The effect of cassava mosaic disease on the genetic diversity of cassava in Uganda

Elizabeth Balyejusa Kizito[1,4], Anton Bua[2], Martin Fregene[3], Thomas Egwang[4], Urban Gullberg[1] & Anna Westerbergh[1,*]

[1]*Department of Plant Biology and Forest Genetics, Swedish University of Agricultural Sciences, Box 7080, SE-750 07, Uppsala, Sweden;* [2]*Cassava Programme, Namulonge Agricultural and Animal Production Institute, P.O. Box 7084, Kampala, Uganda;* [3]*Centro Internacional de Agricultura Tropical (CIAT), AA 6713, Palmira Recta, Cali, Colombia;* [4]*Med Biotech Laboratories, P.O. Box 9364, Kampala, Uganda*

([*]author for correspondence: e-mail: anna.westerbergh@vbsg.slu.se)

Received 30 November 2004; accepted 8 September 2005

Key words: cassava, cassava mosaic disease, CMD, genetic diversity, *Manihot esculenta*, SSR markers

Summary

Cassava (*Manihot esculenta*) is a tropical crop that is grown in Africa, Latin America and Southeast Asia. Cassava was introduced from Latin America into West and East Africa at two independent events. In Uganda a serious threat to cassava's survival is the cassava mosaic disease (CMD). Uganda has had two notable CMD epidemics since the introduction of cassava in the 1850s causing severe losses. SSR markers were used to study the effect of CMD on the genetic diversity in five agroecologies in Uganda with high and low incidence of CMD. Surprisingly, high gene diversity was detected. Most of the diversity was found within populations, while the diversity was very small among agroecological zones and the high and low CMD incidence areas. The high genetic diversity suggests a mechanism by which diversity is maintained by the active involvement of the Ugandan farmer in continuously testing and adopting new genotypes that will serve their diverse needs. However, in spite of the high genetic diversity we found a loss of rare alleles in areas with high CMD incidence. To study the effect of the introgression history on the gene pool the genetic differentiation between East and West Africa was also studied. Genetic similarities were found between the varieties in Uganda and Tanzania in East Africa and Ghana in West Africa. Thus, there is no evidence for a differentiation of the cassava gene pool into a western and an eastern genetic lineage. However, a possible difference in the genetic constitution of the introduced cassava into East and West Africa may have been diminished by germplasm movement.

Introduction

Cassava (*Manihot esculenta)* is a neotropical crop that is also grown in tropical Africa and southeast Asia. It is cross-fertilizing but is vegetatively propagated through stem cuttings. However, in traditional agroecosystems in Latin America small-scale farmers also multiply plants from volunteer seedlings produced by sexual reproduction (Elias et al., 2001). Spontaneous recombination and farmer selection from volunteer seedlings seem also to occur in Africa (Fregene et al., 2003). When visiting farmers in different parts of Uganda, seedlings were seen growing in plots near cassava fields (E. Balyejusa Kizito & U. Gullberg, personal communication). Cassava is believed to have originated from wild *M. esculenta* populations growing along the southern rim of the Amazon Basin in Brazil (Olsen & Schaal, 1999, 2001). Cassava was introduced from Brazil into Africa by Portuguese slave ships and arrived to the western coast in the 17th century and to the eastern coast in the 18th century (Jones, 1959). In Uganda cassava was introduced relatively late and had limited distribution by the 1900s (Jones, 1959; Langlands, 1966). It was most likely introduced from the eastern introgression route by European and Arab traders (Jones, 1959). It was first grown in Buganda, the central part of Uganda, around 1870 and was soon spread to Bunyoro, the northwestern region. Cassava arrived much later to the eastern (Busoga) and northern Uganda in the early 20th century (Langlands, 1966). Because of its excellent adaptability to erratic rainfall and low fertility soils, it became a major dietary staple, a famine reserve crop and a source of cash to many small-scale farming communities.

A threat to cassava's survival as a dominant crop in subsistence communities in Uganda is its vulnerability to the cassava mosaic disease (CMD). The disease is caused by several viruses belonging to the genus *Begomovirus* in the family *Geminiviridae*, transmitted by whiteflies (*Bemisia tabaci*; Pita et al., 2001). The first outbreak of CMD was in the late 1920s (Martin, 1928; Jameson, 1964). At that time there were also frequent famines caused by drought and locusts, especially in northern Uganda. To alleviate these situations the then colonial government vigorously encouraged cassava multiplication schemes and land area planted to cassava increased greatly in Uganda (Jones, 1959). Also, this led to development of the regional breeding program for CMD resistance in Amani, Tanzania and the distribution of resistant plant material in the 1950s. In Uganda this plant material was released as the Bukalasa series in the early 1960s (Otim-Nape et al., 1994). In this program, *M. glaziovii*, a wild cassava species, was used as a donor of genes conferring resistance to CMD. Open pollinated seed of progenies from crosses that had been made between *M. glaziovii* and different cassava varieties were sent to many parts of Africa notably Uganda (Jameson, 1964), Kenya, Nigeria (Jennings, 1994), and Ghana (Doku, 1969). However, the first breeding activities took place just before the first outbreak of CMD and in 1927 six varieties were introduced to Uganda from the West Indies (Jameson, 1964). Another CMD outbreak that occurred in 1989 has led to development and release of the new varieties in the NASE series in the 1990s. The recent CMD incidence differed significantly between agroecologies with locations in northern parts of Uganda being more affected than those in the southern parts

