Analysis of cassava brown streak viruses reveals the presence of distinct virus species causing cassava brown streak disease in East Africa

Stephan Winter, Marianne Koerbler, Beate Stein, Agnes Pietruszka, Martina Paape and Anja Butgereitt

DSMZ Plant Virus Department, Braunschweig, Germany

Cassava brown streak virus (CBSV) isolates were analysed from symptomatic cassava collected between 1997 and 2008 in the major cultivation regions of East Africa. An analysis of complete RNA genomes of seven isolates from Kenya, Tanzania, Mozambique, Uganda and Malawi revealed a common genome structure, but the isolates clearly clustered in two distinct clades. The first comprised isolates from Kenya, Uganda, Malawi, north-western Tanzania and the CBSV described previously, and shared between 87 and 95% nucleotide sequence identity, whilst the second included isolates from coastal regions of Mozambique and Tanzania, which shared only 70% nucleotide sequence identities with isolates of the first clade. When the amino acid sequences of viral proteins were compared, identities as low as 47% (Ham1) and 59% (P1) between the two clades were found. An antiserum obtained against the capsid protein of a clade 1 isolate identified a 43 kDa protein in clade 1 isolates and a 45 kDa protein in clade 2 isolates. Several cassava cultivars were susceptible to isolates of clade 2 but resistant to those of clade 1. The differences observed both in biological behaviour and in genomic and protein sequences indicate that cassava brown streak disease in East Africa is caused by at least two distinct virus species. It is suggested that those of clade 1 retain the species name *Cassava brown streak virus*, whilst those of clade 2 be classified as *Cassava brown streak Mozambique virus*.

INTRODUCTION

Cassava brown streak disease (CBSD) is currently considered to be the most devastating disease of cassava (*Manihot esculenta* Crantz) in coastal areas of East Africa. The disease was first reported in 1936 from Tanzania (Storey, 1936) and in 1950 from Malawi (Nichols, 1950) and is now found in all areas, from Kenya to Mozambique, where this important food crop is cultivated (Alicai et al., 2007).

Cassava brown streak virus (CBSV) was first confirmed as the causal agent of CBSD by sap inoculation from cassava to herbaceous hosts and back transmission to cassava (Lister, 1959). Filamentous virus-like particles of 650 nm in length and pinwheel inclusions found in cells of CBSV-infected plants (Harrison et al., 1986; Were et al., 2004b) indicated CBSV as a putative member of the family *Potyviridae*. The taxonomic affiliation of the species *Cassava brown streak virus* as a member of the genus *Ipomovirus*, family *Potyviridae*, was clarified by partial sequencing of the 3′-terminal regions of the ssRNA genomes of virus isolates from Tanzania (Monger et al., 2001b), whilst a complete genome sequence for a CBSV isolate from north-western Tanzania, CBSV MLB 3, was determined recently (Mbanzibwa et al., 2009a).

A high genetic diversity of CBSV isolates has been indicated both by variations in symptoms in *Nicotiana benthamiana* and from the diversity of short coat protein (CP) sequences (Monger et al., 2001a). The recent study by Mbanzibwa et al. (2009b) comparing the complete CP of CBSV isolates collected in the Lake Victoria basin in north-western Tanzania and Uganda with CP sequences from viruses collected in Mozambique and coastal Tanzania (Monger et al., 2001a) indicated two geographically separated and distinct virus populations implicated in CBSD. However, from the limited sequence data, a definite taxonomic classification of the geographical isolates of CBSV was not possible. This raised the question of whether CBSD in East Africa is caused by a single virus with diverse isolates and strains, or whether several distinct viruses are implicated in this disease.

In this study, we present the results of a biological and molecular characterization of CBSV isolates collected between 1997 and 2008 from all major cassava regions in East Africa. The phylogenetic analysis involving complete genome sequences of seven CBSV isolates is presented as molecular evidence for the existence of two virus species causing CBSD. Analysis of symptoms on indicator hosts
and differential resistance in cassava cultivars are further discussed to support a demarcation of CBSV as two separate virus species.

