Plant Disease • 2019 • 103:818-824 • https://doi.org/10.1094/PDIS-08-18-1445-RE

e-Xtra*

Prevalence and Genome Characterization of Field Isolates of Sugarcane Mosaic Virus (SCMV) in Nigeria

Adama Yahaya,[1] Danladi B. Dangora,[1] P. Lava Kumar,[2] Matthew D. Alegbejo,[3] Lori Gregg,[4] and Olufemi J. Alabi[4,†]

[1]Department of Botany, Ahmadu Bello University, Zaria, Kaduna State, Nigeria; [2]International Institute of Tropical Agriculture (IITA), PMB 5320, Ibadan, Nigeria; [3]Department of Crop Protection, Ahmadu Bello University, Zaria, Kaduna State, Nigeria; and [4]Department of Plant Pathology & Microbiology, Texas A&M AgriLife Research & Extension Center, Weslaco, TX 78596, USA

Abstract

Maize and sugarcane are two economically important crops often grown in adjacent fields or co-cultivated in the northern guinea savannah agroecological zone, a major cereal production region of Nigeria. This study was conducted to determine the prevalence of mosaic disease in sugarcane and maize fields in the northern guinea savannah agroecological zone and to molecularly characterize the associated sugarcane mosaic virus (SCMV, genus *Potyvirus*) isolates. Surveys were conducted from June to July 2015, and sugarcane mosaic disease (SCMD) incidence was assessed across 21 farmer's fields. Mean SCMD incidence varied across states with ~82% (308/376), ~66% (143/218), and ~67% (36/54) recorded in Kaduna, Kano, and Katsina states, respectively. RT-PCR analysis of 415 field-collected samples using genus-specific primers confirmed potyvirus infection in 63.7% (156/245) of sugarcane, 29.7% (42/141) of maize crops, and 45% (13/29) of itch grass samples. Cloning and sequencing of gene-specific DNA amplicons from a subset of 45 samples (sugarcane = 33, maize = 9, itch grass = 3) confirmed their specificities to SCMV. Phylogenetic analysis of the partial gene sequences showed that they all belong to a single monophyletic clade of SCMV. These results were supported by analysis of complete polyprotein sequences of representative maize and sugarcane isolates from Nigeria. Both isolates shared 94.9%/97.3% complete polyprotein nucleotide (nt)/amino acid (aa) identities with each other and 75.2%/97.6% nt/aa identities with corresponding sequences of global SCMV isolates. The detection of identical populations of SCMV isolates in both crop species and a weed host suggests possible vector mediated interspecies spread within cereal landscapes in the study area with implications for the integrated and sustainable management of SCMD in cereal cropping systems in Nigeria.

Maize (*Zea mays* L.) and sugarcane (*Saccharum officinarum* L.) (family Poaceae) are two economically important crops in Nigeria. The former is grown primarily as a basic staple food crop for human and livestock consumption, while the latter is produced as both a food (chewing cane) and cash crop (sugarcane) in the country. Although both crops are cultivated across all seven agroecological zones of Nigeria, their production occurs predominantly in the northern guinea savannah zone of the country encompassing states such as Kaduna, Kano, and Katsina (Iken and Amusa 2004; USDA 2010). Both sugarcane and maize play very important roles in the livelihood of rural farmers, serving as main income sources.

Maize and sugarcane are often cultivated in adjacent fields, in mixed cropping systems, or as rotation crops, thus providing opportunities for the perpetuation of common viruses infecting both crops. Several diseases, including mosaic and streak disease caused by viruses, affect sugarcane and maize production, but studies on the estimation of yield losses due to mosaic-inducing viruses infecting both crops are generally lacking in Nigeria. However, studies elsewhere have shown that *Potyvirus*-associated mosaic disease may cause serious yield losses in maize and sugarcane crops (CABI 2018). Hence, efforts by national and international research institutes with mandates for both crops in Nigeria have focused on understanding the epidemiology of virus diseases and the development of control strategies.

Among several viruses documented from maize and sugarcane worldwide, the most economically important in Africa are causative agents of maize streak and sugarcane mosaic diseases. However, previous studies of viral diseases of both crops in Nigeria have focused largely on the widely prevalent maize streak disease, caused by maize streak virus (MSV; genus *Mastrevirus*; family *Geminiviridae*) (Gordon and Thottappilly 2003; Oluwafemi et al. 2014). Besides MSV, other so-called ''African streak viruses'' (AfSVs), including *Sugarcane streak Reunion virus*, *Urochloa streak virus*, *Maize streak Reunion virus*, *Axonopus compressus streak virus*, and *Sugarcane chlorotic streak virus*, have also been documented in maize plants and other grass species showing streak symptoms (Oluwafemi et al. 2014; Yahaya et al. 2017). These AfSVs are currently delimited to the continent of Africa and its adjoining islands. In a recent study, the maize-adapted strain MSV-A was detected in sugarcane under field conditions in Nigeria (Yahaya et al. 2017), suggesting that the sugarcane crop could be epidemiologically important as a reservoir host for MSV.

Unlike maize streak disease, sugarcane mosaic disease (SCMD) is widespread across all cereal-producing regions of the world (Silva-Rosales et al. 2015). SCMD has been implicated in significant yield losses of tons of cane and sucrose per hectare (Yao et al. 2017), with most sugarcane genotypes being susceptible to the disease (Addy et al. 2017; Thorat et al. 2015). SCMD also causes reduced grain and forage yield of affected plants, and susceptible cultivars that become infected early in the season may be barren (Leng et al. 2015). The primary causative agent of SCMD is sugarcane mosaic virus (SCMV; genus *Potyvirus*; family *Potyviridae*) (Silva-Rosales et al. 2015 and references therein). Other SCMD-causal viruses include *Maize dwarf mosaic virus*, *Johnsongrass mosaic virus*, *Sorghum mosaic virus*, *Zea mosaic virus*, *Pennisetum mosaic virus*, and *Cocksfoot strike virus* (Wu et al. 2012). SCMV and some of the SCMD-causal viruses may also interact synergistically with *Maize chlorotic mottle virus* (genus *Machlomovirus*; family *Tombusviridae*) to cause maize lethal necrosis disease (synonyms corn lethal necrosis disease), an emerging debilitating disease of maize (Niblett and Claflin 1978; Wangai et al. 2012).

†Corresponding author: O. J. Alabi; E-mail: alabi@tamu.edu

Funding: This research was supported by Norman E. Borlaug Leadership Enhancement in Agriculture Program (Borlaug-LEAP) funded by the USAID, project no. 016258-121; the Tertiary Education Trust Fund (TETFUND), Ahmadu Bello University, Zaria, Kaduna State, Nigeria; the CGIAR Research Program on MAIZE; and the Amina Ado Foundation.

*The *e-Xtra* logo stands for "electronic extra" and indicates that three supplementary figures and one supplementary table are published online.

Accepted for publication 25 September 2018.

The occurrence of SCMV in Nigeria has been serologically demonstrated (Dangora et al. 2014; Muhammad et al. 2016; Yahaya et al. 2014), but information is lacking on the molecular characteristics of field isolates of the virus. Studies from different cereal-producing regions of the world have shown that SCMV exists as genetic variants belonging to distinct phylogroups (Alegria et al. 2003; Gao et al. 2011; Gemechu et al. 2005; Ha et al. 2008; Xu et al. 2008). Hence, it is important to further characterize field isolates of SCMV from Nigeria to determine their genetic relationship to global isolates of the virus. In addition, given the mixed cropping system nature of the Nigerian northern guinea savannah agroecological zone, a comparative evaluation of SCMV isolates derived from sugarcane and maize may provide epidemiological insights pertinent to developing ecologically relevant control measures to tackle SCMD.

This study was conducted to determine the prevalence of mosaic disease in sugarcane and maize fields in the northern guinea savannah agroecological zone, to molecularly characterize field isolates of the associated SCMV, and to determine the genetic relationships between isolates of SCMV from Nigeria and global isolates of the virus.

