University of Pretoria etd – Ndunguru, J (2005)

CHAPTER 3

Restriction and sequence analysis of PCR-amplified viral DNAs suggests the existence of different cassava mosaic geminiviruses associated with cassava mosaic disease in Tanzania

Joseph Ndunguru, [1,4] James P. Legg, [3] Theresa A. S. Aveling, [4] Graham Thompson [5] and Claude M. Fauquet [2]

[1]Plant Protection Division, P.O. Box 1484, Mwanza, Tanzania

[2]International Laboratory for Tropical Agricultural Biotechnology, Danforth Plant Science Center, St. Louis, MO 63132 USA

[3]International Institute of Tropical Agriculture-Eastern and Southern Africa Regional Center and Natural Resource Institute (UK), Box 7878, Kampala, Uganda

[4]Department of Microbiology and Plant Pathology, University of Pretoria, Pretoria 0002, South Africa

[5]ARC-Institute for Industrial Crops, Private Bag X82075, Rustenburg 0300, South Africa

---

The molecular variability and geographical distribution of cassava mosaic geminiviruses (CMGs) occurring in Tanzania was investigated. Infected cassava leaves were collected from cassava plantings in different geographical areas of the country and PCR and RFLP were used for molecular characterization of the CMGs. RFLP analysis using *Eco*RV and *Mlu*I endonucleases of the CMGs revealed more molecular variability in the genome of *East African cassava mosaic virus* (EACMV) than in *African cassava mosaic virus* (ACMV) and partial sequence comparison of the replication associated gene (AC1) strongly reinforced this observation. EACMV-[Tanzania] (EACMV-[TZ]) and eleven other EACMV-like virus types, designated here as EACMV-[TZ1]to-[TZ11] were

identified. These viruses were associated with distinct symptoms on different cassava cultivars, and had non-overlapping and overlapping geographical distributions with most of them occurring in the coastal regions of the country. Of the EACMV-like virus types, EACMV-[TZ] was the most widespread followed by EACMV-[TZ1] and [EACMV-[TZ2]. EACMV-[TZ4], EACMV-[TZ5] and EACMV-[TZ8] were associated with the most severe mosaic symptoms. Another as yet undescribed CMG provisionally designated EACMV-[TZ Mara] was identified in the Mara region in northwestern Tanzania. An ACMV-like virus designated ACMV-[TZ], was also detected in the Mara region. The EACMV-[UG2] associated with the pandemic of severe cassava mosaic disease (CMD) in Uganda and elsewhere has expanded its range into Tanzania, covering most of the Lake Victoria region. Coinfection frequently involved ACMV and EACMV-[UG2] and occurred mainly at the advancing front of the CMD pandemic and plants infected displayed more severe symptoms than those infected by either of the two viruses alone. ACMV was not found in any of the coastal regions or in the south of the country.

3.1 INTRODUCTION

Cassava mosaic geminiviruses (CMGs) of the genus *Begomovirus*, family *Geminiviridae* have been known since the 1970s to cause cassava mosaic disease (CMD) in Tanzania (Swanson and Harrison, 1994) and were recently reported by Ogbe *et al*. (1997) and Legg (1999). Six CMG species are known to affect cassava in Africa but the two most widely occurring species are:- *African cassava mosaic virus* (ACMV), which occurs in most cassava-producing areas in Africa and *East African cassava mosaic virus* (EACMV), which is mainly found in coastal East Africa, although recent reports of its presence have been made in some west African states (Fondong *et al*., 2000; Ogbe *et al*., 2003) and South Africa (Berry and Rey, 2001). More recently, four new CMG species have been described: *South African cassava mosaic virus* (SACMV) (Rey and Thompson, 1998; Berrie *et al*., 2001), *East African cassava mosaic Cameroon virus* (EACMCV) (Fondong *et al*., 2000), *East African cassava mosaic Malawi virus* (EACMMV) isolates MH and MK and *East African cassava mosaic Zanzibar virus* (EACMZV) (Maruthi *et al*., 2002). These CMGs possess two DNA molecules, DNA-A and DNA-B. DNA-A encodes all the viral proteins

University of Pretoria etd – Ndunguru, J (2005)

necessary for replication and encapsidation of both components (Rogers *et al*., 1986; Townsend *et al*., 1986; Sunter *et al*., 1987), while DNA-B encodes for two proteins required for efficient systemic spread of the virus throughout the plant and symptom development (Hanley-Bowdoin *et al*., 1999). Based on serology, only EACMV was detected in the coastal regions of Tanzania and ACMV in the inland parts of the country (Ogbe *et al*., 1996). In northwestern Tanzania, both ACMV and EACMV infect cassava and currently the spread of the pandemic of severe CMD in this area has been associated with the EACMV-[UG2] strain (recombinant between ACMV and EACMV also named the Uganda variant) (Legg, 1999; Pita *et al*., 2001a). EACMV-[UG2] has spread to other countries of Africa including southern Sudan (Harrison *et al*., 1997), Rwanda (Legg *et al*., 2001), Democratic Republic of Congo and the Republic of Congo (Neuenschwander *et al*., 2001), Burundi (Bigirimana *et al*., 2004) and Gabon (Legg *et al*., 2004). A survey of cassava virus diseases in Tanzania recorded CMD throughout the country at low to moderate incidences in fifteen districts, with higher incidence in the coastal plain than at higher altitudes in the interior (Legg and Raya, 1998). In 1998, severe yellow mosaic symptoms were observed on cassava in the Ruvuma region of southern Tanzania, particularly in the Mbinga district near Lake Nyasa (J. Ndunguru, unpublished data), but the identity of the viruses associated with the disease was not established. This paper provides molecular evidence for the diversity of CMGs in Tanzania based on extensive sample collection from all of the main cassava-growing areas. The extent of the spread of the pandemic of severe CMD associated with EACMV-[UG2] in the country was also investigated.

3.2 MATERIALS AND METHODS

3.2.1 Virus collection. Cassava was surveyed in three regions in the Lake Victoria basin in the northwest (Mwanza, Kagera and Mara), five regions in the coastal zones (Tanga, Dar-es-Salaam, Coast, Lindi and Mtwara) and one region (Ruvuma) in the south, all of which comprise the major cassava-growing areas in Tanzania. Surveying was done from August to September 2002 in each of the 31 districts covered. A total of 205 cassava fields of three

to six months of age were selected at regular intervals along the roads within each region. Within each field, two symptomatic plants (one expressing mild and the other severe symptoms) were sampled by removing the youngest symptomatic leaf and placing it in a 1.5 ml eppendorf tube, which was then stored in a cool box. In fields where all the diseased plants showed either severe or mild symptoms, two randomly selected plants were sampled. Symptoms were described and photographs taken for further reference. The severity of CMD was assessed using a 0-5 (0 = symptomless and 5 = severe leaf size reduction) scoring scale. In each field, geographical location, longitude and latitude were recorded using a GPS 4300 device (Magellan, Taiwan) and the coordinates were used for mapping the virus types/species in the surveyed areas. Farmers provided information on cassava variety, name of the location and crop age but could not name the cassava varieties sampled in 52 fields.

3.2.2 DNA extraction. Total DNA was extracted during the survey from young cassava leaves that showed CMD symptoms. The method used was essentially as described by Dellaporta *et al*. (1983), with some modifications. Symptomatic young leaf samples (0.1 to 0.5 g) were ground with Kontes pestles in a 1.5 ml microfuge tube containing 500 µl of extraction buffer (100 mM Trizma base, 8.5 mM EDTA, 500 mM NaCl, 10 mM β-mercaptoethanol, pH 8.0). To each tube, 33 µl of 20% sodium lauryl sulphate solution was added, mixed and the contents were incubated at 65°C in a water bath for 10 min. Then 160 µl of 5 M potassium acetate solution was added to each tube, mixed and the tubes were stored at -20°C for 10 min and then centrifuged at 13000 g for 10 min. The supernatant (450 µl) was transferred into a new 1.5 ml tube to which 450 µl of cold isopropanol was added and the centrifugation repeated to precipitate the DNA. After washing with 70% ethanol, the DNA was air dried for 30 min, re-suspended in 300 µl of distilled water and stored at 4°C.