**RESULTS**

**CBSV infections in cassava**

Leaf symptoms of CBSV-infected cassava collected from diverse geographical regions did not differ significantly in this plant, but when herbaceous hosts were inoculated, differential symptoms were observed, which were most pronounced in *N. benthamiana* (Fig. 1a–f, inserts). Inoculation of CBSV isolates from Kenya, Uganda and Malawi caused vein clearing, systemic mottling and leaf curling, whilst infections of isolates collected at sites in Mozambique and Tanzania resulted in local lesions on inoculated leaves and severe vein chlorosis and necrosis on systemically infected leaves. The most severe infections of *N. benthamiana* were recorded for isolates from Mozambique with plants dying 5–8 days after inoculation.

From infected *N. benthamiana*, the disease was back transmitted to cassava by mechanical inoculation of sap. The cassava cultivars TME 1, TME 3, TME 4, TMS 96/0304, TMS 96/0160 and TMS 96/0529 developed typical brown streak symptoms 6–8 weeks after inoculation (Fig. 1).

The cassava cultivar response to virus infection was evaluated by graft transmission of CBSV isolates using scions from TMS 96/0304, a susceptible breeding line used for *in situ* maintenance and propagation of CBSV. All cassava lines were susceptible to infection with isolates from Mozambique and Tanzania and developed distinct disease symptoms on older leaves approximately 6–10 weeks after grafting. Cassava inoculation with isolates from Kenya, Uganda and Malawi resulted in virus infections in TME 1, TME 3, TME 4, TMS 96/0160 and TMS 96/0529, whilst graft transmission failed to infect cultivars Albert, Kibaha and TMS 30572. These cultivars, irrespective of their susceptibility to isolates from Mozambique and Tanzania, were resistant to infections with those isolates from Kenya, Uganda and Malawi. Based on similarities of symptoms in *N. benthamiana* and the resistance response in cassava, CBSV isolates were assigned to clades. Viruses from Kenya, Uganda and Malawi were assigned to clade 1, whilst those of Tanzania and Mozambique were assigned to clade 2.

**Structure of CBSV genomes**

The complete genome sequences of isolates from clades 1 and 2 were assembled to reveal genome sizes of 9070 nt for CBSV isolates from Kenya, Uganda and Malawi (clade 1) and somewhat smaller RNA genomes for clade 2 isolates from Mozambique (9008 nt) and Tanzania (8997 nt).
Translation of the RNA genome sequences revealed a single large open reading frame (ORF) of 2902 aa for clade 1 isolates. For clade 2 virus genomes, an ORF of 2912 aa for a CBSV isolate from Tanzania and an ORF of 2916 aa for a CBSV Mozambique isolate was determined starting after a short 5’ untranslated region (5’UTR) with the consensus motif AAAAAATAAA. The polyprotein ORFs were followed by a 3’UTR of 230 nt (clade 1) or 133 nt (clade 2), followed by a poly(A) tail. The polyprotein encoded the proteins typically found in genomes of the members of the family Potyviridae (P1, P3, 6K1, CI, 6K2, VPg, Nla-Pro, Nlb and CP), but with the gene for HC-Pro missing. An additional sequence inserted between Nlb and CP, identified upon Pfam analysis (http://pfam.sanger.ac.uk) as encoding a Ham1-like protein (Ham1), was present within the polyprotein of all CBSV genomes.

**CBSV CPs**

The CP cistron was predicted by aligning the CBSV polyprotein with cucumber vein yellowing virus (CVYV) amino acid sequences to identify cleavage sites for the Nla protease between the replicase Nlb and Ham1 and further downstream between Ham1 and the putative CP. The cleavage site Q/A in the context IDVQ↓A would result in CPs of 367 aa for clade 1 isolates starting with Q↓ALNQEEIE to yield predicted CPs of approximately 41.5 kDa. In the polyprotein sequence of clade 2 isolates, a CP of 387 aa starting with Q↓AIDKDEIE and resulting in a CP of 41 kDa was determined. Expression of the putative CP gene of a clade 1 isolate (Ke_125) in bacteria and a CP of 43 kDa was determined. Expression of the putative CP of 387 aa starting with Q

41.5 kDa. In the polyprotein sequence of clade 2 isolates, a Q

CPs of 367 aa for clade 1 isolates starting with cleavage site Q/A in the context IDVQ

downstream between Ham1 and the putative CP. The protease between the replicase NIb and Ham1 and further amino acid sequences to identify cleavage sites for the NIa protease.