Materials and Methods

Field incidence of SCMD and provenance of samples. Surveys were conducted from June to July 2015 in 21 farmer's fields across three leading cereal producing states of Kaduna, Katsina, and Kano (Supplementary Figure S1). The survey sites were selected to include farms that were frequently cultivated with sugarcane and/or maize. SCMD incidence was assessed in each field on 30 to 32 plants that were chosen on each of the two longest opposite sides of the field and a diagonal, after every 20 to 30 steps (depending on the size of the field) (Barnett 1986). Disease incidence per field was determined as number of plants showing typical mosaic symptoms (Fig. 1) divided by number of plants assessed, expressed as a percentage. However, leaf tissues were sampled from 18 to 19 randomly selected plants per field (sugarcane = 245; maize = 141) for laboratory analysis. An additional 29 samples were collected from itch grass (*Rottboellia cochinchinensis*) growing within or in proximity to the assessed fields due to virus-like symptoms observed on some of the plants (Fig. 1B). This resulted in the collection of a total of 415 samples. Each sample was preserved dry on $CaCl_2$ at room temperature to be analyzed later.

Preliminary screening. Total RNA was extracted from each of the 415 dried leaf samples using a modified CTAB method (Silva et al. 2015) at the Virology and Molecular Diagnostics Laboratory of the International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria. The RNA was suspended in 50 µl of RNase-free water and stored at −20°C until use. Two microliters of 100 ng/µl RNA per sample was used in a 12.5-µl one-step RT-PCR assay containing 0.2 µM of each universal potyvirid primers CIFor and CIRev (Ha et al. 2008) targeting partial cylindrical inclusion (CI) gene of

Fig. 1. Virus-like symptoms observed on leaves of (**A**) sugarcane, (**B**) itch grass (*Rottboellia cochinchinensis*), and (**C**) maize in the northern guinea savannah agroecological zone of Nigeria.

potyviruses (Table 1), 0.2 mM of each dNTP, 1 U GoTaq DNA Polymerase (Promega, USA), 12 U M-MLV reverse transcription (Promega, USA), and 1× GoTaq Reaction Buffer (plus 1.5 mM $MgCl_2$). The thermal cycle conditions were 42°C for 30 min, 35 cycles of 95°C for 5 min, and 35 cycles of 95°C for 30 s, 40°C for 30 s, and 72°C for 45 s, and a final extension step of 72°C for 10 min. The obtained DNA products were resolved in a 1.2% agarose gel in Tris-Acetate-EDTA (TAE) buffer (pH 8.3) containing 0.05 μl/ml EZ-Vision Blue Light DNA Dye (Amresco, USA), and visualized under UV light using a Gel Doc EZ Imager (Bio-Rad, France). Samples were scored as either potyvirus-positive or -negative based on an appearance of the expected ~700 bp DNA fragment.

DNA fingerprinting of potyvirus-positive samples. Based on the results of the initial screening, dried leaf tissues from a subset of 45 samples (sugarcane = 33, maize = 9, and itch grass = 3) were shipped under permit from the U.S. Department of Agriculture-Agricultural Research Service Plant Protection Quarantine (P526P-14-04321) to the Texas A&M University (TAMU) AgriLife Research and Extension Center facility in Weslaco, TX for further analysis. The 45 samples were selected such that they were representative of the different plant species and field locations. Total nucleic acids (TNA) was extracted from each dried leaf sample (Dellaporta et al. 1983) and used as template for cDNA synthesis using random hexamers, with the ThermoScript RT-PCR System (Thermo Fisher Scientific, Waltham, MA, USA) as per manufacturer's instructions. Approximately 400 ng/μl cDNA aliquots per sample were subjected to PCR using three pairs of degenerate primers. The primer pairs HPFor and HPRev, and CIFor and CIRev (Ha et al. 2008) each target a ~700 bp DNA fragment specific to the helper component protease (HC-Pro) and cylindrical inclusion (CI) genes of potyviruses, respectively, while primers NIb2F and NIb2R (Zheng et al. 2010) target a ~350 bp fragment of the nuclear inclusion b (NIb) gene of most potyviruses (Table 1, Supplementary Figure S2). Each 25-μl reaction mixture contained 1× buffer with $MgCl_2$ (Roche Diagnostics, Indianapolis, IN), 0.4 μM each forward and reverse primers, 0.2 mM dNTPs, 5 U Taq polymerase (Roche Diagnostics, Indianapolis, IN), and 2 μl cDNA aliquot. Amplicons of expected sizes were verified by resolving 10 μl PCR product aliquots, along with the 1 kb Plus DNA ladder (Thermo Fisher Scientific), on 1.5% agarose gels prestained with ethidium bromide. The gels were visualized under a UV-transilluminator, and the gene-specific amplicons were gel-eluted (Zymoclean Kit; Zymo Research, Irvine, CA). Each eluate was ligated into pCR2.1 TOPO-TA vector (Thermo Fisher Scientific) following the manufacturer's instructions. The ligation products were used to transform One Shot TOP10 chemically competent *Escherichia coli* cells (Thermo Fisher Scientific), following which plasmid DNA was purified from positive recombinant clones with the GenElute Plasmid Miniprep Kit (Sigma-Aldrich, MO, USA). Two to three independent clones were sequenced per isolate in both orientations for each gene-specific DNA fragment using M13F/R primers. A consensus sequence was derived when all isolate-specific independent clones from each genomic region showed ≥99% nucleotide identities after pairwise comparisons were performed with the BioEdit Sequence Alignment Editor version 7.2.5 (Hall 1999). The aligned gene-specific amino acid sequences were phylogenetically analyzed along with corresponding sequences of global SCMV isolates using the MEGA7 program (Kumar et al. 2016).

Complete genome characterization of representative maize and sugarcane SCMV isolates from Nigeria. To further characterize SCMV isolates from Nigeria, one representative virus isolate from sugarcane and maize each were selected based on results obtained from the analyses of partial gene fragments which showed very high sequence identities among SCMV isolates from Nigeria. Complementary DNA samples from each of the two isolates were subjected to PCR to amplify the near complete genome of SCMV as four pieces of overlapping DNA fragments using virus-specific primers (Table 1). The 25-μl PCR reaction volume for each reaction consisted of 1× PrimeSTAR GXL buffer, 0.2 mM dNTPs, 0.2 μM each primer, 1.5 U PrimeSTAR GXL DNA polymerase (Takara-Bio USA, Inc., Mountain View, CA, USA) and 2 μl of a 1:10 dilution of the template cDNA. Thermal cycling conditions were 35 cycles of 98°C for 10 s, annealing 55°C for 15 s, and extension 68°C for 10 s/kb. The correct size amplicons on a prestained (10 mg/ml ethidium bromide) 1% agarose gel were excised and gel-eluted as described above. Each eluate was A-tailed before cloning as previously described (Alabi et al. 2016). Three plasmids with the correct size inserts per DNA amplicon were isolated and initially sequenced in both directions with the M13F/R primers as described above. Additional primers were used to walk each plasmid DNA sample (data not shown) until the entire sequence fragment was completed. A consensus sequence was obtained for each PCR fragment, and the putative near-full-length viral genome was assembled from all four genome fragments using the CAP contig assembly program of the BioEdit software (Hall 1999).

Sequence analysis, estimation of genetic distance, and test for recombination. Multiple alignments were performed for the partial gene sequences and the complete polyprotein nucleotide (nt) and amino acid (aa) sequences with the program MUSCLE (Edgar 2004; https://www.ebi.ac.uk/Tools/msa/muscle/), with default settings along with corresponding sequences of 88 global SCMV isolates. The complete polyprotein alignment files were used to generate sequence identity matrices using the BioEdit program and for reconstruction of the phylogeny of global SCMV sequences with the MEGA7 program (Kumar et al. 2016) using the best fit substitution model as described previously (Alabi et al. 2014). The estimates of evolutionary divergences over sequence pairs (d) were calculated for

Table 1. Name, sequence, and genome position of oligonucleotides used for the amplification of partial and complete genome fragments of sugarcane mosaic virus (SCMV) isolates analyzed in this study and their expected product sizes

Primer name	Sequence (5′ – 3′)[a]	Genome position[b]	Size (bp)	Reference
HPFor	TGYGAYAAYCARYTIGAYIIIAAYG	1441–1465	683	Ha et al. 2008
HPRev	GAICCRWAIGARTCIAIIACRTG	2123–2101		
CIFor	GGIVVIGTIGGIWSIGGIAARTCIAC	3730–3755	683	Ha et al. 2008
CIRev	ACICCRTTYTCDATDATRTTIGTIGC	4412–4387		
NIb2F	GTITGYGTIGAYGAYTTYAAYAA	7441–7463	350	Zheng et al. 2010
NIb2R	TCIACIACIGTIGAIGGYTGNCC	7790–7768		
SCMV-63F	CAGATTGTAGTGAACGGCTCG	63–83	2,769	Gao et al. 2011
SCMV-2831R	ACAGCGCCTAAATCTACG	2831–2814		
SCMV-2711F	ACCARTCATGGGCAGAWTTATC	2711–2732	2,640	Gao et al. 2011; modified
SCMV-5329R	CCATGAGTTCTGACGGATCTTG	5350–5329		
SCMV-5209F	AATGGGTTGGTRTCCATGAT	5209–5228	2,350	Gao et al. 2011; modified
SCMV-7558R	CTTCCGGCAACTTCTCAAGTA	7558–7538		
SCMV-7408F	GCAGCTCCATTGGAAACTTTG	7408–7428	2,167	Gao et al. 2011; modified
SCMV-9574R	TGTCTCTTACCAGGAGACTCG	9574–9554		