(Otim-Nape et al., 1997; Figure 1). Therefore the impact of governmental and non-governmental mitigation programs to distribute CMD-free or CMD resistant planting material differed accordingly. Studies done on varietal diversity on the basis of the morphology of the cassava varieties showed that there were also marked changes in the predominating varieties being grown, as susceptible ones

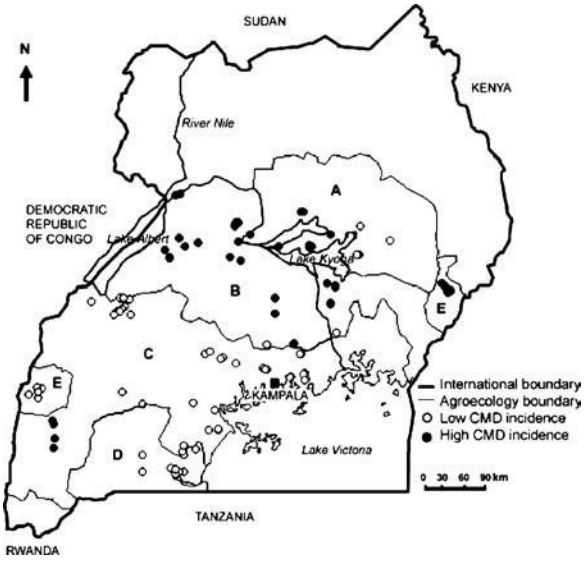

Figure 1. Map of Uganda showing the collection sites with high and low CMD incidence, and the agroecological zones: A: northern agroecosytem, B: banana – millet – cotton agroecosystem, C: banana – coffee agroecosystem, D: pastoral agroecosystem, E: montane agroecosystem.

were being lost and replaced with other CMD-free or CMD tolerant or resistant varieties (Otim-Nape et al., 2001). It is, however, not known how the CMD has affected the genetic diversity and the gene pool of cassava in Uganda. We have therefore studied the genetic diversity in different agroecological zones and areas with high and low incidence of CMD using 35 simple sequence-repeat (SSR) markers. In addition, because of the two independent introductions of cassava in Africa one might to expect loss of genetic diversity as a result of founder effects and genetic differentiation between East and West Africa. To study the effect of the introgression history on the gene pool of cassava the genetic differentiation between Latin America, East and West Africa was studied. The neighboring countries Uganda and Tanzania was chosen from East Africa, and the neighboring countries Ghana and Nigeria from West
Africa.

Materials and methods

Plant materials

The cassava varieties in Uganda were collected in September through to December 2002, from five agroecological zones (Figure 1). These zones lie between 2°12′N and 0°44′ S, and 29 °56′ E and 34°21′ E, and on altitudes from 664 to 1484m above sea. Locations were visited as in Otim-Nape et al. (1998) except the northwestern region of Uganda. These locations were classified as high and low CMD incidence based on Otim-Nape et al. (1998), where a location of high CMD incidence had greater than 50% occurrence of the disease (Figure 1). Based on morphological characteristics plants were collected with assistance from the cassava-breeding program and labeled by the name as given by the farmer. Varieties from the recent breeding work of the 90s, the NASE series, were avoided. These varieties did not exist at the time of the last CMD epidemic in 1989, and therefore would not reflect on the impact of CMD. A total of 245 plants were sampled with 17 from the northern agroecology, 16 from the montane, 63 from the banana–millet– cotton agroecology, 123 from the banana–coffee agroecology and 26 from the pastoral agroecology. The cassava was planted in a screen house in Namulonge Agricultural and Animal Research Institute (NAARI) from which leaves for DNA extraction were obtained. A summary of the plant materials and their source can be viewed at http://www.ciat.cgiar.org/molcas/estudios.jsp? Code=6 & pais=Uganda.