3.2.3 PCR analysis. In order to amplify DNA-A sequences of CMGs, universal oligonucleotide primers UNI/F (5'KSGGGTCGACGTCATCAATGACGTTRTAC3') and UNI/R (5'AARGAATTCATKGGGGCCCARARRGACTGGC 3'), where K = G+T, R = A + G, S = G + C (Invitrogen, Life Science, USA), designed to amplify near full-length DNA-A, were used. PCR was performed with a first cycle of 1 min at 94°C followed by 30

cycles of 1 min at 94°C, 1.5 min at 55°C and 2 min at 72°C and a final extension at 72°C for 10 min. PCR amplification was performed in a 20 µl reaction volume following the manufacturer's protocol. To confirm the presence of EACMV-[UG2] (recombinant for the coat protein) (Zhou *et al*., 1997) in Tanzania, specific primer pairs UV-AL1/F1 (5' TGTCTTCTGGGACTTGTGTG3') and ACMV-CP/R3 (5'TGCCTCCTGATGATTATATGTC 3') were used with PCR conditions as described above.

After completion of PCR, samples (12 µl) were loaded onto a 1.2% agarose gel in TAE buffer and the ethidium bromide stained DNA-bands identified using ultraviolet light.

3.2.4 DNA precipitation and RFLP analysis. PCR–amplified DNA was precipitated with ethanol and digested with 10 U *Mlu*I in 10.2 µl reaction volumes containing SH buffer (10 mM Tris-HCl, 100 mM NaCl, 5 mM $MgCl_2$, 1 mM 2-mercaptoethanol, pH 8.0) and 10 U *Eco*RV in the same reaction volume containing SB buffer (50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 1mM dithioerythritol, pH 7.5) in 0.5 ml tubes. The endonuclease-digested DNA was electrophoresed in a 1.5% agarose gel in TAE buffer and visualized with ultraviolet light after staining in ethidium bromide.

3.2.5 Molecular cloning of viral DNA. The PCR- amplified viral DNA fragments were eluted from the agarose gel using the QIAgen Gel Extraction Kit (Invitrogen, Life Science, USA) as per the manufacturer's protocol. PCR products were cloned into the TOPO plasmid vector (pCR. 2.1TOPO.) using the TA Cloning Kit (Invitrogen, Life Sciences, USA) between *Eco*RI restriction sites and transformed into chemically competent *Escherichia coli* strain DH5αTM-T1R. Clones were confirmed positive for viral inserts using *Eco*RI restriction digestion of the extracted plasmids and PCR primers used for initial amplification.

3.2.6 Sequence determination and analysis. Cloned PCR products were sequenced essentially as described by Pita *et al*. (2001a). Sequence data were assembled with the aid of the DNAStar software and GenBank searches performed with the BLAST programme. The annealing position of the primer UNI/F and UNI/R in the replication associated (Rep)

AC1 gene of DNA-A component of the different CMGs identified in this study and of 16 selected CMGs from GenBank were determined to be between 1824-1834 and 1773-1834 nucleotides, respectively. Multiple sequences were aligned using the CLUSTAL-W option of the Mac Vector 7.2 package (Accelry, San Diego, Inc, USA). Phylogenetic trees were constructed using a neighbour-joining majority rule consensus with 1000 bootstrapped replicates. The phylogenetic trees generated showed virus clusters typical of those given by the International Committee for the Taxonomy of Viruses (Fauquet *et al*., 2003) except that when forward primer sequences were used (Fig. 3.3a), ACMCV-[CM] (AF112352) paired closely to ACMV-[Nig-Ogo] (AJ427910) instead of ACMV Cameroon-[Ivory Coast] (AF249894) because of the AC2 gene sequence inclusion upstream of the 360 nts fragment (Fondong *et al*., 2000). Thus, the 360 nts AC1 nucleotide sequence fragments from the RFLP-categorized virus types obtained as described above were compared with equivalent GenBank nucleotide sequences. The EACMVs selected for comparison with their accession numbers were *East African cassava mosaic virus*-Tanzania (EACMV-[TZ], Z83256); *East African cassava mosaic Zanzibar virus* (EACMZV, AF422174); *East African cassava mosaic Cameroon virus*-Cameroon [Ivory Coast] (EAMCV-CM[CI], AF259896); *East African cassava mosaic Malawi virus* (EACMMV-MW[MH], AJ006459; EACMMV-MW[K], AJ006460); *East African cassava mosaic* virus-Kenya (EACMV-[KE-K2B], AJ006458), *East African cassava mosaic Cameroon virus* (EACMCV-[CM], AF112354) and *East African cassava mosaic virus*-Uganda (EACMV-[UG2], Z83257). ACMVs were *African cassava mosaic virus*-Nigeria (ACMV-[Nig-Ogo], AJ427910); *African cassava mosaic virus*-Kenya (ACMV-[KE], J02057); *African cassava mosaic virus*-Uganda severe (ACMV-[UGSvr], AF126802); *African cassava mosaic virus*-Cameroon (ACMV-[CM], AF112352); *African cassava mosaic virus*-Cameroon/DO2 (ACMV-[CM/D02], AF366902) and *African cassava mosaic virus*-Ivory Coast (ACMV-[IC], AF259894). *Sri Lankan cassava mosaic virus*-Colombo (SLCMV-[Col], AF314737), *Indian cassava mosaic virus*-Maharashtra (ICMV-[Mah], AJ314739) and *South African cassava mosaic virus* (SACMV, AF155807) were chosen as outgroups.

3.3 RESULTS

3.3.1 Cassava sample collection

Samples were obtained from each of the four main cassava-growing regions: southern, coastal, eastern, and Lake Victoria basin in the northwest. The initial plan was to collect samples from plants showing both mild and severe symptoms in each field. However, severe symptoms were observed in most of the sampled fields whilst mild symptoms were only occasionally found in the Mara, Tanga and Lindi regions. In most fields therefore, samples were only available from plants showing moderate to severe disease symptoms.

3.3.2 PCR and RFLP-based molecular variability of CMGs in Tanzania

Using primers UNIF/UNIR, PCR of total DNA extracted from CMD-diseased cassava plants gave amplified fragments of *c*. 2.8 kbp, as expected for CMGs. Three hundred and twenty five of the 510 samples gave PCR products. RFLP analysis of the CMGs revealed greater molecular variability in EACMV genomes than in those of ACMV. CMGs that gave RFLP patterns specific to EACMV when digested with *Eco*RV but a different pattern when cut with *Mlu*I were classified as EACMV-like virus types. Any CMG that gave RFLP patterns that were neither ACMV nor EACMV-like with both *Eco*RV and *Mlu*I was provisionally designated as an EACMV-[TZ] type, with the addition of the location name from which it was sampled. Thus EACMV-[TZ Mara] was detected in samples collected from Mara region. Based on RFLP data, 11 EACMV-like viruses were identified in addition to EACMV-[TZ], and designated EACMV-[TZ1] to -[TZ11] (Fig. 3.1a). ACMV-[KE] and ACMV-[NG] were the main isolates of ACMV detected in this study. *Mlu*I cuts ACMV-[KE] but not ACMV-[NG]. However, another type of ACMV that was different from these two isolates was detected and designated ACMV-Tanzania (ACMV-[TZ]). A total of 325 PCR products obtained from amplified samples were digested with *Eco*RV and *Mlu*I and analysed on agarose gels to detect fragment length polymorphism. EACMVs yielded two fragments of ~ 2199 and ~ 585 bp following digestion with *Eco*RV (Fig. 3.1b).

The EACMV that gave a fragment length pattern similar to those expected for previously sequenced EACMV isolates (Swanson and Harrison 1994) was named EACMV-[TZ]. *Mlu*I digests, used to distinguish the EACMVs, are shown in Fig. 3.1a. EACMV-[TZ], which yielded three different fragments (~ 1212, ~ 1057 and ~ 515 bp) when digested with *Mlu*I*,* was detected in 63 of the 325 samples (20%) (Fig. 3.1a panel a). The total size of the fragments was as expected for EACMV. EACMV-[TZ1] also yielded three fragments with *Mlu*I, but the sizes differed markedly from those of EACMV-[TZ] (panel b). Seventeen percent (55/325) of the PCR-amplified samples gave this pattern. Another RFLP pattern (~ 1550 and ~ 1212 bp fragments) was observed for EACMV-[TZ2] in 15% (48/325) of the samples (panel c). EACMV-[TZ3] gave a pattern that differed from all the EACMV-like virus types described above (panel d). It was found in only 0.9% (3/325) of the analyzed samples. Another type, EACMV-[TZ4] gave only two RFLP fragments following digestion with *Mlu*I and was detected in 0.9% of the tested samples. EACMV-[TZ5] and EACMV-[TZ6] produced four and three RFLP fragments respectively with sizes that differed from each other (Panels e and f). They were found in 0.6 and 0.3% of the samples, respectively. Two other types, EACMV-[TZ7] and EACMV-[TZ8] (panel g) each gave different RFLP fragments in 0.3% of the samples. The remaining three types (EACMV-[TZ9], EACMV-[TZ10] and EACMV-[TZ11]) each yielded four *Mlu*I RFLP fragments that differed in size from each other (Panels h-j) as well as from EACMV-[TZ]. Only 0.3% of the restriction-analyzed samples contained these types. About 30% of the PCR-amplified DNA samples were found to contain EACMV-[UG2]. It produced four distinct fragments following digestion with *Mlu*I (Fig. 3.2).

