Arch Virol
DOI 10.1007/s00705-016-3148-5

ANNOTATED SEQUENCE RECORD

# 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

**Adama Yahaya[1] · Danladi B. Dangora[1] · Matthew D. Alegbejo[2] · P. Lava Kumar[3] · Olufemi J. Alabi[4]**

Received: 31 August 2016 / Accepted: 31 October 2016


**Abstract** Sugarcane and maize plants showing symptoms typical of those described for the so-called "African streak viruses" (AfSVs) were encountered during field surveys conducted from February to July 2015 to document viruses infecting both crops across the northern Guinea savannah region of Nigeria. As part of this study, two categories of complete mastrevirus-like genome sequences were obtained from nine samples (maize = 2; sugarcane = 7). In pairwise comparisons, the full-length genomes of the first sequence category (2,687 nt each; maize = 2; sugarcane = 2) shared 96 to 99% identity with global isolates of the A-strain of maize streak virus (MSV-A), indicating that sugarcane may also serve as a reservoir host to MSV-A. Analysis of the complete genomes belonging to the second sequence category (2,757 nt each; sugarcane = 5) showed that they shared 42 to 67% identity with their closest AfSV relatives, thus indicating that they represent sequences of a novel mastrevirus. Both sequence categories shared 61-62% sequence identity with each other. Further analysis revealed that the novel sugarcane-infecting virus, tentatively named as sugarcane chlorotic streak virus (SCSV), arose from a putative interspecific recombination event involving two grass-infecting mastreviruses, eragrostis streak virus and urochloa streak virus, as putative parental sequences. The results of this study add to the repertoire of diverse AfSVs present in cereal and sugarcane mixed cropping landscapes in the northern Guinea savannah region of Nigeria, with implications for disease epidemiology.

The genus *Mastrevirus* (family *Geminiviridae*) consists of leafhopper-transmitted circular ssDNA viruses with ~2.7-kb monopartite genomes that are capable of infecting and causing diseases in monocotyledonous and dicotyledonous plants [10]. Of the several currently recognized mastreviruses, maize streak virus (MSV) is widespread across the African continent and ranks among the top 10 economically most important plant viruses [14]. Up to 11 strains of MSV, named in alphabetical order from MSV-A to MSV-K, have been characterized to date, among which the maize-adapted MSV-A is the most prevalent, with a debilitating impact on maize crops across Africa [7]. In addition to MSV, other 'African' monocot-infecting mastreviruses have also been characterized in recent years, including sugarcane streak Reunion virus (SSRV), urochloa streak virus (USV), maize streak Reunion virus (MSRV), and Axonopus compressus streak virus (ACSV) [12], thus indicating a greater diversity of mastreviruses on the African continent.

A major characteristic of the cereal-production landscape in Nigeria is mixed cropping systems in which maize and other cereals (grown mainly as food staples) are cultivated together in the same field with sugarcane (grown

**Electronic supplementary material** The online version of this article (doi:10.1007/s00705-016-3148-5) contains supplementary material, which is available to authorized users.

✉ Olufemi J. Alabi
alabi@tamu.edu

[1] Department of Biological Sciences, Ahmadu Bello University, Zaria, Kaduna State, Nigeria

[2] Department of Crop Protection, Ahmadu Bello University, Zaria, Kaduna State, Nigeria

[3] International Institute of Tropical Agriculture (IITA), PMB 5320 Ibadan, Nigeria

[4] Department of Plant Pathology & Microbiology, Texas A&M AgriLife Research & Extension Center, Weslaco, TX 78596, United States of America

Published online: 04 November 2016

primarily as a cash crop). Both crops may also be grown as monocrops in adjacent fields or in rotation with each other. Unfortunately, there are a number of common viruses that cause debilitating diseases in both cereal and sugarcane crops and pose a major threat to their production. Several of the so-called ''African streak viruses'' (AfSVs) have been characterized from maize and grass samples in Nigeria [12]. In addition, recent studies have documented the occurrence of potyviruses – sugarcane mosaic virus (SCMV), maize dwarf mosaic virus (MDMV) and sorghum mosaic virus (SrMV) – in maize, sugarcane and weed samples from the northern Guinea savannah region of Nigeria [2]. To the best of our knowledge, there is a dearth of information on mastreviruses infecting sugarcane in the country. Here, we report the identification and molecular characterization of field isolates of MSV-A infecting maize and sugarcane and a novel mastrevirus from sugarcane in Nigeria.