Twenty-two Nigerian accessions were included from the international collection at the International Institute of Tropical Agriculture (IITA) with 10 being improved genotypes from the Institute's cassava breeding program or the 1950s breeding efforts at the Moor Plantation Experiment Station, Ibadan, Nigeria. Based on Fregene et al. (2003) 20 Tanzanian accessions and 38 Latin American holdings from the core collection at CIAT; 6 from Colombia, 3 from Brazil, 3 from Peru, 2 from Mexico, 2 from Venezuela, 2 from Argentina and 20 from Guatemala were selected. In addition 20 accessions from Ghana (courtesy of Elizabeth Okai; CIAT, 2003) were included. In order to cover as much of the genetic diversity as possible, these accessions were selected to represent the genetic distribution found by principal component analysis (PCA) based on SSR markers in each study.

DNA analysis

DNA was isolated from young leaf tissue by CTAB method (Doyle & Doyle, 1987). A subset of 35 SSR markers with broad genome coverage and high polymorphism information content (PIC) was selected from 67 markers from an earlier study (Fregene et al., 2003).
To obtain a maximum amount of information on allelic diversity in cassava 30 unlinked markers are required (Fregene et al., 2003). PCR was carried out using 10 ng of DNA per reaction following Mba et al. (2001). The PCR product was denatured and electrophoresed on 6% polyacrylamide gels using Bio-Rad sequencing apparatus (Bio-Rad Inc., USA) and visualized by silver staining according to the Promega manufacturer's guide. Allele sizes were

then determined based upon an internal gel molecular marker size standard using both manual scoring and the computer software "quantity One" (Bio-Rad Inc.). In addition, a few plants with known genotypes were used as controls on each gel. The genotype data was exported to Microsoft Excel (Microsoft Inc) for further formatting as input files for statistical analysis.

Statistical analysis

Genetic diversity parameters within and among populations (agroecological zones or areas with high and low CMD incidence) were estimated with data from the 35 SSR marker loci using the GEN-SURVEY program (Vekemans & Lefe`bvre, 1997). The average expected heterozygosity (H_e) and the distribution of genetic diversity within and among populations were calculated according to Neil (1978).For all loci and populations the total diversity estimate (H_t =total heterozygosity in the entire dataset) was partitioned into within-population diversity (H_s = heterozygosity within populations averaged over the entire dataset) and between-population diversity (D_{st}) estimates, where $H_t = H_s + D_{st}$. Gene diversity between populations was expressed relative to total population diversity as $G_{st} = D_{st}/H_t$. Standard deviations sampled by jackknifing (200 replications) and 95% confidence intervals sampled by bootstrapping (1000 replications; Quenoille, 1956; Efron, 1982) were estimated over loci for the above parameters.

Genetic differentiation between countries was analyzed using Wright's *F*-statistics (1965) and pairwise calculations of F_{st} overall loci between pairs of country variety groups were estimated using FSTAT 2.9 (Goudet, 1995). The pairwise F_{st} estimates was used to construct a dendrogram on the basis of the unweighted paired group method with the arithmetic mean (UPGMA) using NTSYS-PC (Rohlf, 1993). In addition, pairwise genetic distances between individual cassava plants were calculated from the allele size-data based on the 1-proportion of shared alleles (PSA) using the computer program "microsat" available at http://hpgl.stanford.edu/projects/microsat/microsat.htm. The distance matrix was analyzed by principal component analysis (PCA) using the JMP program to deduce multivariate relationships among the cassava genotypes. Several measures of genetic distance have been developed for SSR markers on the basis of the stepwise mutation model (SSM; Kimura and Crow,1964).SSM assumes that alleles mutate back and forth by adding or subtracting a repeat motif, so that the same allelic states are formed repeatedly over time. An alternative model is the infinite allele model (IAM; Ohta & Kimura, 1973), which assumes that each mutation creates a new allele in the population. Because of the relatively short evolutionary divergence times for the cassava varieties, and thus a smaller number of mutations are expected, we have based the estimates of genetic distance on the IAM. In addition, the difference in number of base pairs between alleles within a SSR locus observed in this study is variable and does not follow any pattern. This may indicate complex patterns of mutation and that the SSR variability in cassava may not fit the stepwise mutation model (SSM; Kimura & Crow, 1964). In fact, a number of studies have shown that many SSR marker loci do not evolve according to the SSM (e.g. Valdes et al., 1993; Matsuoka et al., 2002).

Results

Number of alleles in SSR loci

The observed number of alleles at each locus in the whole dataset (Africa and Latin America) was relatively high and ranged from 2 at SSRY102 and SSRY132, to 12 at SSRY19 (Table 1). Six alleles or more were found in 63% of the 35 studied marker loci and the average number of alleles per locus was 6.1. *Allelic distribution in Uganda*

Due to some missing genotype data the number of plants scored per marker locus varied. The average number of plants scored per locus is given in Tables 1 and 2. A total of 183 alleles were found in the Ugandan collection (Table1).Ofthe183 alleles found 38 alleles occurred in less than 1% within Uganda. These alleles are defined as rare. The number of rare alleles found was highest in the agroecological zone of the banana– coffee agroecosystem (Table 2), which had the largest sample size. The lower number of rare alleles found in the other agroecological zones is likely to be influenced by a smaller sample size. The number of non-rare alleles (occurred in more than 1%) was about the same in the agroecological zones with the exception of the northern agroecosystem. None of the non-rare alleles was unique to any of the five agroecological zones, while some rare alleles were unique to each of the zones: 14 in the banana–coffee agroecosystem, 3 in the banana–millet agroecosystem, 2 in the montane agroecosystem, 2 in the northern agroecosystem and 1 in the pastoral agroecosystem.