Using the same restriction enzymes, PCR products from 10 samples gave a RFLP pattern that differed from both EACMVs and ACMVs. This virus type referred to here as EACMV-[TZ Mara] produced two fragments (~ 1625 and ~ 1125 bp) following restriction with *Mlu*I and three fragments after digestion with *Eco*RV (Fig. 3.2). *Mlu*I did not cleave ACMV-[NG] (Fig. 3.2) that was found in 7% of the samples but cleaved ACMV-[KE] into two fragments of ~ 1550 and 1212 bp. ACMV-[KE] was identified in only 0.6% of the samples. *Eco*RV detected a previously undescribed ACMV type (ACMV-[TZ]) in 0.9% of

the samples. When it was digested with *Mlu*I, a RFLP pattern distinct from that of ACMV-[KE] or ACMV-[NG] was produced (Fig. 3.2).

3.3.3 Phylogenetic analysis of AC1 gene sequences and comparison with selected viruses

A phylogenetic tree of selected CMGs is shown in Fig. 3.3a, b. The EACMV-like types separate into clusters of the reference sequences of EACMV isolates from GenBank regardless of whether sequence segments were generated by the forward (UNIF) or reverse (UNIR) primers, broadly confirming the validity of the RFLP results. Thus, ACMV-[TZ] is closely related to other ACMVs with > 90% nucleotide sequence identity. Phylogenetic analysis of the AC1 gene partial (360 nts) nucleotide sequences revealed molecular variability among the EACMV–like virus types from Tanzania. However some isolates appear to be closely related to one another. For example, when a comparison was made using sequence obtained with UNIF, EACMV-[TZ8] clustered close to EACMV-[TZ9], EACMV-[TZ3] to EACMV-[TZ5], EACMV-[TZ1] to EACMV-[TZ7], and EACMV-[TZ4] to EACMV-[TZ11] (Fig. 3.3a). EACMV-[TZ10] is closely related to a reference sequence of EACMV-[UG2] (Z83257) with sequence identity of > 90%. Phylogenetic comparison of the sequences generated with UNIR (Fig. 3.3b) showed all the EACMV-like virus types to be closely related to EACMV sequences from GenBank and only ACMV-[TZ] grouped closely to the reference sequences of ACMV types (Fig. 3.3b). As expected, none of the CMGs identified in this study clustered with ICMV from the Indian continent.

3.3.4 Relationship between virus type, symptom expression and disease severity

Symptoms were of different types in the different geographical locations and in different cassava cultivars. Affected cassava in the coastal regions (Tanga, Pwani, Dar-es-Salaam, Lindi and Mtwara) showed predominantly yellow mosaic that varied from mild light

yellow mosaic to severe yellowing and leaf distortion. In some fields, two types of mosaic could be observed: light-green and yellow mosaic in plants of the same cultivar. In the southern region of Ruvuma, particularly in the Mbinga district, CMD symptoms were predominantly severe yellow mosaic associated with leaf abscission. Over 80% of the affected cassava fields had plants that had very small leaves at the apex giving the plants the appearance of a 'sweeping broom'. In contrast to the symptoms observed elsewhere in the country, in the northwestern region of Kagera and in some parts of Mwanza region to the south of Lake Victoria, CMD symptoms were generally severe green mosaic with a yellow mosaic in isolated fields. In many instances, affected plants had extremely small leaves with distortion at the base of the leaflets, distinct chlorosis, leaf desiccation and death of some young plants. In the Mara region, severe symptoms were observed on the cultivars Lyongo, Lwabakanga and Kigoma.

An important consequence of the molecular diversity of CMGs could be variability of CMD symptom expression. In this study, CMD symptoms differed with virus type and ecological zone. Mild and severe symptoms or even different types of symptoms were sometimes observed on cassava plants of the same cultivar in the same field when infected by different types of EACMV like virus. ACMV infection was typically associated with green mosaic symptoms (Fig. 3.4a). EACMV-[TZ]-infected samples expressed mild (score 2) to severe (score 4 or 5) symptoms depending on cassava cultivar and location. Usually, symptoms consisted of yellow mosaic often resulting in severe distortion of leaves starting at the bases of the leaflets. Plants singly infected with EACMV-[TZ1] displayed moderate to severe yellow mosaic, of which 37% had a mean severity score of 3, with the rest either 2 or 4. Chlorosis was uniformly distributed on the whole leaflet with leaf abscission in severely affected plants (Fig. 3.4b). This symptom expression was consistent in all the affected cassava fields in the southern region. EACMV-[TZ2] was associated with moderate to severe green mosaic leading to severe leaf distortion and size reduction (Fig. 3.3f). Out of 48 plants infected by EACMV-[TZ2], 71% had CMD severity scores of 4 and 5. Cassava plants infected by EACMV-[TZ3] only displayed moderate (score 3) yellow mosaic with prominent small flecks on some parts of the leaflet (Fig. 3.3g). EACMV-[TZ4] caused severe downward cupping of leaves with chlorosis becoming more apparent

on the leaf margins while cassava plants infected by EACMV-[TZ5] showed severe to very severe yellow mosaic. EACMV-[TZ6] was associated with mild symptoms, often as flecks or spots that were distributed throughout the leaf area (Fig. 3.3c). Similarly, EACMV-[TZ7] induced mild yellow mosaic. In contrast, EACMV-[TZ8] was associated with severe green mosaic on all leaves. The other EACMV-like types, (EACMV-[TZ9], EACMV-[TZ10] and EACMV-[TZ11]), caused similar symptoms consisting of generalized green and yellow leaf mosaic. Most EACMV-[UG2]-infected plants had small leaves, plant stunting and desiccation of young leaves with an average score of 4 (Fig. 3.3d) although 13% showed mild (score 2) symptoms. EACMV-[TZ Mara] caused severe leaf distortion (severity score of 5) often associated with leaf abscission and desiccation, particularly of the young leaves (Fig. 3.3h).

3.3.5 Geographical distribution of CMGs in Tanzania

The CMGs detected during this study had largely separate but sometimes overlapping geographical distributions. EACMV-[TZ] occurred infrequently in the inland regions (Kagera, Mara and Mwanza) but occurred widely in the coastal regions (Dar-es-salaam, Pwani, Tanga, Mtwara) most particularly to the south near Lindi (Fig. 3.5). ACMV-[KE] and ACMV-[NG] were only found in the northwest. Many of the EACMV-like types had a localized geographical distribution. EACMV-[TZ1] was rarely found in the inland regions, was more common in the coastal regions but was particularly widespread in the Ruvuma region, where it occurred in 70% of all samples. EACMV-[TZ2] was virtually absent inland, but was common in the coastal regions, particularly in Tanga, Pwani and Lindi. It occurred at a single location in the Ruvuma region (Fig. 3.5). EACMV-[TZ3] was only detected in the Temeke district of Dar-es-Salaam, and EACMV-[TZ4] occurred only in Lindi and Mtwara. A similar localized pattern was observed for EACMV-[TZ5] and EACMV-[TZ6], which occurred only at Rufiji in the Pwani region and at Temeke, respectively. EACMV-[TZ8] and EACMV-[TZ11]) were found in Rufiji in the Pwani region, resulting in a total of three virus types including EACMV-[TZ5] identified in the same area. EACMV-[TZ7] was detected at only one location in Ruvuma. EACMV-like

types were generally rare in inland regions. However, EACMV-[TZ9] and EACMV-[TZ10] occurred in Kagera region. EACMV-[UG2] was not found in the southern part of the country. However, it was present in the Kagera (in 74% of the total collected samples), Mwanza and occasionally in the Mara region. RFLP analysis also suggested that it occurred in the coastal region of Tanga (Fig. 3.5), although given the distance from the known area of distribution of EACMV-[UG2], full length DNA-A sequences will be required to confirm this.