[a] N = A+C+G+T; V = A+C+G; R = A+G; W = A+T; Y = C+T; M = A+C; K = G+T; H = A+T+C; D = G+A+T; S = G+C; I = deoxyinosine.
[b] Nucleotide position based on SCMV isolate Brisbane (AJ278405) from Australia.

the entire polyprotein amino acid dataset and for the within and between identified phylogroups using the JTT matrix-based model (Jones et al. 1992), with rate variation among sites modeled with a gamma distribution (shape parameter = 0.6) in MEGA7 (Kumar et al. 2016). The complete polyprotein nucleotide sequence alignment file was scanned for putative recombination events involving SCMV isolates from Nigeria since recombination has been reported for other isolates of the virus (Moradi et al. 2016; Xie et al. 2016). The suite of programs implemented in the RDP v.4.56 software (Martin et al. 2005) was used to scan the complete polyprotein alignment file, with Bonferroni correction. Default settings were used for each RDP analysis with the exceptions that the linear genome option was selected. Only events detected by at least five of the seven detection methods, with strong statistical support ($P \leq 0.05$) were considered.

Results

SCMD prevalence. Typical mosaic symptoms were observed on leaves of sugarcane (Fig. 1A), itch grass (Fig. 1B), and maize (Fig. 1C) plants encountered in farmer's fields during the survey. Based on visual assessment of symptoms, mean disease incidence varied across states, with ~82% (308/376), ~66% (143/218), and ~67% (36/54) recorded for Kaduna, Kano, and Katsina states, respectively. Information gathered through informal interviews of farmers encountered during the survey indicated that, whereas there is a consensus that presence of mosaic symptoms on maize and sugarcane plants is linked to poor yield and quality attributes of both crops, many farmers attributed the mosaic patterns to varietal properties (~43% or 9/21) or nutritional imbalances (~10% or 2/21). The remaining 48% of farmers (10/21) have no opinion about the potential cause of the observed mosaic symptoms. Eleven of 21 surveyed fields (~52%) cultivated maize or sugarcane as monocrops, while the remaining 10 fields (~48%) practiced mixed cropping whereby sugarcane is grown in mixtures with maize and/or other vegetable crops such as tomato, peppers, or watermelon.

Potyvirus incidence and DNA fingerprinting of positive samples. The ~700-bp product specific to the CI gene of most potyviruses was successfully amplified by RT-PCR from ~51% (211/415) of samples, indicating presence of a *Potyvirus* in the samples. Further analysis revealed greater potyvirus incidence in sugarcane (~64% or 156/245) than in maize (~30% or 42/141) samples. Thirteen (~45%) of 29 itch grass samples collected during the survey also tested positive by RT-PCR. Based on these results, 45 samples (sugarcane = 33, maize = 9, and weed = 3) were advanced for further molecular analyses.

DNA band of the expected size was successfully amplified from each of the 45 samples with at least one of the three generic potyvirid primer pairs (Table 1) in agreement with results of the initial screening. However, the three gene-specific primers varied in their relative performance with the HC-Pro-specific primers yielding DNA band of the expected size in ~76% (34/45) of samples, followed by the NIb-specific primers (~71% or 32/45) and the CI-specific primers (~42% or 19/45). BLASTN analysis of the derived gene-specific sequences confirmed their specificities to SCMV. The sequences have

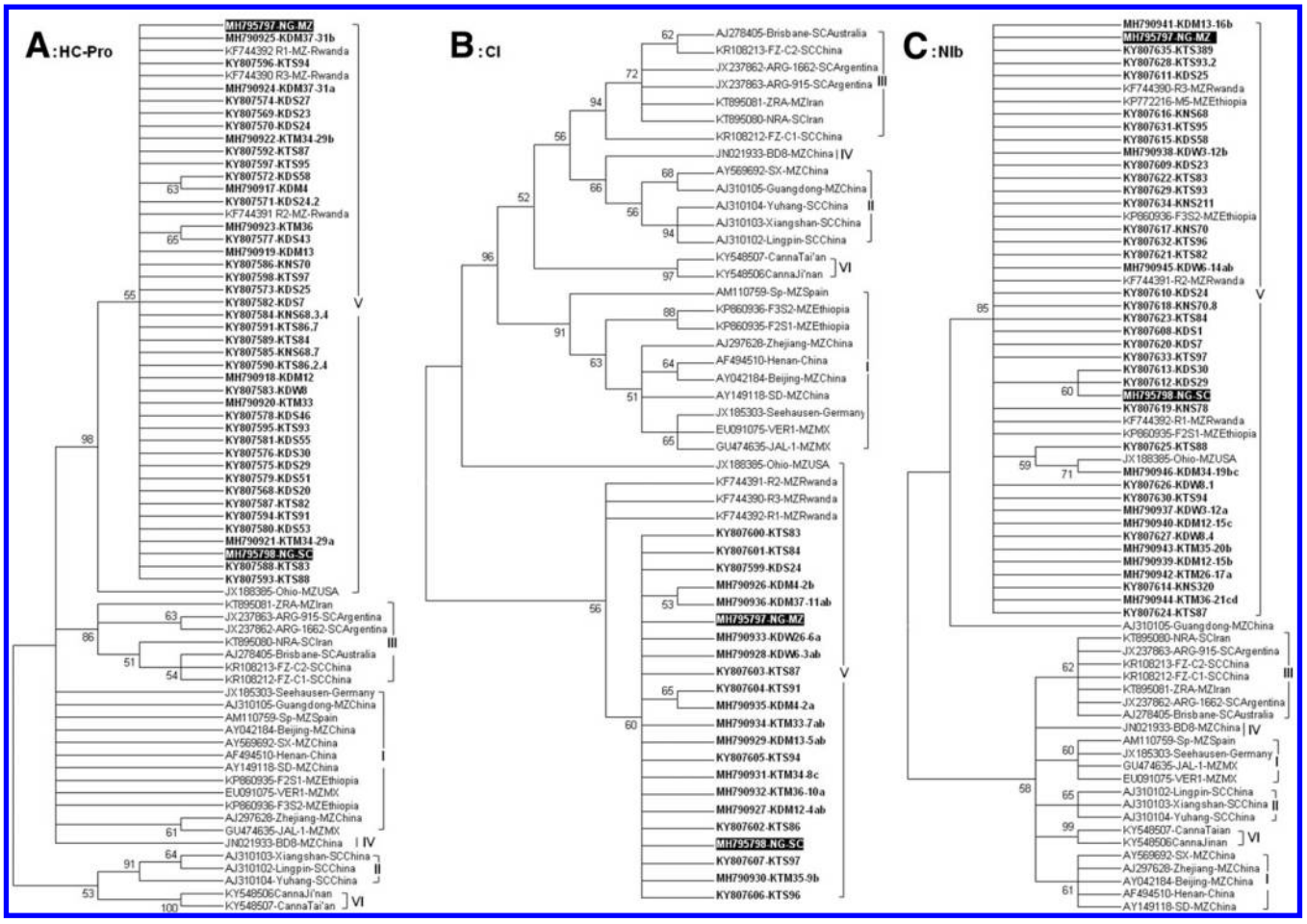


Fig. 2. Unrooted neighbor-joining phylogenetic tree depicting evolutionary relationships among global isolates of sugarcane mosaic virus (SCMV) based on analysis of partial amino acid sequences specific to **A,** helper component protease (HC-Pro), **B,** cylindrical inclusion (CI) protein, and **C,** nuclear inclusion B (NIb) protein. The analysis was performed with the MEGA7 program (Kumar et al. 2016) using the Jones-Taylor-Thornton (JTT) evolutionary model (Jones et al. 1992) that best-fitted the sequence dataset. Bootstrap values (1,000 replicates) are shown at the branch nodes and branches corresponding to partitions reproduced in <50% of bootstrap replicates were collapsed. The GenBank accession number of each isolate is followed by the taxon ID, country of origin, and host plant (M or MZ = maize; S or SC = sugarcane). Sequences obtained in this study are in bold fonts.