When the collection sites were classified with respect to high (>50%) or low (≤50%) CMD incidence according to Otim-Nape et al. (2001) the sample sizes were more equal and the number of scored plants was 89 and 106 from the high and low CMD incidence groups, respectively. Thirty-three rare alleles were found in the areas with low CMD incidence compared to 13 rare alleles in the high CMD incidence areas (Table 2), and a significant difference in the number of rare alleles was found using a Chi-square test ($P < 0.05$). Eight of the rare alleles were common to both the low and high CMD incidence classification groups. The number of non-rare alleles was about the same in the two groups.

Genetic diversity in Uganda

Of the 35 marker loci studied there were on an average 93.7% polymorphic loci across all agroecological zones (Table 2), using the criterion that the frequency of the most common allele does not exceed 0.98. The observed heterozygosities were high in all agroecological zones and ranged from 0.536 to 0.594, at an average of 0.559 (Table 2). The expected heterozygosity (corrected for small sample sizes; Nei, 1978) in the agroecological zones ranged from 0.487 to 0.594, at an average of 0.544. This implies that the probability that two randomly selected alleles in Uganda are different is more than half. The observed and expected heterozygosity was about the same in the high and low CMD incidence groups and close to 0.5. Compared to previous gene diversity studies on cassava where allozyme markers

were used (Lefevre & Charrier, 1993; Resende et al., 2000), the observed proportion of heterozygotes is about three-fold greater with SSR markers. Similarly, in domesticated sunflower accessions the observed proportion of heterozygotes was found to be two- to four-fold greater for SSR marker loci than for allozymes (Tang & Knapp, 2003), and in sorghum landraces 20-fold greater for SSR marker loci than for allozymes (Dje` et al., 1999). The higher

Table 1. Number of alleles in each SSR locus for cassava varieties in the different countries

	Population					
Locus	Uganda (195)[1]	Tanzania (19)[1]	Ghana (19)[1]	Nigeria (20)[1]	Latin America (34)[1]	Number of alleles
SSRY4	5	5	4	4	6	6
SSRY9	4	4	4	5	5	5
SSRY12	6	4	4	4	4	6
SSRY19	8	7	7	6	10	12
SSRY20	9	6	7	6	7	10
SSRY21	6	5	4	4	7	7
SSRY34	4	3	3	2	3	5
SSRY38	4	2	2	3	4	5
SSRY51	5	5	4	4	6	6
SSRY59	5	4	3	2	6	7
SSRY63	4	2	3	3	3	4
SSRY64	6	4	4	4	5	7
SSRY69	8	7	6	5	7	8
SSRY82	9	7	8	7	8	9
SSRY100	6	5	6	4	7	7
SSRY102	2	2	2	2	2	2
SSRY103	5	5	5	4	5	5
SSRY105	5	4	3	3	5	6
SSRY106	5	5	5	4	5	5

[1] Average number of scored plants per locus.

SSRY108	6	3	4	4	6	8
SSRY110	6	4	4	4	5	6
SSRY132	1	1	1	1	2	2
SSRY135	3	3	3	3	3	3
SSRY147	4	2	2	2	3	4
SSRY148	3	3	3	2	5	5
SSRY151	8	6	9	7	8	9
SSRY155	5	2	5	3	5	6
SSRY161	5	4	4	5	4	6
SSRY164	6	3	5	5	6	8
SSRY169	4	2	2	3	4	5
SSRY171	4	2	2	2	5	7
SSRY177	6	4	6	5	6	6
SSRY179	7	6	6	6	6	7
SSRY180	5	3	3	3	6	7
SSRY181	4	3	4	4	4	4
Total alleles	183	137	147	135	183	

Proportion of observed heterozygotes found with SSR marker loci is to be expected because of their high polymorphism.

The total heterozygosity over all loci was high in the agroecological zones (H_t = 0.564, Table 2). Only a very small fraction of this was due to differentiation among zones (G_{st} = 0.035), while most of the diversity was found within zones (H_s = 0.545). The same pattern was found for the CMD incidence groups (G_{st} = 0.004).