The CP cistron was predicted by aligning the CBSV polyprotein with cucumber vein yellowing virus (CVYV) amino acid sequences to identify cleavage sites for the Nla protease between the replicase Nlb and Ham1 and further downstream between Ham1 and the putative CP. The cleavage site Q/A in the context IDVQ↓A would result in CPs of 367 aa for clade 1 isolates starting with Q↓ALNQEEIE to yield predicted CPs of approximately 41.5 kDa. In the polyprotein sequence of clade 2 isolates, a CP of 387 aa starting with Q↓AIDKDEIE and resulting in a CP of 41 kDa was determined. Expression of the putative CP gene of a clade 1 isolate (Ke_125) in bacteria and production of an antiserum against the recombinant protein resulted in experimental confirmation of the predicted CP. In Western blot analysis using these antibodies, CPs of approximately 41 kDa were determined for clade 1 isolates (Fig. 2a, lanes 1–3) which, as predicted from the amino acid sequence, were distinctly smaller than those of clade 2 isolates (Fig. 2a, lanes 4, 5 and 7).

A phylogenetic analysis of CBSV CP sequences, including complete CP sequences of CBSV isolates from the recent study by Mbanzibwa et al. (2009b), provided further evidence to support the existence of two distinct CBSV populations. CBSV clade 1 isolates together with virus isolates from the Lake Victoria zone in Uganda and from north-western Tanzania (Mbanzibwa et al., 2009b) formed a cluster, with amino acid identities between 94 and 96% to the Ke_125 isolate (Table 1). In contrast, CP sequences of clade 2 isolates shared only 79–81% amino acid identities with CP of clade 1 isolates (80% amino acid identities with the Ke_125 isolate). This divergence between clade 1 and clade 2 isolates was because of striking differences in the N-terminal aa 1–120 portions of the CP. The differences in the CP were most pronounced at aa 70–110 with an insertion of 9 aa (aa 101–110 in clade 2 isolates), resulting in a CP slightly larger than that of clade 1 isolates (Fig. 2a). Sequence conservation in the core and C-terminal regions of CBSV CP was also found when CVYV CP sequences were included in clustal_x alignments.

This was confirmed in Western blot analysis with protein extracts of CVYV-infected plants showing a clear signal at approximately 41.5 kDa (Fig. 2b). In contrast, a serological cross-reaction with sweet potato mild mottle virus (SPMMV), the type species of the genus Ipomovirus, was not predicted from sequence analysis nor found experimentally (Fig. 2b).

### CBSV P1 proteins

As with the genomes of CVYV (Janssen et al., 2005) and squash vein yellowing virus (SqVYV; Li et al., 2008) lacking an HC-Pro gene, P1 was directly followed by P3 and separated by a predicted autocatalytic cleavage site Y↓S in the consensus sequence I(E/D)(L/M)Y↓S. However, unlike the previously reported P1 protein from CVYV formed by two homologous proteins, P1a and P1b (Valle et al., 2007), CBSV P1 is a single serine protease and acts as viral suppressor of RNA silencing (Mbanzibwa et al., 2009a). P1 proteins of CBSV were divergent and only about 60% amino acid identity was found between clade 1 and clade 2 isolates (Table 1).

### CBSV Ham1-like proteins

The putative Maf/Ham1 pyrophosphatase identified in the genomes of CBSV isolates from Uganda and Tanzania...
Table 1. Identities of genome (% nucleotide identity), polyprotein (% amino acid identity) and mature proteins (% amino acid identity) between a clade 1 isolate, Ke_125, and CBSV isolates from Kenya, Uganda, Malawi, Tanzania and Mozambique