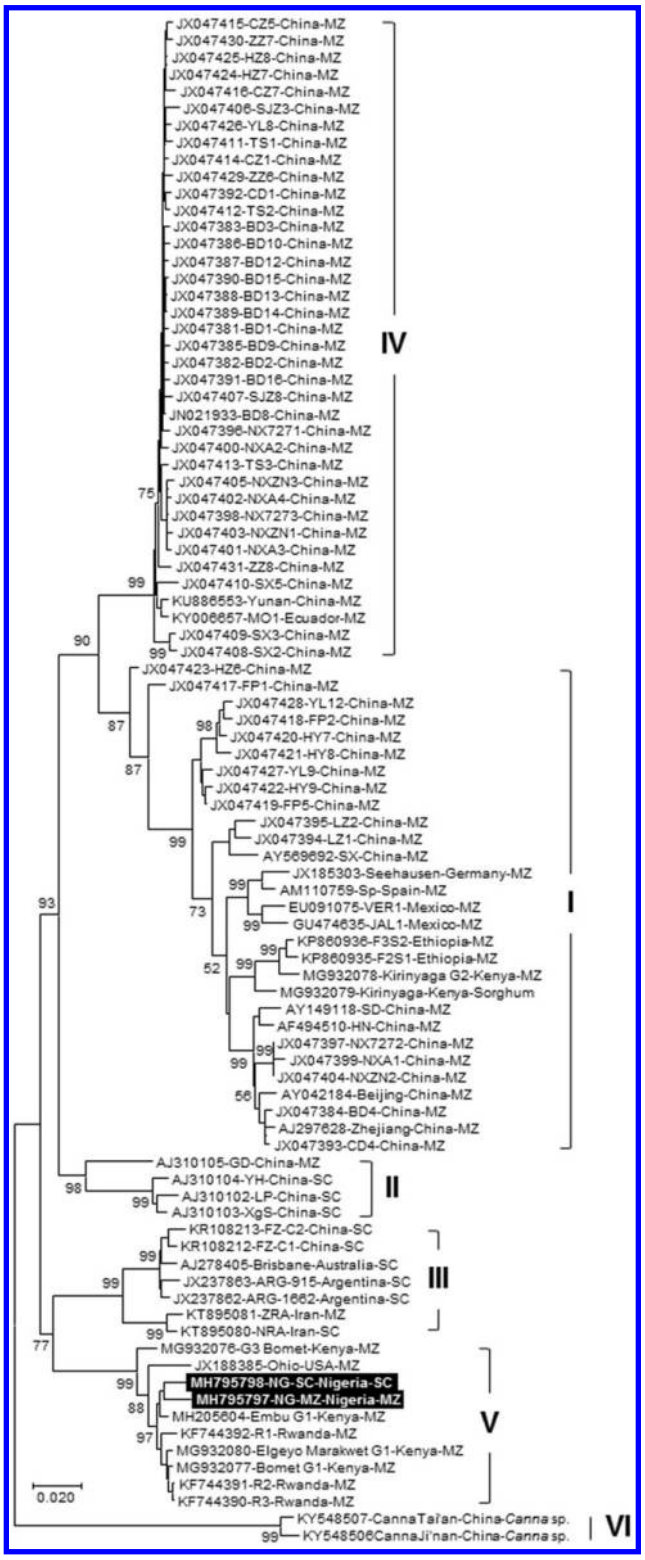


Fig. 3. Unrooted neighbor-joining phylogenetic tree depicting evolutionary relationships among global isolates of sugarcane mosaic virus (SCMV). The tree was constructed using aligned complete polyprotein amino acid sequences of 90 SCMV isolates (This Study = 2; GenBank = 88). The phylogenetic analysis was performed with the MEGA7 program (Kumar et al. 2016) using the Jones-Taylor-Thornton (JTT) evolutionary model (Jones et al. 1992) that best-fitted the sequence dataset. Bootstrap values (1,000 replicates) are shown at the branch nodes and branches corresponding to partitions reproduced in <50% of bootstrap replicates were collapsed. The GenBank accession number of each isolate is followed by the taxon ID, country of origin, and host plant (MZ = maize; SC = sugarcane). Sequences obtained in this study are highlighted.

been deposited in GenBank under the accession numbers KY807568-598 and MH790917-925 (partial HC-Pro), KY807599-607 and MH790926-936 (partial CI), and KY807608-635 and MH790937-946 (partial NIb). In pairwise comparisons, the partial HC-Pro-specific sequences derived in this study were 94 to 100% and 97 to 100% identical to each other at the nt and aa levels, respectively. The partial CI-specific sequences derived in this study were 94 to 100% and 98 to 100% identical to each other at the nt and aa levels, respectively. The partial NIb-specific sequences derived in this study were 92 to 100% and 96 to 100% identical to each other at the nt and aa levels, respectively. All partial gene-specific SCMV sequences derived in this study clustered into one of the five phylogroups (group V) reported earlier (Gao et al. 2011; Wang et al. 2010) (Fig. 2). We detected an additional phylogroup VI formed by two newly characterized SCMV isolates derived from *Canna* sp. (Fig. 2). Within group V, the maize, sugarcane, or itch grass SCMV sequences from Nigeria were intermingled regardless of the gene under consideration (Fig. 2). These results suggest that SCMV isolates from Nigeria are genetically identical regardless of their crop host or sample location. Consequently, one isolate each from maize and sugarcane were studied further via analyses of their near complete genome sequences.

Complete genome characterization of representative maize and sugarcane SCMV isolates from Nigeria. The near complete genome of SCMV was amplified as four pieces of overlapping DNA fragments from one sugarcane (isolate NG-SC) and one maize (isolate NG-MZ) sample. Size variations were observed in the near complete genome length of both isolates such that the length of NG-SC was determined to be 9,526 nt (GenBank accession no. MH795798), while that of NG-MZ was 9,558 nt (GenBank accession no. MH795797). Further analyses of both isolates revealed that each contain the expected single large open reading frame (ORF) with a putative ORF start at AUG (86-88 nt). However, whereas isolate NG-SC encodes a polyprotein of 3,067 amino acids (nt 86…9289) with an estimated molecular weight of 347.3 kDa, isolate NG-MZ has a slightly larger polyprotein of 3,078 amino acids (nt 86…9,322) with an estimated molecular weight of 347.9 kDa. The difference in genome length between the two isolates is due to the presence of an extra 11 amino acid residue "GAKPAASGAGS" within the carboxyl-terminus of the NIb protein of isolate NG-MZ. Both isolates NG-SC and NG-MZ are similar in seven of the nine putative protease cleavage sites based on comparison of their putative coding region with the consensus protease recognition motifs in other SCMV isolates (data not shown). Our attempts to verify the 5′ and 3′ extremities of both isolates with the SMARTer RACE 5′/3′ Kit failed perhaps due to suboptimal RNA template quality. Thus, both isolates are putatively missing 62 nucleotides in the 5′ untranslated region (UTR) and a poly(A) tail at their 3′UTR based on homology with complete SCMV isolates available in the databases.

Phylogenetic and recombination analysis of complete polyprotein. Analysis of global SCMV isolates using complete polyprotein amino acids sequences (This study = 2; GenBank = 88) revealed their segregation into six phylogroups (Fig. 3) similar to results obtained for the partial gene-specific sequences (Fig. 2). Support for the existence of the six phylogroups could be found in the greater between-group genetic distances (0.091 ± 0.006 to 0.184 ± 0.010) relative to within-group genetic distances (0.012 ± 0.002 to 0.043 ± 0.003). The overall genetic distance was estimated to be 0.081 ± 0.003 (Supplementary Figure S3).

Using a stringent threshold of events detected by at least five of the seven RDP4-implemented programs (Martin et al. 2005), a total of 20 acceptable putative recombination events were detected when the aligned complete polyprotein sequences of global SCMV isolates (This Study = 2; GenBank = 88) were scanned using the software (Supplementary Table S1). Notably, one of the putative recombination events involved maize isolates NG-MZ (MH795797) from Nigeria and Embu G1 (MH205604) from Kenya as putative recombinant sequences (Fig. 4). The event spanned the P1 and HC-Pro with the potential major and minor parental sequences being maize isolates Bomet G1 (MG932077) from Kenya and R3 (KF744390) from Rwanda (Fig. 4). This phylogroup V specific putative recombination

event was predicted by all the seven RDP4-implemented programs with very strong statistical support. Thirteen of the 20 acceptable putative recombination events were interphylogroup events, while seven involved isolates belonging to the same phylogroup. None of the putative recombination events involved the Nigerian sugarcane isolate NG-SC (MH795798).