Table 2. Genetic diversity of cassava varieties in different agroecological zones and areas with high or low CMD incidence in Uganda

Agro-ecology[1]	Mean no. of scored plants/locus	No. of loci	No. of rare alleles[2]	No. of non-rare alleles[3]	Percent of poly-morphic loci	Mean no. of alleles/ locus	Mean no. of alleles/poly-morphic locus	H_o4	H_e5	$He-p$[6]
A	13	35	2	113	88.6	3.3	3.3	0.550	0.467	0.487
B	49	35	6	137	94.3	4.3	4.5	0.536	0.537	0.542
C	99	35	29	145	94.3	4.9	5.1	0.552	0.534	0.537
D	21	35	7	137	94.3	4.1	4.2	0.565	0.547	0.561

E	13	35	8	133	97.1	4.1	4.2	0.594	0.571	0.594
Mean					93.7	4.1	4.3	0.559	0.532	0.544
SD[2]					3.1	0.6	0.6	0.022	0.038	0.039
CMD incidence										
High	89	35	13	139	94.3	4.4	4.6	0.545	0.535	0.538
Low	106	35	33	145	94.3	5.0	5.2	0.557	0.548	0.550
Mean					94.3	4.7	4.9	0.551	0.541	0.544
SD[7]					0.0	0.4	0.4	0.009	0.009	0.009

Agroecology	H_t[34]	H_s	D_{st}	G_{st}
Mean	0.564	0.545	0.019	0.035
SD[7]	0.192	0.188	0.012	0.020
95% CI[9]	0.499	0.482	0.015	0.029
95% CI[9]	0.626	0.606	0.023	0.042
CMD incidence				
Mean	0.562	0.560	0.002	0.004
SD[7]	0.197	0.197	0.004	0.007
95% CI[9]	0.499	0.497	0.001	0.001
95% CI[9]	0.623	0.621	0.003	0.006

[1]A:northernagroecosytem,B:banana–millet–cottonagroecosystem,C:banana–coffeeagroecosystem,D:pastoralagroecosystem,E:montane agroecosystem.

[2]An allele occurring in less than 1% in the Ugandan population.

[3]An allele occurring in more than 1% in the Ugandan population.

[2] Standard deviation estimated by jackknifing over loci (200 replications).

[3] H_t = total heterozygosity in the dataset, H_s = heterozygosity within populations averaged over the entire dataset, D_{st} = average gene diversity between populations, G_{st} = coefficient of genetic differentiation. These parameters are given over loci and over populations.

[4] % confidence interval for the mean estimated by bootstrapping over loci.

[4]Average observed heterozygosity within populations.

[5]Average expected heterozygosity within populations.

[6]Average expected heterozygosity within populations corrected for small sample sizes (Nei, 1978).

Allelic distribution in Latin America and the other African countries

Despite the small sample size Latin America had the same number of alleles (183) as Uganda (Table 1). Nineteen of the 38 rare alleles in Uganda occurred in Latin America in frequencies of 2.5–18.8%, 14 in Ghana (2.5–18.8%), 10 in Tanzania (2.5–13.2%) and 8 in Nigeria (2.4–31.8%). Five non-rare alleles found in Uganda occurring in frequencies from 1.1 to 16.7% were not found in any of the other African countries but in Latin America in a range of 4.1 to 16.2%. There were 24 unique alleles to Latin America in frequencies of 1.4 to 28%. Twenty-seven alleles were unique to Africa of which 13 were rare alleles and 3 non-rare alleles only found in Uganda. One of the 27 alleles was unique to Tanzania while 6 occurred in two or three African countries. Four of the 27 alleles were common to all of the African countries.

*Table3.*PairwiseestimatorofF_{st} between pairs of country groupings of cassava varieties

Population	Uganda	Tanzania	Nigeria	Ghana	Latin America
Uganda	–				
Tanzania	0.051	–			
Nigeria	0.113	0.104	–		
Ghana	0.040	0.050	0.049	–	
Latin America	0.097	0.059	0.057	0.054	–

Genetic differentiation between Latin America and Africa

Pairwise calculations of F_{st} overall loci between country variety groups are shown in Table 3. Small differentiation was found between the varieties in Uganda and Tanzania in East Africa and Ghana in West Africa, and between varieties in Ghana and Nigeria in West Africa (average F_{st} = 0.048). Interestingly, larger differentiation was found between Ugandan and Nigerian cassava varieties, and between Tanzanian and Nigerian varieties, than between varieties in Latin America and the African countries (average F_{st} = 0.067). Among the African countries Uganda showed the largest differentiation with Latin America.

The genetic relationship between the country variety groups is visualized in a UPGMA dendrogram on the basis of the pairwise F_{st} estimates (Figure 2). The varieties in Uganda and Tanzania in East Africa and Ghana in West Africa clustered closely together.