EACMV-[TZ Mara] occurred only in the Bunda and Musoma districts of the Mara region and was detected less commonly on the island of Ukerewe in Lake Victoria near Mwanza. It did not occur in any of the coastal regions. ACMV-[TZ] also occurred only in the Mara region.

3.3.6 Detection of co-infection in cassava plants by RFLP analysis

Twenty-four samples of CMD-infected cassava collected at widely separated locations during this study had dual infection (Table 3.1). Dual infections of ACMV-[NG] and EACMV-[UG2] were the most common. Dual infection was not found in samples from Mtwara, Lindi or Dar-es-Salaam in the coastal areas and rarely (1 of 71 samples) in the Ruvuma region in the south. In the Mwanza region (epidemic front of the severe CMD), however, six plants collected in and around Mwanza city had a dual infection of ACMV-[NG] and EACMV-[UG2], characteristic of the CMD epidemic front. More than 90% of the plants having dual virus infection expressed severe CMD symptoms. Symptoms of CMD on plants dual-infected with ACMV-[NG] and EACMV-[UG2] are shown (Fig. 3.4e). Mixed infection of different types of EACMV was very rare and no plants were infected by three types of CMG.

3.4 DISCUSSION

The universal oligonucleotide primer pair UNIF/UNIR successfully amplified near full-length fragments of CMG DNA-A. Restriction digestion analysis of the amplified DNA-A fragments obtained from 325 samples suggested that several viruses were present. In this study, and using RFLP as a criterion, we found that EACMV was more variable than ACMV although the latter was found in fewer samples. A total of 11 different EACMV-like virus types were distinguished from their RFLP polymorphic patterns, following digestion with *Eco*RV and *Mlu*I, and by partial nucleotide sequence comparisons. Another virus designated EACMV-[TZ Mara] that gave RFLP patterns different to those typically produced by ACMV or EACMV types was also detected. Only about 7% of the samples contained one of the three types of ACMV: ACMV-[KE], ACMV-[NG] and ACMV-[TZ]. Previous work has suggested that EACMV genomes are more genetically diverse than those of ACMV due to the frequent occurrence of recombination within their two components (Pita *et al*., 2001a, b; Ogbe *et al*., 2003). The results of the current study seemed to support this assertion.

In this paper, we describe 12 different EACMV-like virus types, all of which gave similar RFLP patterns, with the exception of EACMV-[TZ Mara]. The EACMV-like types were associated with distinct symptoms on cassava, probably reflecting differences in their virulence. Although symptom expression, particularly on different cultivars, may not provide significant evidence regarding pathogenicity of these virus types, it was interesting to note that the influence of virus on symptom expression was very apparent. For example, it was found that regardless of cultivar, more than 70% of all the plants singly-infected by EACMV-[TZ1] showed severe to very severe symptoms. Similarly, EACMV-[TZ2], EACMV-[TZ4], EACMV-[TZ5] and EACMV-[TZ8] were each associated with distinct and severe symptoms. In contrast, EACMV-[TZ6] and EACMV-[TZ7] were associated with only mild to moderate symptoms. Harrison *et al*. (1997) also reported significant differences between the symptoms of plants infected with ACMV (mild to moderate) and those infected with EACMV-[UG2] (severe) using cassava material collected in Uganda and kept under greenhouse conditions. More recently, both mild and severe strains of EACMV-[UG2] have been identified from Uganda (Pita *et al*., 2001a; Sseruwagi *et al*.,

2004). Due to the high disease severity associated with many of the Tanzanian CMGs described here, they have the potential to become a threat to cassava production in the country.

Phylogenetic analysis of the AC1 gene sequences of the CMGs strongly supports the classification based on the RFLP analysis. All the EACMV-like types cluster separately from ACMV and ACMV-[TZ] clusters with published sequences of ACMV isolates. Although complete nucleotide sequences of the viral DNA genome provide the most reliable means of geminivirus classification (Padidam *et al*., 1995; Fauquet *et al*., 2003), it is time consuming and expensive. We have shown that PCR and RFLP are useful tools for providing a preliminary yet useful categorization. EACMV-like types categorized based on RFLP as being EACMV clustered with published EACMV sequences obtained from GenBank with the exception of EACMV-[TZ Mara] that was categorized by RFLP as being different from EACMV yet clustered with EACMV isolates. This points out the limitation of RFLP, since a single base pair mutation can change a DNA sequence such that it forms a new restriction enzyme site, or alternatively removes a previously existing site leading to differences in RFLPs even between very closely related viruses. Importantly however, the power of the technique can be significantly enhanced by increasing the number of restriction enzymes used. However, RFLP analysis can serve as a useful 'first step' tool in investigating virus diversity as reported elsewhere for other geminiviruses (Willment *et al*., 2001). An important characteristic of the technique is that it can be readily done in a modestly equipped laboratory, a significant advantage in developing countries where facilities are limited.

Similar smaller studies using the RFLP approach for the investigation of CMG diversity have been undertaken in Guinea and Senegal (Okao-Okuja *et al*., 2004), Burundi (Bigirimana *et al*., 2004), Gabon (Legg *et al*., 2004) and Uganda (Sseruwagi *et al*., 2004). The current study of CMGs in Tanzania, however, has revealed a uniquely high level of diversity not encountered in any of these earlier studies conducted in diverse locations across Africa. The relative uniformity of ACMV, recorded here as elsewhere, when compared with the EACMVs, probably reflects the different molecular biology of the two

viruses, with the EACMVs appearing to be much more promiscuous and open to recombination events (Pita *et al*., 2001b). The unusual level of diversity within the EACMVs of Tanzania may have evolutionary significance, suggesting that this part of East Africa might have been the location for the first association of EACMVs with cassava. In this regard, it is significant that the first ever record of CMD was made from the Tanga region of Tanzania (Warburg, 1894).

Although 12 different EACMV-like virus types were reported here, five of these predominated. EACMV-[TZ], EACMV-[TZ1] and EACMV-[TZ2] all occurred throughout the country, but were most frequent in coastal and southern areas. By contrast, EACMV-[UG2] and EACMV-[TZ Mara] had more defined distributions, the former mainly down the western shores of Lake Victoria in the north-west and the latter on the eastern shores of the same lake. The distribution of EACMV-[UG2] largely matched that previously described for the CMD pandemic (Legg and Thresh, 2000), although RFLP results did suggest the occurrence of EACMV-[UG2] in northern coastal areas. If confirmed, this could represent a worrying possible expansion in the range of a CMG that has consistently been associated with rapidly spreading severe CMD (Harrison *et al*., 1997). Similar reports of EACMV-[UG2] have been made from parts of southern Africa, however, although there is no evidence in these cases to suggest an association with epidemic conditions (Berry and Rey, 2001). Full DNA-A sequences will be required to confirm the identify of the apparent EACMV-[UG2] obtained from north-eastern Tanzania. It is notable that prior to 1998, only ACMV had been recorded from the Kagera region (Ogbe *et al*., 1997), and the current study confirms the dramatic expansion in the range of EACMV-[UG2] that has taken place in recent years, in north-western Tanzania, as elsewhere in East and Central Africa (Harrison *et al*., 1997; Legg *et al*., 2001; Neuenschwander *et al*., 2002; Bigirimana *et al*., 2004).

Previous work using serology failed to detect ACMV in coastal Tanzania (Ogbe *et al*., 1997) and the present study has confirmed this. ACMV was not detected in any of the samples from the coastal regions (Tanga, Pwani, Dar-es-Salaam, Mtwara and Lindi) and Ruvuma in southern Tanzania. At the front of the severe CMD pandemic, ACMV-[NG]

was found frequently in a mixed infection with EACMV-[UG2] as reported by Zhou *et al.* (1997) and was associated with very severe symptoms, a phenomenon attributed to synergism between the two viruses (Harrison *et al*., 1997; Fondong *et al*., 2000; Pita *et al*., 2001a). The severity of disease caused by these mixed virus infections, the rapidity with which EACMV-[UG2] is spreading in Tanzania, and the potential for yet further virus interactions and recombination events all highlight the continued importance of the study of CMGs in Tanzania and the wider East African region. A sound understanding of the ecology and character of these viruses represents essential baseline information for the development and effective targeting of control measures.