Sequences of CBSV MLB 3 (Mbanzibwa et al., 2009a) and CBSV Tan Z (GenBank accession no. GQ329864) and of the respective proteins of CVYV, SqVYV and SPMMV were included in the analysis. Isolates similar to Ke_125 representing clade 1 isolates are shown in bold.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>P1</th>
<th>P3</th>
<th>6K1</th>
<th>CI</th>
<th>6K2</th>
<th>Nla-VPg</th>
<th>Nla-Pro</th>
<th>NIb</th>
<th>Ham 1</th>
<th>CP</th>
<th>Polyprotein</th>
<th>Genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ke_54</td>
<td>95</td>
<td>99</td>
<td>96</td>
<td>98</td>
<td>100</td>
<td>96</td>
<td>98</td>
<td>98</td>
<td>92</td>
<td>95</td>
<td>97</td>
<td>93</td>
</tr>
<tr>
<td>Ug_23</td>
<td>97</td>
<td>98</td>
<td>99</td>
<td>100</td>
<td>97</td>
<td>98</td>
<td>98</td>
<td>99</td>
<td>94</td>
<td>96</td>
<td>98</td>
<td>95</td>
</tr>
<tr>
<td>Ma_42</td>
<td>96</td>
<td>98</td>
<td>96</td>
<td>99</td>
<td>96</td>
<td>97</td>
<td>99</td>
<td>97</td>
<td>89</td>
<td>96</td>
<td>97</td>
<td>93</td>
</tr>
<tr>
<td>Ma_43</td>
<td>96</td>
<td>99</td>
<td>98</td>
<td>99</td>
<td>96</td>
<td>97</td>
<td>99</td>
<td>97</td>
<td>90</td>
<td>96</td>
<td>97</td>
<td>93</td>
</tr>
<tr>
<td>MLB 3</td>
<td>86</td>
<td>88</td>
<td>94</td>
<td>97</td>
<td>94</td>
<td>91</td>
<td>96</td>
<td>95</td>
<td>89</td>
<td>94</td>
<td>93</td>
<td>87</td>
</tr>
<tr>
<td>Mo_83</td>
<td>59</td>
<td>64</td>
<td>87</td>
<td>83</td>
<td>75</td>
<td>72</td>
<td>82</td>
<td>83</td>
<td>47</td>
<td>80</td>
<td>74</td>
<td>70</td>
</tr>
<tr>
<td>Tan Z</td>
<td>60</td>
<td>64</td>
<td>85</td>
<td>83</td>
<td>79</td>
<td>73</td>
<td>80</td>
<td>83</td>
<td>48</td>
<td>80</td>
<td>74</td>
<td>71</td>
</tr>
<tr>
<td>Tan_70</td>
<td>61</td>
<td>63</td>
<td>89</td>
<td>84</td>
<td>86</td>
<td>72</td>
<td>79</td>
<td>83</td>
<td>48</td>
<td>80</td>
<td>74</td>
<td>70</td>
</tr>
<tr>
<td>CVYV</td>
<td>9/30*</td>
<td>47</td>
<td>65</td>
<td>67</td>
<td>36</td>
<td>52</td>
<td>47</td>
<td>64</td>
<td>–</td>
<td>67</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>SqVYV</td>
<td>9/30*</td>
<td>43</td>
<td>56</td>
<td>65</td>
<td>35</td>
<td>50</td>
<td>46</td>
<td>64</td>
<td>–</td>
<td>64</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>SPMMV</td>
<td>30</td>
<td>31</td>
<td>33</td>
<td>49</td>
<td>22</td>
<td>28</td>
<td>29</td>
<td>54</td>
<td>–</td>
<td>35</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

*Values from comparing P1a and P1b separately.

(Mbanzibwa et al., 2009a) was confirmed for all CBSV isolates. Ham1 genes from clade 1 isolates were highly similar (89–94 % amino acid identity) and aligned well with reported Ham1 sequences of the earlier study. In contrast, only a limited relationship (47–48 % amino acid identity) among Ke_125 Ham1 clade 1 and clade 2 proteins was found (Table 1). The genetic distance between the clades was as far as between CBSV Ham1 sequences and those of other eukaryotes (Fig. 3a). Thus, apart from size (226 aa) and short amino acid motifs conserved in Ham1 proteins (Mbanzibwa et al., 2009a), there were no additional informative motifs indicating a common origin of Ham1 in the genomes of CBSV clade 1 and clade 2 isolates. The only other plant virus Ham1-like sequence in Euphorbia ringspot virus (EuRV) represented an even more divergent sequence (34 % amino acid identity).