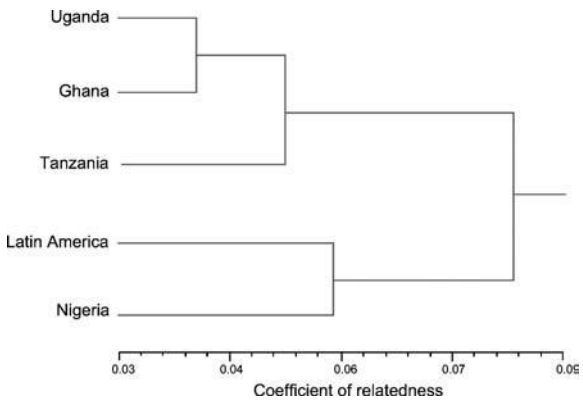

Figure 2. A UPGMA dendrogram illustrating the genetic relationship among cassava variety groups as revealed by pairwise F_{st} estimates between countries.

Interestingly, the Nigerian varieties grouped closer to the Latin American varieties than to the varieties in the other African countries. A principal component analysis was also performed on the genetic distance estimator 1-proportion of shared alleles (1-PSA) and a score plot of the first two principal components (explaining 39 and 10% of the variance, respectively, data not shown) gave essentially the same pattern of relationship among the country variety groups as the dendrogram.

Discussion

Cassava mosaic virus disease (CMD) is one of the most serious cassava diseases in Africa. It is unknown in Latin America and believed to be of African origin. The virus is transmitted by whiteflies and infects all cassava varieties in epidemics, although susceptibility varies greatly among varieties (Fauquet & Fargette, 1990). Due to the severe outbreaks of CMD in Uganda followed by large-scale multiplication schemes of some cassava varieties and release of CMD resistant varieties, one might expect to find a great loss of local varieties and thereby a reduction of genetic diversity in the Ugandan cassava gene pool. In addition, the introductions of cassava from Latin America to West Africa in the 17th century and to East Africa in the 18th century may have resulted in founder effects with loss of genetic diversity and genetic differentiation between the East and West African countries. Being mainly a vegetatively propagated crop a spread of a few vigorous and well-adapted local varieties among farmers would be facilitated and further reduce the genetic diversity of cassava. To study the impact of the history on the genetic diversity of the Ugandan cassava we have used population genetic models. These models are based on a number of assumptions, i.e. sexual reproduction and random mating. This has implications on the appropriateness of using population genetic models on cultivated plants, in particular vegetatively propagated crops, whose diversity is mainly influenced by humans. However, several genetic diversity studies on crops have based their analyses on population genetic models (e.g. Dubreuil & Charcosset, 1998; Dje` et al.,

2000; Fregene et al., 2003). In cassava local varieties are likely to be a result of spontaneous recombination because of incorporation of sexually produced plants by the Ugandan farmers. Thus, we found it appropriate to use population genetic models to analyze the genetic diversity in cassava.

Genetic diversity in Uganda

In spite of the severe outbreaks of CMD and the largescale multiplication of cassava in Uganda a surprisingly high genetic diversity (H_t = 0.567) of cassava was detected. Most of the genetic diversity was found within populations, while very little diversity was found among agroecological zones and between the high and low CMD incidence groups. This indicates a high gene flow among populations. This gene flow is likely to be mediated by an efficient exchange of varieties among agroecological zones as well as the extensive distribution of varieties in the Amani breeding program from the 1920s to the 1960s throughout Uganda.

The high genetic diversity suggests a high number of local varieties and that the farmers have had a major impact on the composition of the cassava gene pool, while the CMD outbreaks and large-scale multiplication schemes have had a limited effect. Even though varieties were wiped out in certain areas (Otim-Nape et al., 1998) many survived and continue to be used by farmers. In addition, farmers have adopted new genotypes by including volunteer seedlings in their fields and by exchange of new genotypes developed by other farmers. In the Amazonian region of Peru there is a turnover of cassava varieties every 15 years due to pest and disease accumulation in the propagules (Salick, 2001). This dynamic diversity may very well be adaptive and indicative of distinct patterns in the local breeding of cassava and other vegetatively propagated crops.

Although the genetic diversity is high there is a loss of rare alleles in the areas with high CMD incidence. This shows that the CMD has had an effect on the cassava gene pool even though it could not be detected by using the heterozygosity or genetic diversity estimates. When the population size is drastically reduced, as in the high CMD incidence areas, rare alleles are most susceptible to loss (Nei, 1975).

Genetic differentiation between Latin America and Africa

Twenty-four alleles were unique to Latin America and not found in any of the African countries. This may indicate a loss of genetic diversity due to founder effects when cassava was introduced into Africa. However, 27 unique alleles were found in Africa. Some of these alleles might have been introduced when the wild cassava species *M. glaziovii* was used as a donor of genes conferring resistance to CMD in the Amani breeding program, Tanzania, in the 30s and 40s. In spite of the unique alleles the African and Latin American varieties seem not to be separated in two distinct genepools since the Nigerian varieties showed closer genetic similarity with the Latin American varieties than with the varieties in Uganda and Tanzania (Table 3 and Figure 2). In addition, we found small to moderate genetic differentiation between Latin American and African varieties (average F_{st} = 0.067).