Discussion

Previous studies on the prevalence of cereal-infecting viruses in Nigeria were limited to specific crop hosts and production areas (Dangora et al. 2014; Muhammad et al. 2016; Yahaya et al. 2014). In contrast, mixed cropping systems, in which maize (an annual crop) is cultivated in the same or adjacent fields with sugarcane (a perennial crop) are common in the cereal-producing northern guinea savannah zone of Nigeria. Thus, it is important to gain a better understanding of how field isolates of major viruses infecting multiple crops in the same landscape compare genetically with each other and the potential epidemiological significance of their genetic relatedness. We employed the concept of molecular epidemiology to unravel the genetic relatedness of SCMV isolates from maize, sugarcane, and a weed sampled from farmer's fields across the northern guinea savannah agroecological zone of Nigeria. The results show a high prevalence of sugarcane mosaic disease across the study area and the occurrence of SCMV in sugarcane, maize, and a weed host plant.

The results of this study confirmed previous reports of the serological detection of SCMV in sugarcane samples obtained from the same region (Dangora et al. 2014; Yahaya et al. 2014). The results also represent the first confirmed detection of SCMV in maize and *Rottboellia cochinchinensis* (itch grass), a common weed growing in cereal landscapes, in Nigeria. Interestingly, *R. cochinchinensis* was also recently reported as a host to *Maize yellow mosaic virus* (Yahaya et al. 2017), suggesting that this perennial grass species may be a common reservoir for multiple cereal-infecting viruses and could thus be playing an important role in the epidemiology of these viruses. The cosegregation of SCMV isolates derived from different plant species in the study area into a monophyletic clade would suggest cross-species movement of the virus by one or more competent aphid vector species. Given the perennial growth habits of sugarcane and itch grass, it is likely that both plant species are reservoir hosts of SCMV and possibly its aphid vector(s) for initiation of disease epidemics in maize crops during a given season. Further studies are needed to determine the relative ability of different aphid vectors of SCMV to acquire the virus from itch grass and transmit it to cultivated crops such as sugarcane and maize.

Analyses of partial gene-specific sequences derived from the HC-Pro, CI, and NIb proteins (Fig. 2) and the complete polyprotein sequences (Fig. 3) revealed the segregation of global SCMV isolates into six distinct phylogroups. These results agree to a large extent with previous studies that reported the segregation of global SCMV isolates into five well-defined phylogroups (Gao et al. 2011; Wang et al. 2010). However, the detection of an additional phylogroup VI consisting of two recently characterized SCMV genome sequences from *Canna* sp. in China (Figs. 2 and 3) suggests a more complex evolutionary history of SCMV than was previously realized. The analyses of genetic distances show that SCMV isolates from *Canna* sp. are very distant to isolates from other host species, suggesting a divergent evolution of the virus in this host plant. The availability of more complete genome sequences of SCMV from nongrass species will shed further light on the evolutionary aspects of this virus as a precursor to interrogating the biological consequences of the newly evolved sequence variants.

The comparative analyses of isolates NG-SC and NG-MZ derived from sugarcane and maize plants, respectively, show that they differ in the length of their polyprotein nucleotide and amino acid sequences. Further analyses revealed the presence of an extra 11 amino acid residue "GAKPAASGAGS" within the 3′-distal NIb sequences of NG-MZ. Interestingly, these extra amino acid residues have also been documented for other maize isolates of SCMV from the African continent, including isolates from Rwanda (KF744391-92), Ethiopia (KP772216 and KP860936), and Kenya (MH205604). It is unknown what fitness advantage or cost the presence or absence of this residue confers on SCMV isolates, if any. Future reverse genetic studies are needed to determine the biological function of these extra residues.

The possibility that recombination may have contributed to the evolution of SCMV isolates was suggested by Gao et al. (2011) and later confirmed via genome-wide analysis of several isolates (Moradi et al. 2016; Padhi and Ramu 2011; Xie et al. 2016). In the latter studies, recombination events involving few SCMV isolates were identified, leading the authors to conclude that recombination is the major driving force in the evolution and emergence of new variants of SCMV. Recombination analyses performed with complete polyprotein sequences of all available GenBank sequences and isolates derived in this study confirmed the recombinant nature of several isolates identified in previous studies (Moradi et al. 2016; Padhi and Ramu 2011; Xie et al. 2016). Additionally, the maize isolate NG-MZ from Nigeria was also determined to have arisen out of a putative recombination event involving isolates belonging to phylogroup V (Fig. 4). Given that SCMV is neither transmitted via true seeds nor mechanically transmitted, the only plausible explanation for the recombinant nature of NG-MZ is that it arose via interactions between two parental isolates that were coinoculated into a single host plant by aphid-mediated transmission (Louie and Knoke 1975). The perennial nature of sugarcane and *R. cochinchinensis* may provide greater opportunities for recombination among SCMV variants in contrast to the annually propagated maize plant. Regardless of the origin of the putative recombinant SCMV isolates, our results reinforce the notion that genetic recombination influenced the evolution of natural SCMV populations more greatly than was previously realized, in line with the observations made about the recombination-prone nature of positive-sense RNA viruses (Chare and Holmes 2004). Interestingly, Silva-Rosales et al. (2015) also alluded to the occurrence of widespread recombination events in SCMV genomes from five continents based on their yet-to-be published data.

In conclusion, the results obtained in this study have practical implications for cereal cropping systems in Nigeria, and possibly other sub-Saharan African countries, where co-cultivation of sugarcane with maize and other cereals occurs in the same or adjacent fields. The observation that many farmers lack awareness of virus diseases in cereal crops is disturbing. Hence, there is need for greater investment in extension plant pathology support for farmers within the region to educate them on the importance of virus diseases, recognition of virus symptoms, practice of good crop rotation, avoidance of mixed cultivation of susceptible host plant species, and control of grass weed species that may serve as virus reservoirs. These key ingredients, along with the cultivation of disease-resistant or -tolerant cultivars, will be pivotal for an integrated and sustainable management of SCMD in Nigeria.

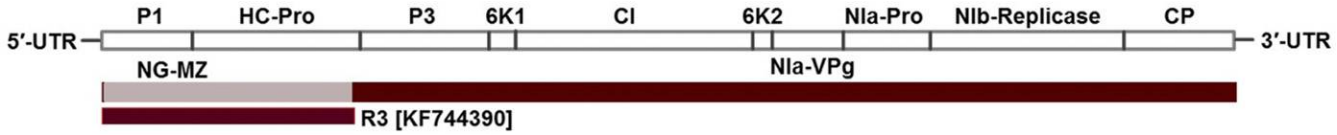


Fig. 4. Putative recombination event in the genome of sugarcane mosaic virus (SCMV) isolate NG-MZ (GenBank accession no. MH795797). The potential major and minor parental sequences are isolates Bomet G1 (MG932077) from Kenya and R3 (KF744390) from Rwanda. The event was detected by seven RDP4-implemented programs (Martin et al. 2005) with very strong statistical supports ($P \leq 3.766 \times 10^{-08}$).

Acknowledgments

Adama Yahaya is a fellow of the Norman E. Borlaug Leadership Enhancement in Agriculture Program. The authors are grateful to Mallam Ibrahim Bello (Ahmadu Bello University, Zaria, Kaduna State, Nigeria) and Dr. Mona Damaj (Texas A&M AgriLife Research & Extension Center, Weslaco, TX, USA) for technical help and all the farmers in the areas surveyed. We are grateful to the Texas A&M AgriLife Research & Extension Center, Weslaco for facility support. We thank the anonymous reviewers for valuable suggestions.

Literature Cited

Addy, H. S., Nurmalasari, Wahyudi, A. H. S., Sholeh, A., Anugrh, C., Iriyanto, F. E. S., Darmanto, W., and Sugiharto, B. 2017. Detection and response of sugarcane against the infection of *sugarcane mosaic virus* (SCMV) in Indonesia. Agronomy (Basel) 7:50.