Diversity of CBSV genomes

A comparison of CBSV proteins with the respective proteins of other members of the genus Ipomovirus (Table 1) showed most similarities with CVYV and SqVYV, whilst the relationship with SPMMV was very limited. A comparison of complete CBSV genome sequences (Fig. 3b) further confirmed that CBSV clade 1 isolates from Kenya, Uganda and Malawi were closely related, with an overall nucleotide sequence identity of 93–95 %. The genome sequence of the CBSV isolate MLB 3 (Mbanzibwa et al., 2009a) from north-western Tanzania grouped with clade 1 isolates but represented a divergent CBSV isolate (87 % nucleotide identity), as had already been indicated from a comparison of viral proteins (Table 1).

Clade 2 isolates from Mozambique formed a sequence cluster separated at a 70 % nucleotide identity level (74 % amino acid identity) from clade 1 isolates. In clade 2 virus genomes, the diversity of the individual genes was dispersed evenly over the entire polyprotein. Interestingly, the genome sequence of a Tan Z isolate (ACT78701) from Musoma district at Lake Victoria in Tanzania was more related to virus isolates collected in Mozambique than to clade 2 viruses collected in the Kibaha region of Tanzania. CBSV isolates collected in this coastal area had genome sequences quite distinct from other clade 2 isolates (Fig. 3b), which was mainly due to strikingly different P1 genes. Whilst the P1 sequences of other CBSV clade 2 virus isolates shared 96 % amino acid identity, P1 proteins of the Kibaha viruses were equidistant to clade 2 viruses (68 % amino acid identity) and to clade 1 viruses (61 % amino acid identity) (Table 1).

DISCUSSION

CBSD is a widely occurring virus disease of cassava and has been reported from a diverse range of agro-ecological zones in East Africa (Alicai et al., 2007). A considerable variation in isolates has been found (Monger et al., 2001a), which has been further substantiated by a phylogenetic analysis of partial CP sequences (Mbanzibwa et al., 2009b) leading to the description of two genetically distinct populations of CBSV occurring in Lake Victoria and the coastal areas of Tanzania and Mozambique.

In this study, we isolated CBSV from cassava, and determined and assembled seven complete genome sequences of isolates collected in all major regions of cassava cultivation in East Africa. The phylogenetic analysis of genome sequences (70 % nucleotide identity) and of mature proteins clearly indicated two clades, with viruses from Kenya, Uganda and Malawi forming clade 1, close to CBSV MLB 3 (Mbanzibwa et al., 2009a), and those collected in Mozambique and Tanzania comprising clade 2. The different sizes of CPs between the two clades, predicted by the
sequence analysis, were confirmed by Western blot analysis using antibody against recombinant CP. Earlier studies to determine the CP size of a CBSV isolate from Tanzania using an antibody against an unknown ‘carlavirus-length’ virus (Monger et al., 2001b) determined a CP of approximately 45 kDa, the size estimated for a clade 2 virus. CBSV clade 1 and clade 2 isolates can be discriminated based on differences in the molecular mass of their CP in Western blot analysis but also from their characteristic N-terminal CP sequences. Nevertheless, the overall CP amino acid sequence identity of 80 % between clade 1 and clade 2 viruses was on the threshold for ipomovirus species demarcation at 80 % amino acid identity (Adams et al., 2005b).

Considering the diversity of CBSV, Ham1 was the most divergent gene among isolates of clade 1 and clade 2. This gene was also found in EuRV, a virus infecting Euphorbia milli, which, like cassava in the genus Manihot, belongs to the plant family Euphorbiaceae. The function of this gene in CBSV and EuRV is unclear, but it can be speculated that Ham1, which belongs to the superfamily of inosine triphosphate pyrophosphatases, might act in preventing excessive RNA mutation, as suggested by Mbanzibwa et al. (2009a). This can be a significant feature for viruses infecting vegetatively propagated, perennial crops with continuous virus replication. Whilst the function of Ham1 for CBSV and also for EuRV requires further investigation, the high genetic distance among Ham1 sequences of CBSV clade 1 and clade 2 isolates provides evidence for independent gene lineages.