Within Africa genetic similarities were found between the varieties in Uganda and Tanzania in East Africa and Ghana in West Africa. Because of the similarity between varieties in these countries there is no evidence for a differentiation of the African cassava gene pool into a western and an eastern genetic lineage. Thus, the two independent introduction events in East and West Africa are not reflected in the African cassava gene pool. However, a larger genetic difference was found between the varieties in Nigeria in West Africa and the varieties in Uganda and Tanzania. A possible difference in the genetic constitution of the introduced cassava material into East and West Africa may have been diminished by movement of germplasm between countries. In the Tanzania Amani breeding program open pollinated seed of progenies from crosses that had been made between *M. glaziovii* and different cassava varieties were sent to many parts of Africa notably Uganda (Jameson, 1964), Kenya, Nigeria (Jennings, 1994) and Ghana (Doku, 1969). Uganda and Ghana may have got more closely related germplasm than what was received in Nigeria. Nigerian farmers tend to prefer bitter varieties that require processing as opposed to Ghana and Uganda where sweet, fresh, and boil varieties tend to predominate (Nweke & Bokanga, 1994; Westby, 2002). Investigation into the pedigree records and movement of germplasm from the Amani breeding programme may bring more clarity on this matter.

In conclusion, one of the important findings in this study is that in spite of severe outbreaks of CMD followed by large-scale multiplication and introduction of bred resistant varieties in Uganda the genetic diversity is high. The mechanism that rapidly compensates the loss of variability is the active involvement of the Ugandan farmer in continuously testing and adopting new genotypes that will serve their diverse needs. However, even though the genetic diversity is high there is a loss of rare alleles in the areas with high CMD incidence. Continuous CMD epidemics and extensive introductions of bred varieties throughout the country with the replacement of local varieties will most likely further reduce the genetic diversity. For food security a large number of local varieties comprising extensive genetic diversity are important to maintain so that coming generations can cope with unpredictable environmental changes and human needs. Extensive breeding programs should therefore be preceded by investigations of the organization of genetic diversity within cassava that would benefit germplasm conservation.

Acknowledgement

We are grateful to Wilson Castelblanco for lab and statistical assistance, and Jacobs Omara and Halima Nanyonjo for their assistance with field collection. This work was supported by grants from the Swedish International Development Cooperation Agency (Sida/SAREC) and from the International Program in the Chemical Sciences (IPICS).

References

CIAT, 2003. Annual Report 2003. Project IP-3: Improved Cassava for the developing world, CIAT, Cali, Colombia.

Dje`, Y., D. Forcioli, M. Ater, Lef`ebvre & X. Vekemans, 1999. Assessing population genetic structure of sorghum landraces from North-western Morocco using allozyme and microsatellite markers. Theor Appl Genet 99: 157–163.

Dje`, Y., M. Heuertz, C. Lef`ebvre & X. Vekemans, 2000. Assessment of genetic diversity within and among germplasm accessions in cultivated sorghum using microsatellite markers. Theor Appl Genet 100: 918–925.

Doku,G.V.,1969.Cassava in Ghana. Ghana University Press, Accra.

Doyle, J.J. & J.L. Doyle, 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19: 11–15.

Dubreuil, P. & A. Charcosset, 1998. Genetic diversity within and among maize populations: A comparison between isozyme and nuclear RFLP loci. Theor Appl Genet 96: 577–587.

Efron, B., 1982. The Jackknife, the Bootstrap and Other Resampling Plans. CBMS-NSF Regional Conference Series in Applied Mathematics, Monograph 38, SIAM, Philadelphia.

Elias, M., L. Penet, P. Vindry, D. MacKey, O. Panaud & T. Robert, 2001. Unmanaged sexual reproduction and the dynamics of genetic diversity of a vegetatively propagated crop plant, cassava (*Manihot esculenta* Crantz), in a traditional farming system. Mol Ecol 10: 1895–1907.

Fauquet, C. & D. Fargette, 1990. African cassava mosaic-virusetiology, epidemiology and control. Plant Dis 74: 404–411.

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

Goudet, J., 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. J Hered 86: 485–486.

Jameson, J.D., 1964. Cassava mosaic disease in Uganda. East Afr Agri J 29: 208–213.

Jennings, D.L., 1994. Breeding for resistance to African cassava mosaic geminivirus in East Africa. Trop Sci 34: 110–122.

Jones, W., 1959. Manioc in Africa. Stanford University Press, Stanford.

Kimura, M. & J.T. Crow, 1964. The number of alleles that can be maintained in a finite population. Genetics 49: 725–738. Langlands, B.W., 1966. Cassava in Uganda 1860–1920. Uganda J 30: 211–218.