Alabi, O. J., Al Rwahnih, M., Mekuria, T. A., and Naidu, R. A. 2014. Genetic diversity of *Grapevine virus A* in Washington and California vineyards. Phytopathology 104:548-560.

Alabi, O. J., Villegas, C., Gregg, L., and Murray, K. D. 2016. Complete nucleotide sequences of a new bipartite begomovirus from *Malvastrum* sp. plants with bright yellow mosaic symptoms in South Texas. Arch. Virol. 161:1729-1733.

Alegria, O. M., Royer, M., Bousalem, M., Chatenet, M., Peterschmitt, M., Girard, J. C., and Rott, P. 2003. Genetic diversity in the coat protein coding region of eighty-six sugarcane mosaic virus isolates from eight countries, particularly from Cameroon and Congo. Arch. Virol. 148:357-372.

Barnett, O. W. 1986. Surveying for plant viruses: Design and considerations. Pages 147-166 in: Plant Virus Epidemics: Monitoring, Modelling and Predicting Outbreaks. G. D.McLean, B. G. Garrett, and W. G. Ruesink, eds. Academic Press, Australia.

CABI. 2018. Sugarcane mosaic virus (mosaic of abaca). In: Invasive Species Compendium. Wallingford, UK: CAB International. https://www.cabi.org/isc.

Chare, E. R., and Holmes, E. C. 2004. Selection pressures in the capsid genes of plant RNA viruses reflect mode of transmission. J. Gen. Virol. 85:3149-3157.

Dangora, D. B., Yahaya, A., Khan, A. U., and Zangoma, M. A. 2014. Identification of virus isolates inducing mosaic of sugarcane in Makarfi Local Government Area of Kaduna State. Afr. J. Biotechnol. 13:1351-1357.

Dellaporta, S. L., Wood, J., and Hicks, J. B. 1983. A plant DNA minipreparation: Version II. Plant Mol. Biol. Report. 1:19-21.

Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792-1797.

Gao, B., Cui, X. W., Li, X. D., Zhang, C. Q., and Miao, H. Q. 2011. Complete genomic sequence analysis of a highly virulent isolate revealed a novel strain of *Sugarcane mosaic virus*. Virus Genes 43:390-397.

Gemechu, A. L., Chiemsombat, P., Attathom, S., Reanwarakorn, K., and Lersrutaiyotin, R. 2005. Cloning and sequence analysis of coat protein gene for characterization of *Sugarcane mosaic virus* isolated from sugarcane and maize in Thailand. Arch. Virol. 151:167-172.

Gordon, D. T., and Thottappilly, G. 2003. Maize and sorghum. Pages 295-336 in: Virus and Virus-like Diseases of Major Crops in Developing Countries. G. Loebenstein and G. Thottappilly, eds. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Ha, C., Coombs, S., Revill, P. A., Harding, R. M., Vu, M., and Dale, J. L. 2008. Design and application of two novel degenerate primer pairs for the detection and complete genomic characterization of potyviruses. Arch. Virol. 153:25-36.

Hall, T. A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. 41:95-98.

Iken, J. E., and Amusa, N. A. 2004. Maize research and production in Nigeria. Afr. J. Biotechnol. 3:302-307.

Jones, D. T., Taylor, W. R., and Thornton, J. M. 1992. The rapid generation of mutation data matrices from protein sequences. CABIOS 8:275-282.

Kumar, S., Stecher, G., and Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33:1870-1874.

Leng, P., Ji, Q., Tao, Y., Ibrahim, R., Pan, G., Xu, M., and Lübberstedt, T. 2015. Characterization of *Sugarcane mosaic virus* Scmv1 and Scmv2 resistance regions by regional association analysis in maize. PLoS One 10:e0140617.

Louie, R., and Knoke, J. K. 1975. Strains of *Maize dwarf mosaic virus*. Plant Dis. Rep. 59:518-522.

Martin, D. P., Williamson, C., and Posada, D. 2005. RDP2: Recombination detection and analysis from sequence alignments. Bioinformatics 21:260-262.

Moradi, Z., Mehrvar, M., Nazifi, E., and Zakiaghl, M. 2016. The complete genome sequences of two naturally occurring recombinant isolates of *Sugarcane mosaic virus* from Iran. Virus Genes 52:270-280.

Muhammad, B., Alegbejo, M. D., Kashina, B. D., and Banwo, O. O. 2016. Occurrence and distribution of potyviruses infecting sorghum in Kaduna and Kano States, Nigeria. Arch. Phytopathol. Plant Prot. 49:281-292.

Niblett, C. L., and Claflin, L. E. 1978. Corn lethal necrosis - a new virus disease of corn in Kansas. Plant Dis. Rep. 62:15-19.

Oluwafemi, S., Kraberger, S., Shepherd, D. N., Martin, D. P., and Varsani, A. 2014. A high degree of African streak virus diversity within Nigerian maize fields includes a new mastrevirus from *Axonopus compressus*. Arch. Virol. 159:2765-2770.

Padhi, A., and Ramu, K. 2011. Genomic evidence of intraspecific recombination in *Sugarcane mosaic virus*. Virus Genes 42:282-285.

Silva, G., Bömer, M., Nkere, C., Kumar, P. L., and Seal, S. E. 2015. Rapid and specific detection of *Yam mosaic virus* by reverse-transcription recombinase polymerase amplification. J. Virol. Methods 222:138-144.

Silva-Rosales, L., Alcalá-Briseño, R. I., and Espejel, F. 2015. "*Sugarcane mosaic*". Pages 131-143 in: Virus Diseases of Tropical and Subtropical Crops. P. Tennant and G. Fermin, eds. CAB International, Wallingford, UK.

Thorat, A., Pal, R. K., Shingote, P., Kharte, S., Nalavade, V., Dhumale, R. D., Pawar, B. H., and Babu, H. 2015. Detection of *Sugarcane mosaic virus* in diseased sugarcane using ELISA and RT-PCR technique. J. Pure Appl. Microbiol. 9:319-327.

U.S. Dep. Agric. 2010. Foreign Agricultural Service: Global Agriculture Information Network, Nigeria Grain and Feed Annual Report. https://www.fas.usda.gov/data/nigeria-grain-and-feed-annual-1

Wang, J. G., Zheng, H. Y., Chen, H. R., Adams, M. J., and Chen, J. P. 2010. Molecular diversities of *Sugarcane mosaic virus* and *Sorghum mosaic virus* isolates from Yunnan Province, China. J. Phytopathol. 158:427-432.

Wangai, A. W., Redinbaugh, M. G., Kinyua, Z. M., Miano, D. W., Leley, P. K., Kasina, M., Mahuku, G., Scheets, K., and Jeffers, D. 2012. First report of *Maize chlorotic mottle virus* and maize lethal necrosis in Kenya. Plant Dis. 96:1582-1583.

Wu, L., Zu, X., Wang, S., and Chen, Y. 2012. *Sugarcane mosaic virus* - long history but still a threat to industry. Crop Prot. 42:74-78.

Xie, X., Chen, W., Fu, Q., Zhang, P., An, T., Cui, A., and An, D. 2016. Molecular variability and distribution of *Sugarcane mosaic virus* in Shanxi, China. PLoS One 11:e0151549.

Xu, D. L., Park, J. W., Mirkov, T. E., and Zhou, G. H. 2008. Viruses causing mosaic disease in sugarcane and their genetic diversity in southern China. Arch. Virol. 153:1031-1039.

Yahaya, A., Dangora, D. B., Alegbejo, M. J., Kumar, P. L., and Alabi, O. J. 2017. Identification and molecular characterization of a novel sugarcane streak mastrevirus and an isolate of the A-strain of *Maize streak virus* from sugarcane in Nigeria. Arch. Virol. 162:597-602.

Yahaya, A., Dangora, D. B., Khan, A. U., and Zangoma, M. A. 2014. Detection of sugarcane mosaic disease (SCMD) in crops and weeds associated with sugarcane fields in Makarfi and Sabon Gari local government areas of Kaduna state, Nigeria. Int. J. Curr. Sci. 11:E99-E104.

Yao, W., Ruan, M., Qin, L., Yang, C., Chen, R., Chen, B., and Zhang, M. 2017. Field performance of transgenic sugarcane lines resistant to *Sugarcane mosaic virus*. Front. Plant Sci. 8:104.