P1 varied considerably between clade 1 and clade 2 isolates but also within clade 2 viruses. Striking differences were found in P1 genes as well as in symptoms induced in N. benthamiana. Clade 2 isolates from Mozambique caused severe necrosis and eventually killed N. benthamiana. CBSV Tan_70 never killed the plant, although it caused severe leaf curling and systemic necrosis. In transient assays, P1 of CBSV MLB3 was determined as the viral suppressor of RNA silencing (Mbanzibwa et al., 2009a). The diversity of P1 genes might suggest differences in RNA silencing suppression activity and pathogenicity. This and the role of P1 in symptom expression and pathogenesis in cassava warrants further gene function studies.

In our study on cassava cultivar response, we showed that all lines tested were susceptible to CBSV clade 2 viruses,
whilst several cassava cultivars widely grown in Tanzania and throughout Africa were resistant to clade 1 virus isolates. This provides further support for species differentiation within CBSV, and potential epidemiological and biological differences.

Considering the criteria for species demarcation within the genus Ipomovirus for which only few virus species are known, and with the type species SPMMV possibly reflecting a different taxon (Li et al., 2008), guidelines and thresholds for species demarcation of members of the family Potyviridae (Adams et al., 2005b; Fauquet et al., 2005) can only be followed intuitively for CBSV. Our data support the separation of CBSV into two species. We propose retaining the species name Cassava brown streak virus for clade 1 isolates similar to the CBSV MLB 3 genome described by Mbanzibwa (2009a). Clade 2 viruses with genomes related to isolates from Mozambique should be assigned to a distinct virus species for which we propose the name Cassava brown streak Mozambique virus. The clade 2 virus Tan_70 from Kibaha with its deviating P1 genome described by Mbanzibwa (2005) can only be followed intuitively for CBSV. Our data support the separation of CBSV into two species. We propose retaining the species name Cassava brown streak virus for clade 1 isolates similar to the CBSV MLB 3 genome described by Mbanzibwa et al. (2009a). Clade 2 viruses with genomes related to isolates from Mozambique should be assigned to a distinct virus species for which we propose the name Cassava brown streak Mozambique virus. The clade 2 virus Tan_70 from Kibaha with its deviating P1 sequence and symptoms in indicator plants should therefore be considered a strain of cassava brown streak Mozambique virus (CBSMoV).

From the description of disease symptoms only, it is impossible to associate the history of CBSD occurrence and spread in East Africa with the incidence of a particular virus species. Hence, whilst CBSMoV can be considered to prevail in coastal lowland regions of Tanzania and Mozambique, earlier observations of CBSV at altitudes above 1000 m (Alicai et al., 2007; Hillocks & Jennings, 2003) cannot unequivocally be assigned to highland isolates of CBSV (Mbanzibwa et al., 2009b).

CBSV in our study was first isolated in 1997 and later in 1999 from cassava collected in the coastal regions of Kenya. This region is a secluded geographically isolated area of cassava cultivation, well separated from western Kenyan regions of intensive cassava cultivation. At the time of sample collection, CBSV was found only in the coastal districts of Kenya and not in western Kenya (Were et al., 2004a, b) or in Uganda (Alicai et al., 2007). Similarly, in the sequences of virus isolates collected from Tanzania and Mozambique analysed by Monger et al. (2001a), CBSV was not evident, and only sequences related to CBSMoV were described. CBSV reported from Uganda is highly similar to CBSV isolates from the Lake Victoria basin (Mbanzibwa et al., 2009b) and to CBSV from Malawi (this study). Outbreaks of CBSD in these areas, as already indicated by Mbanzibwa et al. (2009b), are thus independent from strains endemic in coastal regions of East Africa. However, the occurrence of a CBSMoV isolate (CBSV Tan Z) in the Musoma region of Lake Victoria, suggests that CBSV and CBSMoV populations may not remain geographically separated in the future.

**METHODS**

**Virus sources and cassava infections.** CBSV was initially isolated from cassava plants collected in the coastal regions of Kenya during studies to assess the diversity of cassava mosaic geminiviruses (Were et al., 2004a). In the course of the present study, cuttings of cassava plants with suspected symptoms of CBSV were obtained from field locations in Uganda, Tanzania, Malawi and Mozambique (Table 2). Characteristic symptoms of CBSV included typical 'feathery' chlorosis and yellow patch symptoms along secondary and tertiary veins of older leaves of cassava.

