown streak disease in Uganda. Plant Dis. 91:24-29.

Allem AC, Mendes RA, Salomao AN, Burle ML (2001). Primary gene pool of cassava (*Manihot esculenta* Crantz subspecies *esculenta.* Euphorbiaceae). Euphytica. 120:127–132.

Beck BDA (1982). Historical perspectives of cassava breeding in Africa. In: Hahn SK, Ker ADR (eds). Root Crops in Eastern Africa. Proceedings of the workshop held at Kigali, Rwanda. November 1980 (Ottawa: IDRC). pp.13-17.

Bigirimana S, Barumbanze P, Ndayihanzamaso P, Shirima R, Legg JP (2011). First report of cassava brown streak disease and associated Ugandan cassava brown streak virus in Burundi. New Disease Reports. 24:26.

Cours G, Fargette D, Otim-Nape GW, Thresh JM (1997). The epidemic of cassava mosaic virus disease in Madagascar in the 1930s-1940s: lessons for the current situation in Uganda. Trop. Sci. 37:238-248.

Doyle M, Pujol B, Elias M (2001). Ecology and genetics of populations of cassava managed by Amazonian Amerindians: or, how to maximize the benefits and minimize the costs of outcrossing in a vegetatively propagated crop. In: Fauquet C, Taylor NN (eds). Proc 5th Int Scientific Meeting of the Cassava Biotechnology Network, Danforth Plant Science, St. Louis, Missouri.

Excoffier LG, Schneider S (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol. Bioinform Online.1:47-50.

FAO (2011). FAOSTAT|©FAO Statistics Division 2011 | 09 May 2011. http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor. Cited on 20/03/2012.

Fregene MA, Suarez M, Mkumbira J, Kulembeka H, Ndedya E, Kulaya A, Mitchel S, Gullberg U, Rosling H, Dixon AGO, Dean R, Kresovich S (2003). Simple sequence repeat marker diversity in cassava landraces: genetic diversity and differentiation in an asexually propagated crop. Theor. Appl. Genet. 107:1083–1093.

Glaubitz JC (2004). CONVERT: a user friendly program to reformat diploid genetic data set for commonly used population genetic software packages. Mol. Ecol. Notes. 4:309–310.

Hillocks RJ, Jennings DL (2003). Cassava brown streak disease: a review of present knowledge and research needs. Int. J. Pest. Manage. 49:225–234.

IITA (2011). Combating the threat of CBSD. http://r4dreview.org/2011/04/combating-the-threat-of-cbsd/. Cited on 12/12/2011.

IITA (2007). Scientists halt cassava and banana devastation in East and Central Africa. http://www.cgiar.org/newsroom/releases/news.asp?idnews=574. Cited on 20/02/2012.

Jennings DL (1957). Further studies in breeding cassava for virus resistance. The East Afr. Agricul. J. 22:213–219.

Jennings DL (1960b). Observations on virus diseases of cassava in resistant and susceptible varieties. Empire J. Exp. Agricul. 28:261-269.

Jennings DL (1994). Breeding for resistance to African cassava mosaic geminivirus in East Africa. Trop. Sci. 34:110–122.

Kanju E, Mkamilo G, Mgoo V, Ferguson M (2010). Statistical evidence linking the zigzag stem habit with tolerance to cassava brown streak disease. Roots. 12:4–6.

Kawuki RS, Herselman L, Labuschagne MT, Nzuki I, Ralimanana I, Bidiaka M, Kanyange MC, Gashaka G, Masumba E, Mkamilo G, Gethi J, Wanjala B, Zacarias A, Madabula F, Ferguson ME (2013). Genetic Diversity of Cassava (*Manihot esculenta* Crantz) Landraces and Cultivars from Southern, Eastern and Central Africa. Plant Genetic Resources, available on CJO2013. doi:10.1017/S1479262113000014.

Kizito EB, Bua A, Fregene M, Egwang T, Gullberg U, Westerbergh A (2005). The effect of cassava mosaic disease on the genetic diversity of cassava in Uganda. Euphytica. 146:45–54.

Kulembeka HPK (2010). Genetic linkage mapping of field resistance to cassava brown streak disease in cassava (*Manihot esculenta* Crantz) landraces from Tanzania. PhD thesis, University of the Free State, Bloemfontein South Africa. p.281.

Lapidot M, Friedmann M (2002). Breeding for resistance to whitefly transmitted geminiviruses. Ann. Appl. Biol. 140:109 – 127.

Lebot V (2009). Tropical Root and Tuber Crops: Cassava, Sweet potato, Yams and Aroids. Crop production science in horticultural series; 17. MPG Books Group, UK.

Liu K, Muse SV (2005). PowerMarker: Integrated Analysis environment for genetic marker data. Bioinformatics. 21: 2128-2129.

Mohammed UI, Arbashi MM, Muli B, Hillocks RJ, Maruthi MN (2012). The symptom and genetic diversity of cassava brown streak viruses infecting cassava in East Africa. Advances in Virology. 2012: Article ID 795697, p.10.

Montero-Rojas M, Correa AM, Siritunga D (2011). Molecular differentiation and diversity of cassava (*Manihot esculenta*) taken from 162 locations across Puerto Rico and assessed with microsatellite markers. AoB PLANTS. 2011 plr010 doi: 10.1093/aobpla/plr010

Munga TL (2008). Breeding for Cassava Brown Streak Resistance in Coastal Kenya. Ph.D thesis, University of KwaZulu-Natal, South Africa.

Mulimbi W, Phemba X, Assumani B, Kasereka P, Muyisa S, Ugentho H, Legg JP, Laurenson L, Weekes R, Thom FEF (2012). First report of Ugandan cassava brown streak virus on cassava in Democratic Republic of Congo. New Disease Reports 26:11. [http://dx.doi.org/10.5197/j.2044-0588.2012.026.011].

Nassar NMA (2002). Cassava, *Manihot esculenta* Crantz, genetic resources: origin of the crop, its evolution and relationships with wild relatives. Genet. Mol. Res. 1:298–305.

Nicols RFW (1947). Breeding cassava for virus resistance. East Afr. Agricul. J. 12:184–194.

Nicols RFW (1950). The brown streak disease of cassava: distribution, climatic effects and diagnostic symptoms. East Afr. Agricul. J. 15:154–160.

Peakall R, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes. 6:288-295.

Perrier X, Jacquemoud-Collet JP (2006). DARwin software. http://darwin.cirad.fr/darwin

Purseglove JW (1974). Tropical crops. Dicotyledons. English Language Book Society / Longman. p. 719.

Silva R, da Bandel G, Martins PS (2003). Mating system in an experimental garden composed of cassava (*Manihot esculenta* Crantz) ethnovarieties. Euphytica. 134:127-135.

Thresh JM, Otim-Nape GW, Fargette D (1998). The components and deployment of resistance to cassava mosaic virus disease. Integr. Pest Manage. Rev. 3:209–224.

Rausher MD (2001). Co-evolution and plant resistance to natural enemies. Nature. 411:857-864.

Zacarias AM (2008). Breeding potential of cassava (*Manihot esculenta* Crantz) in Mozambique. Ph.D thesis. University of Free State, Bloemfontein South Africa. p.169.