The samples analyzed in this study were obtained from two maize plants (Mz5 and Mz80) and seven sugarcane plants (Sc10, Sc11, Sc13, Sc29, Sc30, Sc57 and Sc97) from six farmers' fields during surveys conducted from February to July 2015 (Supplementary Table 1). The surveys were part of a study conducted to gain an understanding of the landscape of viruses in maize- and sugarcane-production systems in three leading cereal-producing states (Kaduna, Katsina and Kano) of northern Nigeria. Symptoms observed on affected plants consisted mainly of severe chlorotic streaks (Figs. 1A-C and E), but most of the symptomatic sugarcane plants (Sc10, Sc11, Sc13, Sc29 and Sc30) were also stunted due to shortened internodes (Fig. 1D). Symptomatic leaf tissue samples taken from each plant were dried and stored under calcium chloride at room temperature. The dried tissue samples were shipped under permit from the United States Department of Agriculture-Agricultural Research Service Plant Protection and Quarantine (P526P-14-04321) to the Texas A&M AgriLife Research and Extension Center (TAMU AgiLife) Weslaco facility for further analysis. Total nucleic acid (TNA) extracts from each sample [3] was subjected to rolling-circle amplification (RCA) using a TempliPhi Amplification Kit (GE Healthcare Life Sciences, Uppsala, Sweden) essentially as described in the manufacturer's protocol. Two-microliter aliquots of the RCA product of each sample were used as template in PCR with the degenerate primer pair MSV215-234 and MSV1770-1792 [13], which is capable of amplifying a DNA fragment of ~1400 bp spanning the coat protein (CP) and near-full-length RepB sequences of several mastreviruses (Supplementary Table 2). Furthermore, cDNA was synthesized from each TNA sample (ThermoScript RT-PCR System, Life Technologies, Carlsbad, CA, USA) and then used as template for PCR with the generic primer pairs CIFor/CIRev, HPFor/HPRev [5] and NIb2F/NIb3R [16] targeting potyviruses and the primer pair Pol-G-F/Pol-G-R [8] targeting poleroviruses. PrimeSTAR® GXL DNA Polymerase (Takara-Bio USA, Inc., Mountain View, CA, USA) was used in all PCR assays regardless of the template type.

DNA bands of the expected size (1400 bp) were obtained from all nine samples (maize = 2; sugarcane = 7) with the degenerate primer pair MSV215-234 and MSV1770-1792 (data not shown). In addition, four of the nine samples (Mz5, Sc29 Sc30 and Sc97) also tested positive with at least one of the three generic potyvirus primer pairs indicating that the subset of sugarcane (n = 3) and maize (n = 1) source plants harbor mixed infections of a mastrevirus and a potyvirus (Supplementary Table 1). Symptoms observed in the plants with mixed infection consisted of chlorotic streaks and mosaic patterns (Fig. 1E). None of the nine samples tested positive with the Pol-G-F/Pol-G-R primer pair, indicating absence of poleroviruses in the source plants.

The ~1400-bp DNA fragments obtained with the MSV215-234 and MSV1770-1792 primer pair were cloned individually, sequenced, and analyzed as described [1]. BLASTN analysis of these sequences confirmed their viral origin and their specificities to AfSVs. Further analysis revealed that all nine partial genome sequences shared 62 to 100% identity to each other, suggesting that they may likely represent more than one species of AfSV. Based on the identity values, the nine partial mastrevirus genome sequences were categorized into two groups, designated as MSV-like (for sequences closely related to each other and highly similar to strains of MSV) and SSV-like (for sequences closely related to each other but distantly related to all known AfSVs). The within- and between-group partial nucleotide (nt) sequence identities were determined to be 99-100% and 62-63%, respectively, further supporting the proposition that both groups represent distinct mastrevirus taxa. Four isolates (maize, Mz5 and Mz80; sugarcane, Sc57 and Sc97) belonged to the MSV-like group, while the SSV-like group comprised the remaining sugarcane isolates (Sc10, Sc11, Sc13, Sc29 and Sc30). None of the samples were found to harbor a mixture of the identified MSV-like and SSV-like mastrevirus sequences.

To further characterize the MSV-like and SSV-like isolates, the partial nt sequences of each group were aligned using the program MUSCLE [4], and the aligned sequences were used in the design of two pairs of abutting primers (Supplementary Table 2) for each sequence group to amplify the complete mastrevirus genomes from each of the nine isolates. PCR was performed with the abutting primers (Supplementary Table 2) and RCA templates as described above, with the reaction conditions for each primer pair consisting of 30 cycles of 98 °C for 10 s., 55 °C for 15 s., and 68 °C for 30 s. The ~2.7-kb DNA


**Fig. 1** Virus-like symptoms observed on maize and sugarcane plants under field conditions in the northern Guinea savannah region of Nigeria. Plants in the MSV-like group display typical maize streak virus symptoms on maize (A) and sugarcane (B), while those in the SSV-like group show chlorotic streaks (C) and shortened cane internodes (D) on sugarcane. Sugarcane plants were also observed showing mixtures of chlorotic streaks and mosaic patterns (E)