Cuttings were rooted and held under quarantine for several weeks to monitor the presence of pests and diseases other than viruses. CBSV was transmitted from symptomatic cassava leaves by mechanical inoculation to a series of herbaceous hosts including Nicotiana glutinosa, Petunia hybrida, N. benthamiana, Nicotiana hesperis, Nicotiana occidentalis and Nicotiana occidentalis ssp. hesperis.

To fulfil Koch's postulates, transmission tests to cassava were conducted by inoculating leaf homogenates from CBSV-infected N. benthamiana plants to fully expanded leaves of rooted cuttings of cassava land races and breeding lines of the International Institute of Tropical Agriculture (Nigeria). To investigate the genotype response to CBSV, scions of TMS 96/0304 plants that were systemically infected with each of the geographical isolates were side grafted onto uninfected cassava plants for virus transmission. For these experiments, cassava lines TME 1, TME 3, TME 4, TME 117, TMS 96/0160, TMS 96/0529 and TMS 96/1089A and cassava cultivars Albert and Kibaha, popular in Tanzania, and TMS 30572, an elite cultivar widely adopted in Africa (Raji et al., 2008), were used. The presence of typical CBSV symptoms on cassava leaves, CBSV confirmation by RT-PCR using specific primers and reisolation to N. benthamiana were used to confirm virus infection.

**Virus isolation, RNA purification, cDNA synthesis and cloning.** To enrich for virus particles, symptomatic leaves of CBSV-infected N.

---

**Table 2. Origin of CBSV isolates, site and date of collection, and GenBank accession numbers of genome sequences**

<table>
<thead>
<tr>
<th>CBSV isolate</th>
<th>DSMZ accession no.</th>
<th>Collection site/date</th>
<th>Location</th>
<th>Altitude (m)</th>
<th>Cassava cultivar</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ke_125</td>
<td>PV-0912</td>
<td>Kilifi, Kenya/1999</td>
<td>Kenya coast</td>
<td>20</td>
<td>Land race</td>
<td>FN433930</td>
</tr>
<tr>
<td>Ke_54</td>
<td>PV-0910</td>
<td>Kilifi, Kenya/1997</td>
<td>Kenya coast</td>
<td>20</td>
<td>Land race</td>
<td>FN433931</td>
</tr>
<tr>
<td>Ug_23</td>
<td>PV-0924</td>
<td>Namulonge, Uganda/2007</td>
<td>Lake Victoria</td>
<td>1200</td>
<td>TME 204</td>
<td>FN434109</td>
</tr>
<tr>
<td>Ma_42</td>
<td>PV-0927</td>
<td>Chitedze, Malawi/2007</td>
<td>Central Malawi</td>
<td>1100</td>
<td>CH03/044</td>
<td>FN433932</td>
</tr>
<tr>
<td>Ma_43</td>
<td>Pw-0925</td>
<td>Salima, Malawi/2007</td>
<td>Lake Nyasa</td>
<td>500</td>
<td>TME 6</td>
<td>FN433933</td>
</tr>
<tr>
<td>Mo_83</td>
<td>PV-0949</td>
<td>Namacurra, Mozambique/2007</td>
<td>Mozambique coast</td>
<td>20</td>
<td>Mulaleia</td>
<td>FN434436</td>
</tr>
<tr>
<td>Tan_70</td>
<td>PV-0957</td>
<td>Mlandizi, Tanzania/2008</td>
<td>Tanzania coast</td>
<td>60</td>
<td>Land race</td>
<td>FN434437</td>
</tr>
</tbody>
</table>
benthamiana plants were harvested 10–16 days post-inoculation, homogenized in Tris/borate buffer (pH 7.8) with 0.1 % thioglycolic acid and filtered through layers of muslin. Chloroform was added (1:2, v/v) and, after phase separation, the aqueous phase was centrifuged for 2 h at 45 000 g and 4 °C. Pellets were suspended in proteinase K buffer and RNA was extracted using an RNAeasy clean-up protocol (Qiagen) followed by DNase I treatment. To generate the first sequence information, a series of primers was developed from CBSV sequences available in GenBank. From the initial sequence thus obtained, further genome portions were amplified with primers for cDNA synthesis derived from newly analysed genome sequences and following a 5’ rapid amplification of cDNA ends strategy (SMART RACE; Clontech Laboratories). Final 5’ genome sequences of the viral RNA genome and the exact determination of the starting nucleotide of the virus genome were obtained by rolling-circle amplification on CircLigase (Epigenome Technologies)-circularized cDNA (Polidoro et al., 2006) using Φ29 DNA polymerase (New England Biolabs) followed by inverse PCR.