3.5 ACKNOWLEDGEMENTS

The study was funded jointly by the UK's Department for International Development (DfID) through the Tropical Whitefly IPM Project, and a graduate fellowship from the International Institute of Tropical Agriculture (IITA), as part of the first author's Ph.D. research. The views expressed do not represent those of the DfID. The assistance of Mr. Cyprian Alloyce Rajabu of Plant Protection Division, Mwanza, during sample collection is highly appreciated. Special thanks are due to Mr. Okao-Okuja and Mr. Robert Obonyo for their technical assistance during sample analysis. The valuable comments of Professor Mike Thresh during the preparation of this manuscript are gratefully acknowledged.

3.6 REFERENCES

Berrie, L. C., Rybicki, E. P. & Rey, M. E. C. (2001). Complete nucleotide sequence and host range of South African cassava mosaic virus: Further evidence for recombination amongst begomoviruses. *Journal of General Virology* **82**, 53-58.

Berry, S. & Rey, M. E. C. (2001). Molecular evidence for diverse populations of cassava-infecting begomoviruses in southern Africa. *Archives of Virology* **146**,1795-1802.

Bigirimana, S., Barumbanze, P., Obonyo, R. & Legg, J. P. (2004). First evidence for the spread of *East African cassava mosaic virus* – Uganda (EACMV-UG) and the pandemic of severe cassava mosaic disease to Burundi. *Plant Pathology* **53,** 231.

Dellaporta, S. L., Wood, J. & Hicks, J. B. (1983). A plant DNA minipreparation: version II. *Plant Molecular Biology Reporter* **1,**19-21.

Fauquet, C. M., Bisaro, D. M., Briddon, R, W., Brown, J., Harrison, B. D., Rybicki, E. P., Stenger, D. C. & Stanley. J. (2003). Revision of taxonomic criteria for species demarcation in the *Geminiviridae* family, and a new updated list of begomovirus species. *Archives of Virology* **148**, 405-421.

Fondong, V. N., Pita, S. T., Rey, M. E. C., de Kochko, A., Beachy, R. N. & Fauquet, C. M. (2000). Evidence of synergism between African cassava mosaic virus and a new double recombinant geminivirus infecting cassava in Cameroon. *Journal of General Virology* **81,** 287-297.

Hanley-Bowdoin, L., Settlage, S. B., Orozco, B. M., Nagar, S. & Robertson, D. (1999). Geminiviruses: Models for plant DNA replication, transcription, and cell cycle regulation. *Critical Review in Plant Sciences* **18,** 71-106.

Harrison, B. D., Zhou, X., Otim-Nape, G. W., Liu, Y. & Robinson, D. J. (1997). Role of a novel type of double infection in the geminivirus-induced epidemic of severe cassava mosaic in Uganda. *Annals of Applied Biology* **131**, 437-448.

Legg, J. P. & Raya, M. D. (1998). Survey of cassava virus diseases in Tanzania. *International Journal of Pest Management* **44**, 17-23.

Legg, J. P. (1999). Emergence, spread and strategies for controlling the pandemic of cassava mosaic disease in eastern and central Africa. *Crop Protection* **18**, 627-637.

University of Pretoria etd – Ndunguru, J (2005)

Legg , J. P. & Thresh, J. M. (2000). Cassava mosaic virus disease in East Africa: A dynamic disease in a changing environment. *Virus Research* **71,** 135-149.

Legg, J. P., Okao-Okuja, G., Mayala, R. & Muhinyuza, J-B. (2001). Spread into Rwanda of the severe cassava mosaic virus disease pandemic and associated Uganda variant of East African cassava mosaic virus (EACMV-UG). *Plant Pathology* **50**, 796.

Legg, J. P., Ndjelassili, F. & Okao-Okuja. G. (2004). First report of cassava mosaic disease and cassava mosaic geminiviruses in Gabon. *Plant Pathology* **53**, 232.

Maruthi, M. N., Colvin, J., Seal, S. & Thresh, J. M. (2002). First report of a district begomovirus infecting cassava in Zanzibar. *Plant Disease* **86**, 187.

Neuenschwander, P., Hughes, J. d'A., Ogbe, F., Ngatse, J. M. & Legg. J. P. (2001). The occurrence of the Uganda Variant of East African Cassava Mosaic Virus (EACMV-Ug) in western Democratic Republic of Congo and the Congo Republic defines the westernmost extent of the CMD pandemic in East/Central Africa. *Plant Pathology* **51,** 384.

Ogbe, F. O., Songa, W. & Kamau, J. W. (1996). Survey of the incidence of African cassava mosaic and East African cassava mosaic viruses in Kenya and Uganda using a monoclonal antibody based diagnosis test. *Roots* **3**, 10-13.

Ogbe, F. O., Legg, J., Raya, M. D., Muimba-Kankolongo, A., Theu, M. P., Kaitisha, G., Phiri, N. A. & Chalwe, A. (1997). Diagnostic survey of cassava mosaic viruses in Tanzania, Malawi and Zambia. *Roots* **4**, 12-15.

Ogbe, F. O., Thottappilly, G., Dixon, A. G. O., Atiri, G. I. & Mignouna, H. D. (2003). Variants of *East African cassava mosaic virus* and its distribution in double infections with *African cassava mosaic virus* in Nigeria. *Plant Disease* **87**, 229-232.

Okao-Okuja, G., Legg, J. P., Traore, L. & Alexandra, J. M. (2004). Viruses associated with cassava mosaic disease in Senegal and Guinea Conakry. *Journal of Phytopathology* **152**, 69-76.

Padidam, M., Beachy, R. N. & Fauquet, C. M. (1995). Tomato leaf curl geminivirus from India has a bipartite genome and coat protein is not essential for infectivity**.** *Journal of General Virology* **76,** 25-35.

Pita, S., Fondong, V. N., Sangare, A., Otim-Nape, G. W., Ogwal, S. & Fauquet. C. M. (2001a). Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *Journal of General Virology* **82**, 655-665.

Pita, S., Fondong, V. N., Sangare, A., Kokora, R. N. N. & Fauquet. C. M. (2001b). Genomic and biological diversity of the African cassava geminiviruses. *Euphytica* **120,** 115-125.

Rey, M. E. C. & Thompson. G. (1998). Cassava mosaic virus disease in South Africa. *Roots* **5**, 3-5.

Rogers, S. G., Bisaro, D. M., Horsch, R. B., Fraley, R. T., Hoffmann, N. L., Brand, L., Elmer, J. C. & Lloyd, A. M. (1986). Tomato golden mosaic virus A component DNA replicates autonomously in transgenic plants. *Cell* **45**, 593-600.

Sseruwagi, P., Rey, M. E. C., Brown, J. K. & Legg, J. P. (2004). The cassava mosaic geminiviruses occurring in Uganda following the 1990s epidemic of severe cassava mosaic disease. *Annals of Applied Biology* **145**, 113-121.

Sunter, G., Gardiners, W. E., Rushing, A. E., Rogers, S. G. & Bisaro, D. M. (1987). Independent encapsidation of tomato golden mosaic virus A component DNA in transgenic plants. *Plant Molecular Biology* **8**, 477-484.

Swanson, M. M. & Harrison, B. D. (1994). Properties, relationships and distribution of cassava mosaic geminiviruses. *Tropical Science* **34**,15-25.

Townsend, R., Watts, J. & Stanley, J. (1986). Synthesis of viral DNA forms in *Nicotiana plumbaginifolia* protoplasts inoculated with cassava latent virus (CLV): Evidence for independent replication of one component of the CLV genome. *Nucleic Acids Research* **14**, 1253-1255.

Warburg, O. (1894). Die kulturpflanzen usambaras. *Mitteilungen aus den Deutschen Schutzgebieten* **7,** 131.

Willment, J. A., Martin, D. P. & Rybicki, E. P. (2001). Analysis of the diversity of African streak mastreviruses using PCR-generated RFLPs and partial sequence data. *Journal of Virological Methods* **93,** 75-87.

Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G. W., Robinson, D. J. & Harrison B. D. (1997). Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *Journal of General Virology* **78**, 2101-2111.