Lefevre, F. & A. Charrier, 1993. Isozyme diversity within African *Manihot* germplasm. Euphytica 66: 73–80.

Martin, E.F., 1928. Report of the mycologist. In: Annual Report of the Department of Agriculture. Government Printer, Entebbe, Uganda, pp. 31.

Matsuoka, Y., S.E. Mitchell, S. Kresovich, M. Goodman & J. Doebley, 2002. Microsatellites in *Zea* – variability, patterns of mutations, and use for evolutionary studies. Theor Appl Genet 104: 436–450.

Mba, R.E.C., P. Stephenson, K. Edwards, S. Melzer, J. Mkumbira, U. Gullberg,K.Apel,M.Gale,J.Tohme&M.A.Fregene,2001.Simple Sequence Repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: Towards an SSR-based molecular genetic map of cassava. Theor Appl Genet 102: 21–31.

Nei, M., 1975. Molecular Population Genetics and Evolution. Amsterdam: North Holland and American Elsevier Pub.

Nei, M., 1978. Estimation of average heterozygosity and genetic distances from a small number of individuals. Genetics 89: 583– 590.

Nweke, F.I. & M. Bokanga, 1994. Importance of cassava processing for production in sub-Saharan Africa. Acta Hort 375: 401–412.

Ohta, T. & M. Kimura, 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet Res 22: 201–204.

Olsen, K.M. & B.A. Schaal, 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. Proc Natl Acad Sci USA 96: 5586–5591.

Olsen, K. & B. Schaal, 2001. Microsatellite variation in cassava *Manihot esculenta*, Euphorbiaceae and its wild relatives: Evidence for a southern Amazonian origin of domestication. Am J Bot 88: 131–142.

Otim-Nape, G.W., T. Alicai & M. Thresh, 2001. Changes in the incidence and severity of Cassava mosaic virus disease, varietal diversity and cassava production in Uganda. Ann Appl Biol 138: 313–327.

Otim-Nape, G.W., A. Bua & Y. Baguma, 1994. Accelerating the transfer of improved production technologies: Controlling African cassava mosaic virus disease epidemics in Uganda. Afr Crop Sci J 2: 479–485.

Otim-Nape, G.W., A. Bua, J.M. Thresh, Y. Baguma, S. Ogwal, G.N. Semakula, G. Acola, B. Byabakama & A. Martin. 1997. Cassava mosaic virus disease in Uganda: The current pandemic and approaches to control. NARO/DFID, Chatham, UK, Natural Resources Institute.

Otim-Nape, G.W., J.M. Thresh, A. Bua, Y. Baguma & M.W. Shaw, 1998. The incidence and severity of cassava mosaic virus disease in Uganda: 1990–92. Trop Sci 38: 25–37.

Pita, J.S., V.N. Fondong, A. Sangare, G.W. Otim-Nape, S. Ogwal &

C.M. Fauquet, 2001. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. J Gen Virol 82: 655– 665.

Quenoille, M., 1956. Notes on bias in estimation. Biometrika 43: 253–260.

Resende, A.G., P.S.V. Filho & M.P.S. Machado, 2000. Isozyme diversity in cassava cultivars (*Manihot esculenta* Crantz). Biochem Genet 38: 203–216.

Rohlf, F.J., 1993. NTSYS-PC numerical taxonomy and multivariate analysis system. Version 1.8 Exeter Publ., Setauket, NY.

Salick, J., 2001. Diversity, Evolution and Domestication within *Manihot*. In: Abstracts of the Fifth International Scientific Meeting of the Cassava Biotechnology Network (CBN), Danforth Plant Scientific center, St Louis, Missouri, USA, 4–9 November 2001. http://www.danforthcenter.org/iltab/cassavanet/ cbnv/abstracts/session1/ps1-03.htm.

Tang, S. & S.J. Knapp, 2003. Microsatellites uncover extraordinary diversity in native American land races and wild populations of cultivated sunflower. Theor Appl Genet 106: 990–1003.

Valdes, A.M., M. Slatkin & N.B. Freimer, 1993. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. Genetics 133: 737–749.

Vekemans, X. & C. Lefe`bvre, 1997. On the evolution of heavymetal tolerant populations in *Armeria maritima*: Evidence from allozyme variation and reproducive barriers. J Evol Biol 10: 175– 191.

Westby, A., 2002. Cassava utilization, storage and small-scale processing. In: R.J. Hillocks, J.M. Thresh & A.C. Belloti. (Eds.), Cassava: Biology, Production and Utilization, pp. 281–300. CAB International, Wallingford, Oxon, UK.

Wright, S., 1965. The interpretation of population structure by *F*statistics with special regard to systems of mating. Evolution 19:
395–420.