To assemble further CBSV genomes, RNA from virus-infected N. benthamiana plants was isolated using an RNAeasy plant RNA extraction kit (Qiagen) at various time points during infection. RT-PCR was subsequently performed with specific primers designed for each isolate using Thermoscript (InVitrogen) or TaqPlus Precision PCR (Stratagene). All PCR products were purified using a Qiagen gel extraction kit, ligated into the pDrive U/A cloning vector (Qiagen) and subsequently electroporated into Escherichia coli DH5α cells.

Sequencing, assembly of genomes and phylogenetic analysis. Recombinant cDNA clones were selected and sequenced using T7/SP6 primers or internal sequence-specific primers. For each position, two to three independent cDNA clones were sequenced using the services of a commercial company (SEQserve, Germany). Analysis of the nucleotide sequence, annotation, contig assembly to generate full-length genomes and analysis of the coding regions were carried out with Vector NTI Advance 10.01 software (Invitrogen). Putative CBSV polypolyprotein cleavage sites (Adams et al., 2005a) and specifically those of CVYV (Janssen et al., 2005). For evaluation of phylogenetic relationships, CLUSTAL W (Thompson et al., 1994) was used for alignment of CBSV genome sequences and to compare the amino acid sequences of mature CBSV proteins. The complete genome sequence of the CBSV highland isolate MLB3 (Mbanzibwa et al., 2009a) and of a CBSV isolate from Musoma, Lake Victoria (CBSV Tan Z, GenBank accession no. GQ329864), were included in the analysis. Homologous genes of CVYV (GenBank accession no. AY578085), SqVVY (GenBank accession no. EU259611) and SPMVV (GenBank accession no. NC_003797) were used for reference. The evolutionary history was inferred using the minimum evolution method (Zhetsky & Nee, 1992) and the phylogenetic analyses were conducted in MEGA4 (Tamura et al., 2007).

Antiserum production and Western blot analysis. The sequence of the predicted CP gene of CBSV Ke_125 was amplified by RT-PCR from the RNA of CBSV-infected N. benthamiana plants using primers for insertion into the prokaryotic expression vector pET28a (+) (Novagen) carrying an N-terminal His-tag configuration for protein purification, to construct a recombinant His–CP fusion expression vector. This was mobilized into E. coli BL21-CodonPlus (DE3)-RIL competent cells (Agilent) and the cells were grown and induced for protein expression. As the greater part of the expressed protein was in the insoluble cytoplasmic fraction, a protein purification protocol was followed to extract and solubilize inclusion bodies under denaturing conditions (8 M urea). His–CP fusions were affinity purified on NInfTA agarose columns, sequentially reducing the urea concentration to elute the soluble protein fraction under native conditions.

Recombinant CP was used for immunization of New Zealand white rabbits following a protocol of intravenous and intramuscular injections at 2-week intervals. After a last booster injection, blood was taken from the lateral ear vein and IgG was purified and tested for specificity in Western blot analysis essentially as described previously (Dizadji et al., 2008). Membranes were incubated with CBSV IgG (DSMZ AS-0912) at 2 μg ml⁻¹ followed by immunodetection of the bound CBSV IgG with a goat anti-rabbit alkaline phosphatase conjugate (Dako) diluted 1:10 000. To determine serological relationships with other viruses of the genus Ipomovirus, plant materials infected with CVYY (DSMZ PV-0724) and SPMVV (DSMZ PV-0900) were included in the study.

ACKNOWLEDGEMENTS

We are grateful to Drs Ivan Ingelbrecht and Alfred Dixon for inspiring discussions and enthusiastic and critical support of this work. Drs Pheneas Ntawuruhunga, Ibrahim Benesi, Maria Andrade and Mathew Abang are herewith thanked for provision of cassava materials, and Sharon Cheek for critical reading of this manuscript.

REFERENCES