Table 3.1. Extent of CMG co-infection in cassava plants with CMD symptoms

Virus type	No. of samples	Location *	Mean CMD severity score (1-5 scale)
ACMV-[NG] + ECMV-UG2	20	Kagera (13), Mwanza (6), Mara (1)	4.5
ACMV-[KE] + ECMV-UG2	1	Mwanza (1)	4
ECMV-[TZ] + ECMV-[TZ1]	1	Ruvuma (1)	4
ECMV-[TZ2] + ECMV-UG2	2	Pwani (2)	5
Total	24		

* Number in parentheses refer to total samples with the mixed infection

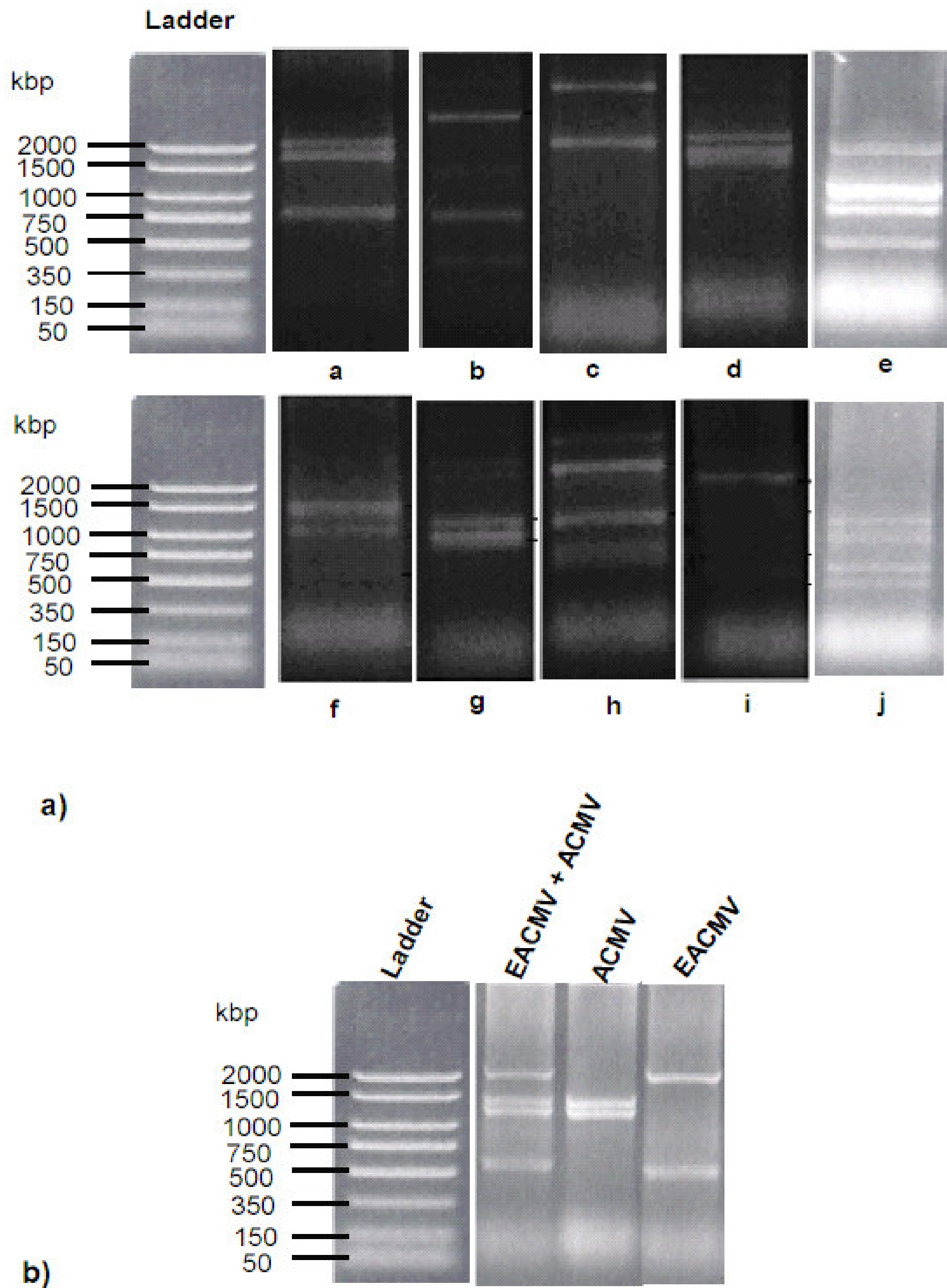


Fig. 3.1. Restriction fragment length polymorphism patterns of 2.8 Kb DNA fragments of selected putative cassava mosaic geminiviruses. EACMV-like virus pattern of DNA fragments digested with *Mlu*I. (Top a): (a) EACMV-[TZ], (b) EACMV-[TZ1], (c) EACMV-[TZ2], (d) EACMV-[TZ3], (e) EACMV-[TZ5], (f) EACMV-[TZ6], (g)

EACMV-[TZ8], (h) EACMV-[TZ9], (i) EACMV-[TZ10], and (j) EACMV-[TZ11]. (Bottom b): Banding patterns showing detection of virus dual infection of EACMV and ACMV by restriction analysis with *Eco*RV endonuclease. DNA size markers are indicated in base pairs in the left margins.