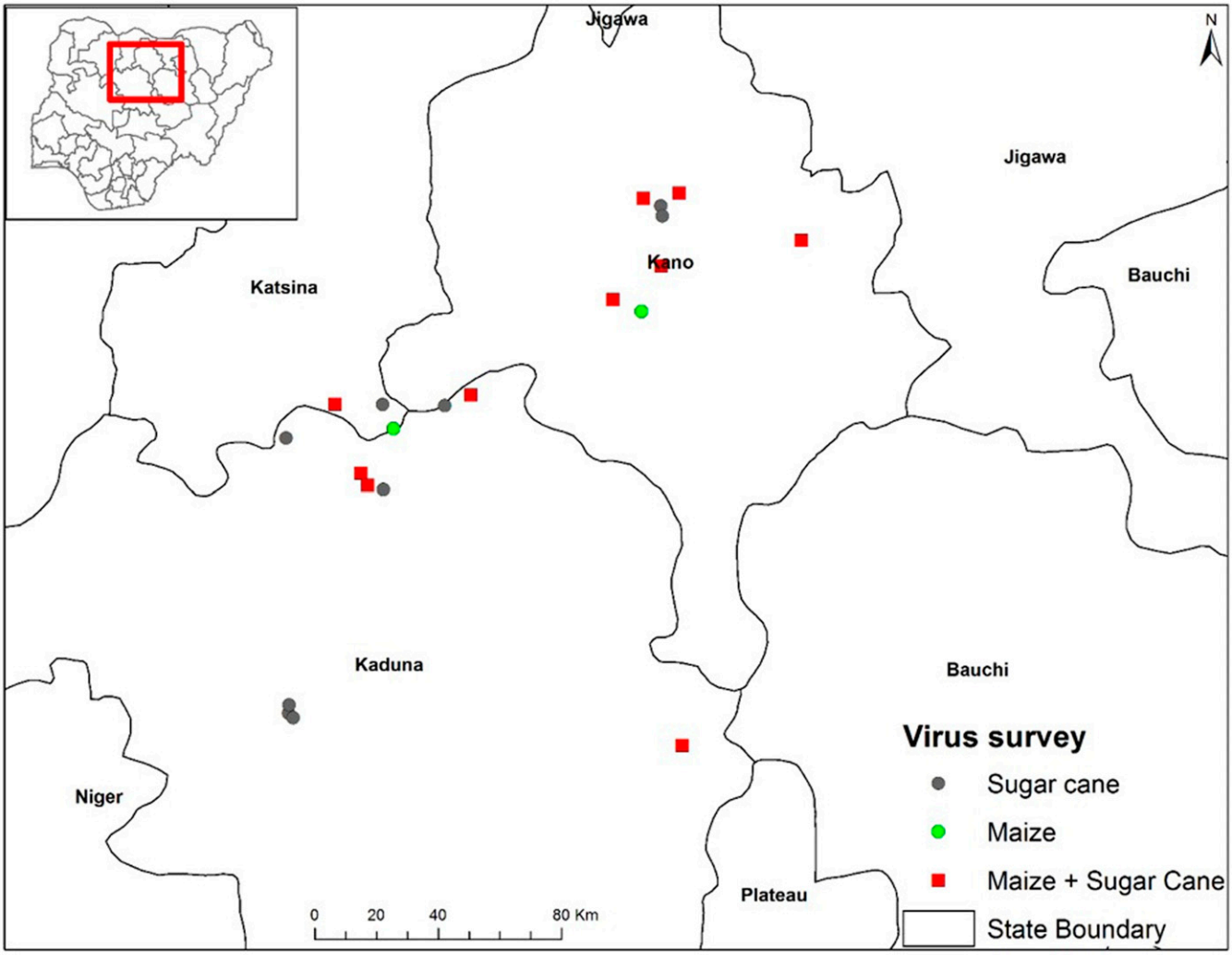
Zheng, L., Rodoni, B. C., Gibbs, M. J., and Gibbs, A. J. 2010. A novel pair of universal primers for the detection of potyviruses. Plant Pathol. 59:211-220.

Supplementary Table 1. Putative recombination events involving isolates of sugarcane mosaic virus (SCMV) based on the analysis of 90 complete polyprotein sequences of the virus (GenBank = 88; This Study = 2; see Fig. 3). Only events detected by five of the suites of programs implemented in the RDP v.4.56 software (Martin et al. 2005), with good statistical support ($P \leq 0.05$) were considered acceptable.

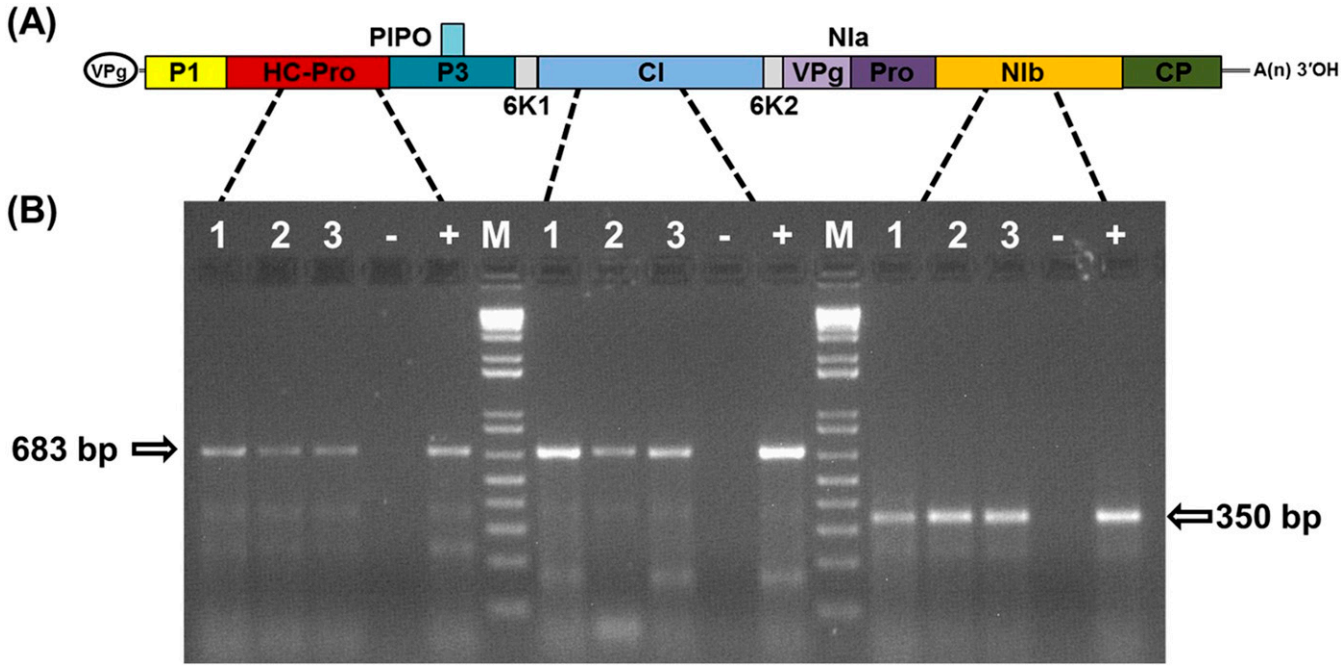
Event #[a]	Putative recombinant[b]	Putative parentals[b]		Site	Detected by[c]	P-value[c]
		Major	Minor			
1	HZ6-CN-MZ [I]	FP5-CN-MZ [I]	HZ8-CN-MZ [IV]	2658 - 9192	RGB**M**CST	8.377×10^{-57}
2	FP1-CN-MZ [I]	HY8-CN-MZ [I]	SX5-CN-MZ [IV]	4304 - 8000	RGBM**C**ST	2.424×10^{-31}
3	LZ1, LZ2, YL9, HY9, FP5, HY8, HY7, YL12, FP2, HZ6, FP1-CN-MZ [I]	Zhejiang-CN-MZ [I]	GD-CN-MZ [II]	1787 - 4929	R**G**BMCST	1.555×10^{-04}
4	YL9, HY9, FP5, HY8, HY7, YL12, FP2, FP1-CN-MZ [I]	SX5-CN-MZ [IV]	LZ1-CN-MZ [I]	8001 - 9109	RGB**M**CST	2.640×10^{-31}
5	G3-Bomet-Kenya-MZ [V]	Elgeyo Marakwet G1-Kenya-MZ [V]	HZ8-CN-MZ [IV]	7952 - 9192	RGB**M**CST	1.921×10^{-25}
6	ZRA, NRA-Iran-MZ [III]	G3-Bomet-Kenya-MZ [V]	Brisbane-Australia [III]	1266 - 7790	RGBM**C**ST	1.401×10^{-08}
7	JAL1-Mexico-MZ, Kiriyanga-Kenya-SrGm [I]	VER1-Mexico-MZ [I]	Ohio-USA-MZ [V]	8260 - 9127	RGBM**C**ST	1.162×10^{-11}
8	Kiriyanga-Kenya-SrGm [I]	Seehausen-Germany-MZ [I]	Embu G1-Kenya-MZ [V]	7948 - 8259	RGBM**C**S	1.181×10^{-08}
9	F3S2, F2S1-Ethiopia-MZ [I]	Kiriyanga-Kenya-SrGm [I]	Embu G1-Kenya-MZ [V]	5396 - 7947	RGBM**C**ST	4.977×10^{-28}
10	F3S2, F2S1-Ethiopia-MZ [I]	Seehausen-Germany-MZ [I]	Embu G1-Kenya-MZ [V]	7955 - 9139	RGB**M**CST	2.757×10^{-20}
11	NG-MZ, Embu G1-Kenya-MZ [V]	Unknown (Bomet G1-Kenya-MZ) [V]	R3-Rwanda-MZ [V]	23 - 2059	RGB**M**CST	3.766×10^{-08}
12	ARG-915-Argentina-SC [III]	Unknown (Brisbane) [III]	ARG-1662-Argentina-SC	261 - 5427	RBMC**S**T	1.772×10^{-02}