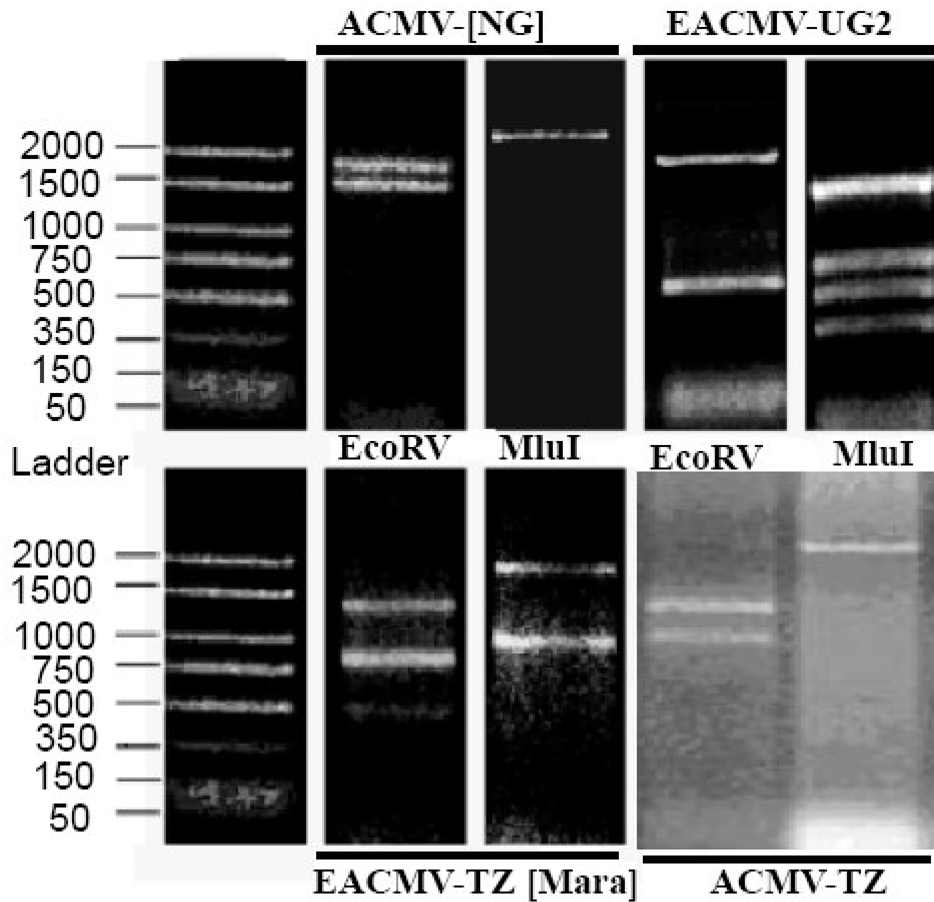
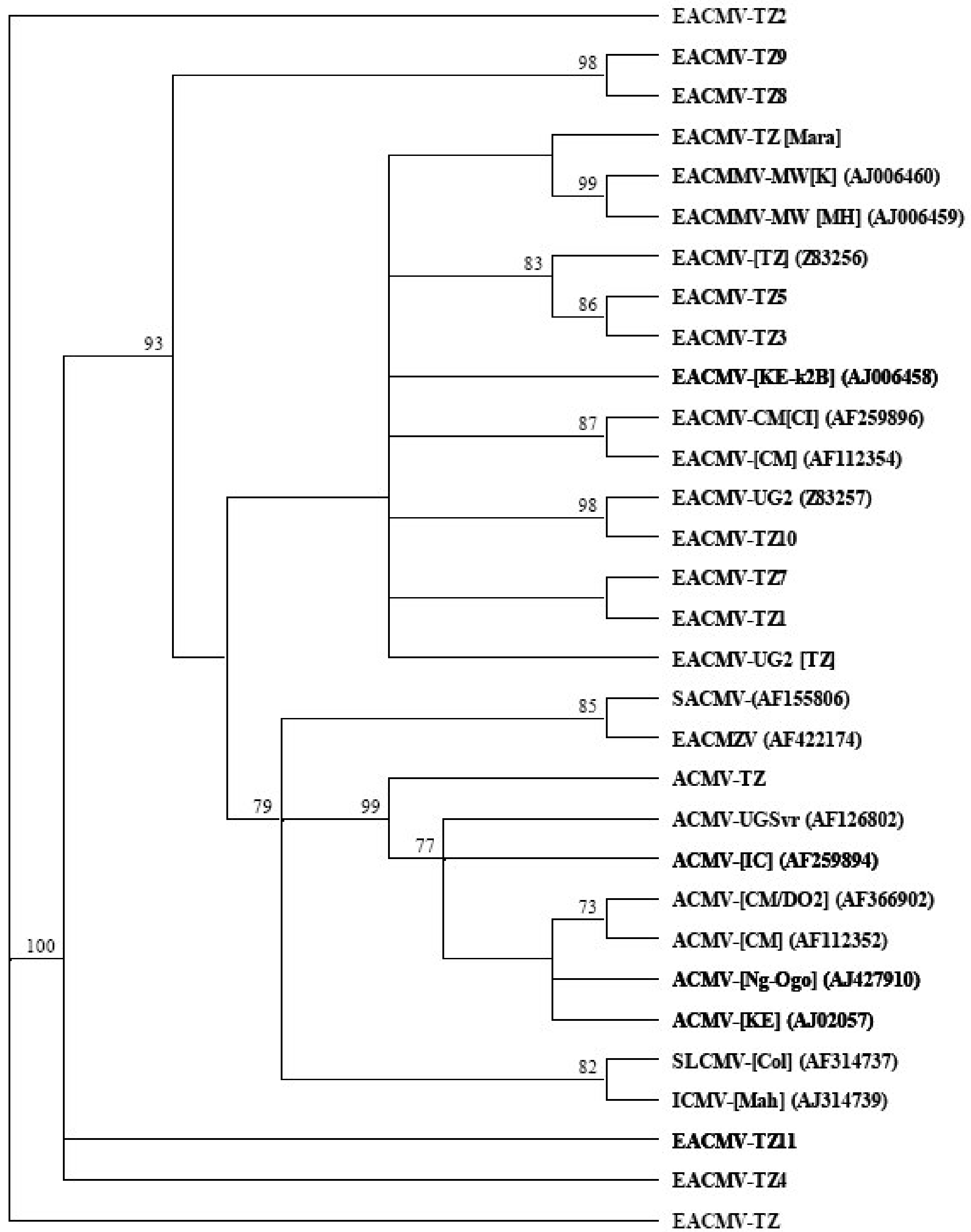
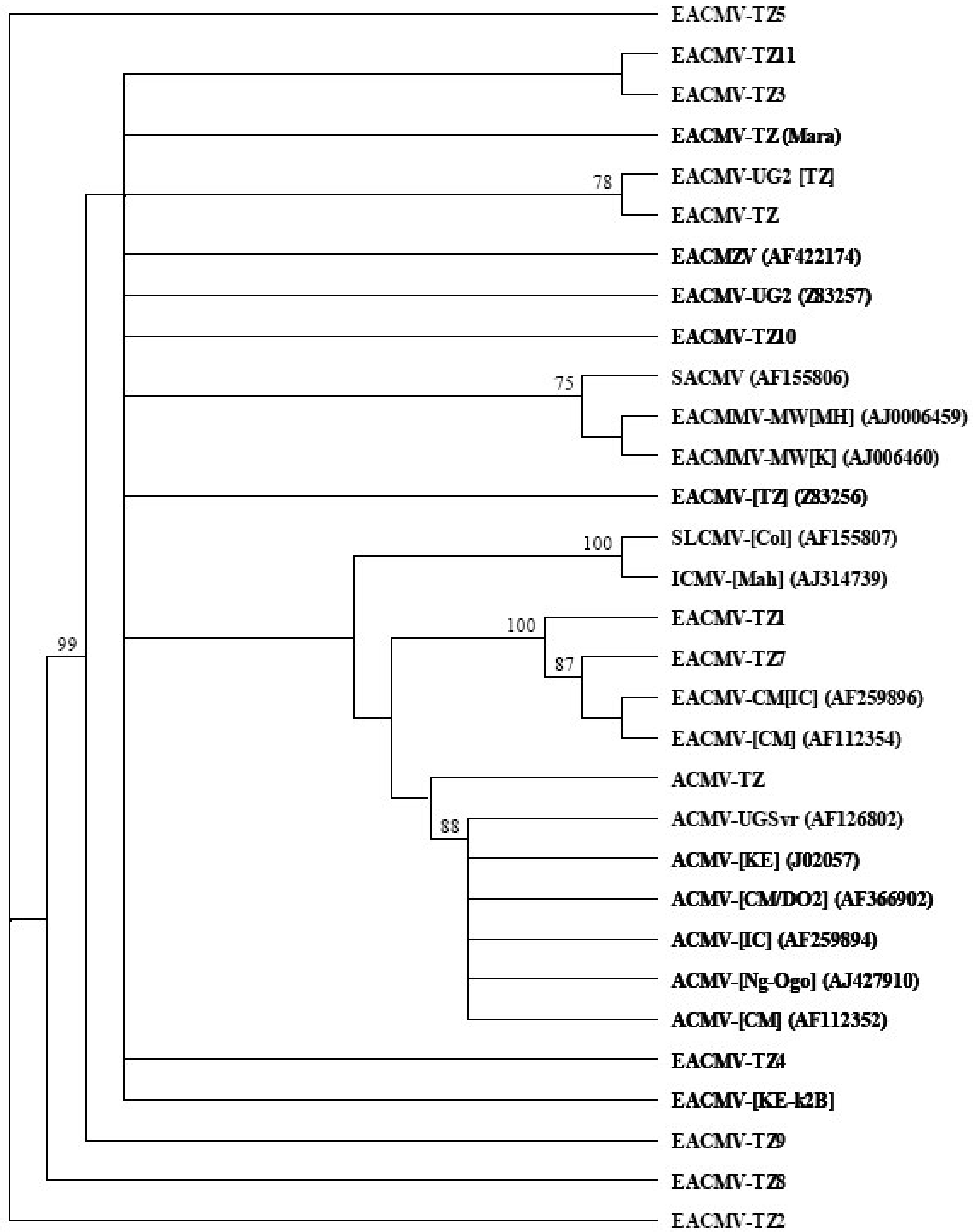


Fig. 3.2. Restriction fragment length polymorphism patterns of 2.8 Kb DNA fragment of cassava mosaic geminiviruses, EACMV-[TZ Mara], EACMV-[UG2], ACMV-[NG] and ACMV-[TZ]. DNA fragments were separately digested with *Eco*RV and *Mlu*I.



a)



b)

Fig. 3.3. Neighbor-joining phylogenetic dendrogram based upon alignments of partial AC1 (360 nts) sequences of CMG DNA-A (Tanzanian isolates) and selected other begomoviruses. The sequences were either generated by forward (UNIF) (a) or reverse (UNIR) (b) primers which anneal at position 1700-1834 of the AC1 gene on the complementary sense of the DNA-A component. Numbers at nodes indicate percentage bootstrap scores (1000 replicates).

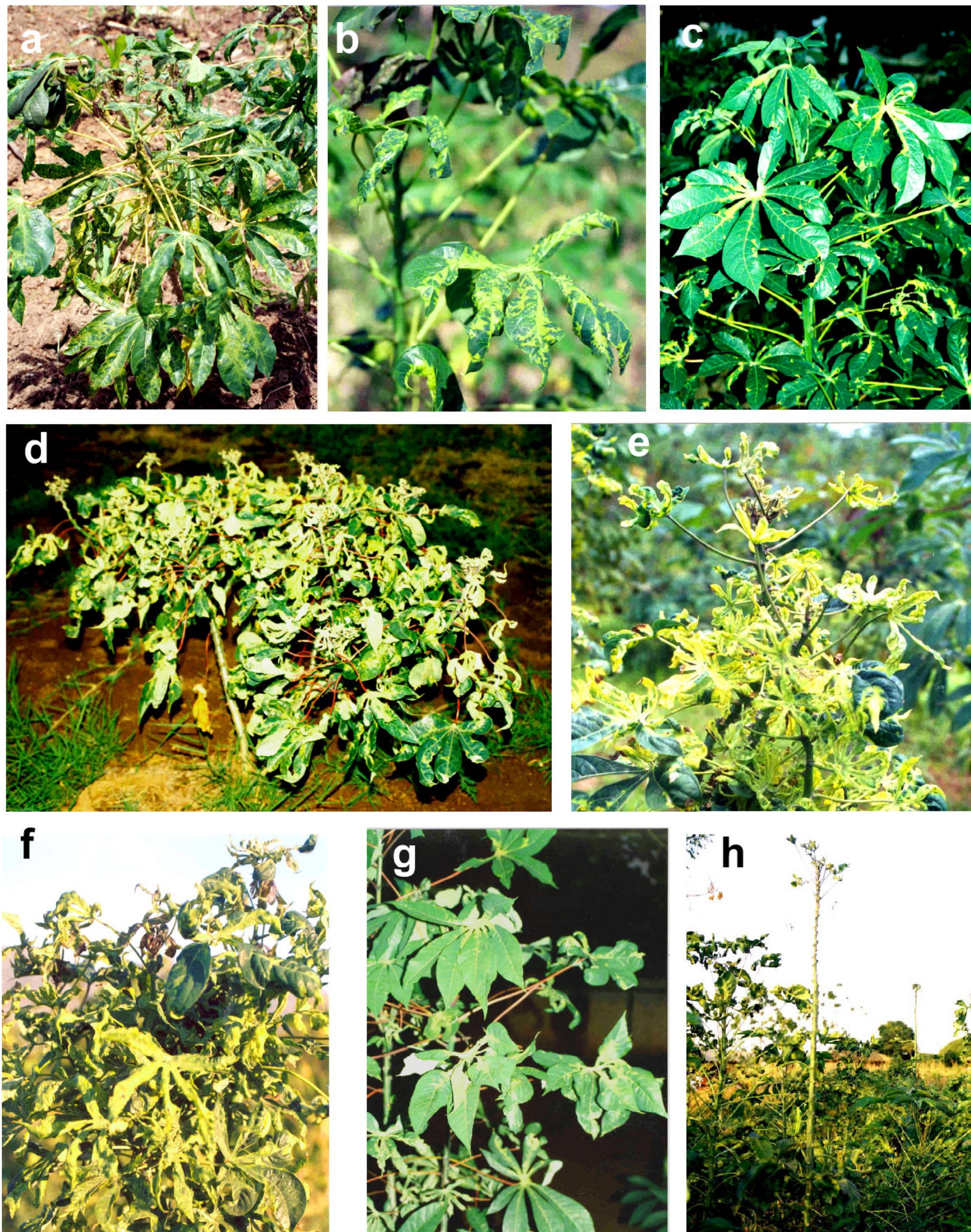


Fig. 3.4. Cassava mosaic disease symptoms caused by selected cassava mosaic geminiviruses in farmers' fields (a-h). Symptoms are on cassava cultivars; (a) Lyongo (caused by ACMV-[NG]), (b) Nakalai (EACMV-[TZ1]), (c) Jawa (EACMV-[TZ6]), (d) Gajaigaja (EACMV-[UG2]), (e) Unnamed local (mixed infection of EACMV-[UG2] and ACMV-[NG]), (f) Msufi (EACMV-[TZ2]), (g) Jawa (EACMV-[TZ3]), (h) Kigoma (EACMV-[TZ Mara]).

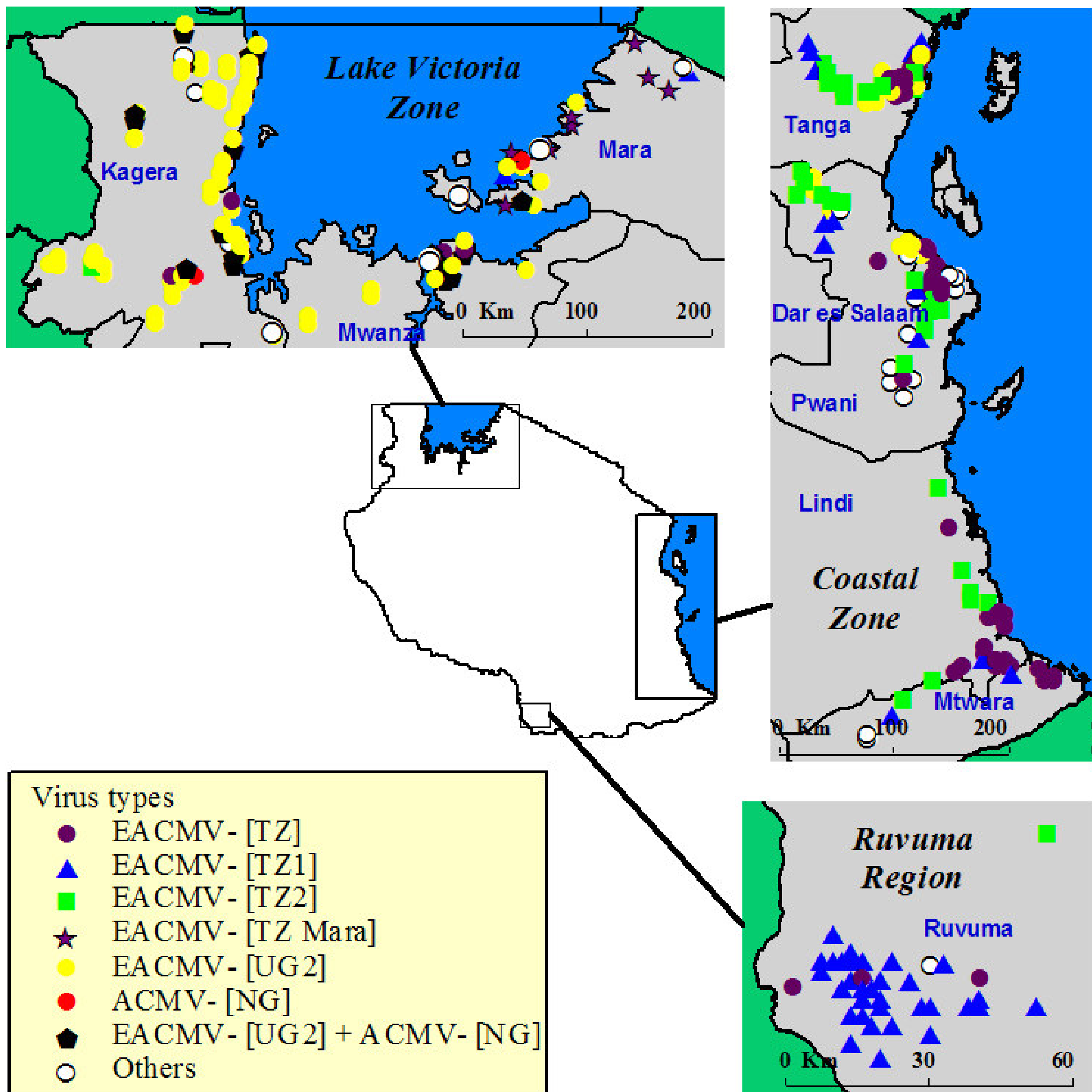


Fig. 3.5. Map of Tanzania illustrating the distribution of the most prevalent cassava mosaic geminiviruses in the major cassava-growing zones. Only the widespread cassava mosaic geminiviruses are listed and the rest are shown as "others".