			[III]			
13	ZRA, NRA-Iran-MZ [III]	Unknown (GD-CN-MZ) [II]	FZ-C1-CN-SC [III]	8039 - 9151	RGBM**C**ST	1.884×10^{-11}
14	SX3, SX2-CN-MZ [IV]	YL8-CN-MZ [IV]	Yunan-CN-MZ [IV]	5012 - 7528	RGBM**C**ST	7.453×10^{-13}
15	Kiriyanga G2-Kenya-MZ [I]	Bomet G1-Kenya-MZ [V]	Kiriyanga-Kenya-SrGm [I]	20 - 5956	RGB**M**ST	2.614×10^{-35}
16	HN, SD-CN-MZ [I]	Sp-Spain-MZ [I]	CD4-CN-MZ [I]	4853 - 9192	R**G**BST	7.391×10^{-03}
17	Beijing-CN-MZ [I]	HN-CN-MZ [I]	Zhejiang-CN-MZ [I]	1674 - 3592	RGBM**C**ST	4.093×10^{-15}
20	SX-CN-MZ [I]	JAL1-Mexico-MZ [I]	Sp-Spain-MZ [I]	2978 - 9126	RG**B**ST	8.448×10^{-13}
21	YL9, HY9, FP5-CN-MZ [I]	SX5-CN-MZ [IV]	FP2-CN-MZ [I]	4304 - 7132	RGB**M**T	1.260×10^{-22}
22	NXZN2, NXZ7272, NXA1-CN-MZ [I]	SX-CN-MZ [I]	Zhejiang-CN-MZ [I]	5460 - 8886	RGBM**C**ST	2.683×10^{-18}

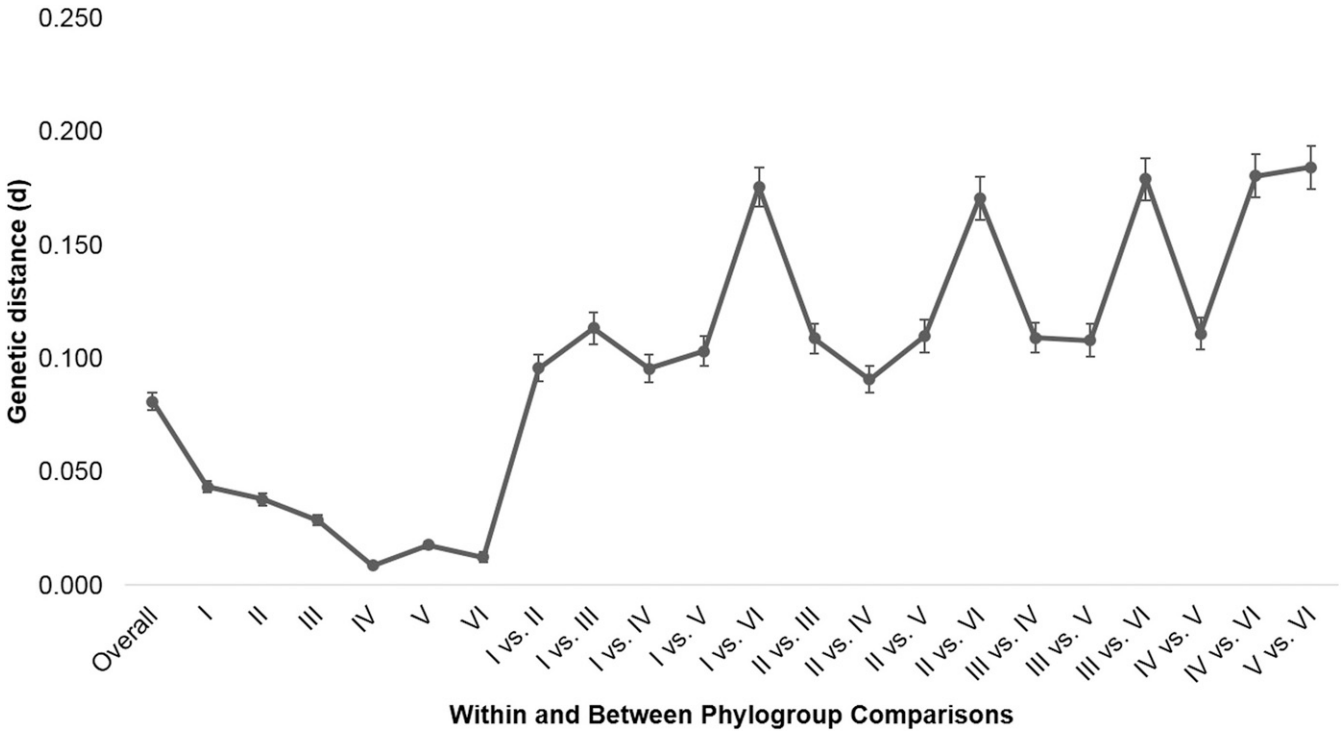
[a]Event #s 18 and 19 are not listed because they did not meet the threshold of events detected by at least five of the seven RDP4-implemented programs.
[b]M = maize; SC = sugarcane; CN = China. The phylogroup to which the isolates belong are provided in square brackets [].
[c]The different recombination detection used in the RDP4 analyses are R, RDP; G, GENECONV; B, BOOTSCAN; M, MAXCHI; C, CHIMAERA; S, SISCAN; and 3S, 3SEQ. P-values are shown only for methods with the least level of statistical significance (bold font).



Supplementary Figure 1. The map of Nigeria (top left corner) showing the northern guinea savannah study area (red rectangular box). The sampled farmer's fields are marked on the enlarged abridged map.



Supplementary Figure 2. (A), Genome map of sugarcane mosaic virus (SCMV) depicting the positions of different open reading frames (ORFs) as colored boxes. The map is drawn based on the complete genome sequence of SCMV isolate Ohio (JX188385). Details of protein abbreviations are available in Adams et al. (2012). (B), Agarose gel electrophoresis of gene-specific RT-PCR products. Potyvirid generic primers (Ha et al. 2008; Zheng et al. 2010) specific to a portion of the helper component protease (HC-Pro), cylindrical inclusion (CI) protein and nuclear inclusion replicase (NIb) gene coding sequences were used to amplify 683, 683 and 350 base pair (bp) DNA fragments (lanes 1 to 3, indicated by arrows), respectively, from each isolate. M, 1Kb Plus DNA Ladder (Thermo Fisher Scientific); –, sample from SCMV-negative sugarcane; and +, sample from SCMV-positive sugarcane.



Supplementary Figure 3. Estimates of genetic distances over complete polyprotein sequence pairs of sugarcane mosaic virus (SCMV). The analysis was performed for the overall dataset and for within or between phylogroups identified in Fig. 3. All analyses were conducted in MEGA7 (Kumar et al. 2016) using the JTT matrix-based model (Jones et al. 1992) and the rate variation among sites was modeled with a gamma distribution (shape parameter = 0.6). The analysis involved 90 complete polyprotein amino acid sequences (This Study = 2; GenBank = 88; see Fig. 3). All positions containing gaps and missing data were eliminated.