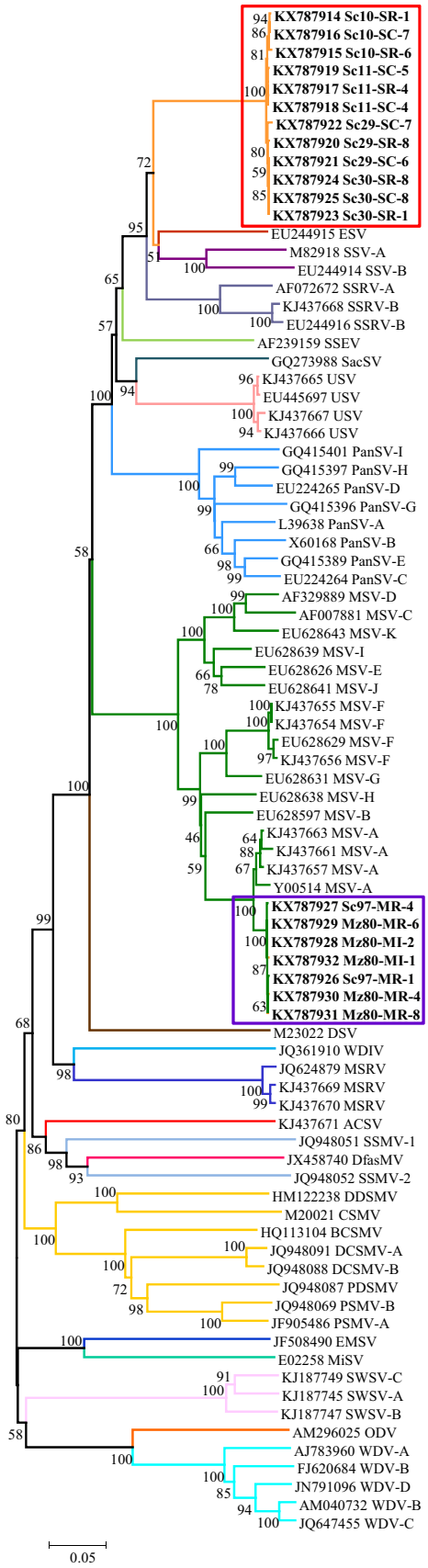
fragments obtained for two representative isolates of the MSV-like group (Mz80 and Sc97) and four of the SSV-like group (Sc10, Sc11, Sc29 and Sc30) were each gel-eluted, A-tailed, and then cloned individually using kits and protocols described previously [1]. Two to five recombinant clones per isolate were sequenced bi-directionally with the universal M13 primers. Additional primers were designed for primer-walking the full genome plasmid DNA of each clone per isolate (Supplementary Table 2). The sequence fragments obtained for each plasmid DNA were edited and assembled to derive the complete virus genome sequences using the suite of programs contained in the BioEdit Sequence Alignment Editor version 7.2.3 [6]. Overall, 12 SSV-like (KX787914-25) and seven MSV-like (KX787926-32) complete genome sequences were obtained in the study.

Each of the seven MSV-like complete genome sequences was 2,687 nt in length, in contrast to the SSV-like sequences, which were each 2,757 nt in length. In pairwise comparisons, the complete genomes of all MSV-like sequences were 100% identical to each other, while the SSV-like sequences shared 99-100% complete-genome sequence identity among themselves. These results indicate that sequences in each group represent distinct virus species, in agreement with results obtained using the partial genome sequences. In addition, the range of complete genome sequence identities between both sequence groups was determined to be 61-62%, further confirming the results obtained with the partial genome sequences.

We adopted the rational mastrevirus species and strain demarcation criteria and suite of programs recommended by Muhire et al. [10] in order to further determine the species status of the isolates belonging to the MSV-like and SSV-like groups. A pairwise alignment of 84 full genome sequences of mastreviruses (19 from this study and 65 from GenBank) was performed using the MUSCLE alignment option of the SDTv1.2 program [11]. The same aligned sequences were also used for inferring evolutionary relationships among the 84 mastrevirus full-genome sequences. The results showed clustering of the MSV-like

sequences determined in this study in the MSV-A clade on the mastrevirus phylogenetic tree (Fig. 2). Similar clustering patterns were observed when the deduced amino acid sequences of the coat protein (CP) and replicase (Rep) were used in phylogenetic analysis (Supplementary Fig. 1). As shown in the color-coded matrix of pairwise identity scores, all seven MSV-like sequences shared 96-97% full-genome sequence identity with global isolates of MSV-A and 76-87% with isolates of the other 10 MSV strains (Supplementary Fig. 2). In contrast, all 12 SSV-like sequences formed a distinct cluster on the matrevirus phylogenetic tree (Fig. 2), with the trees based on CP and Rep amino acid sequences producing similar clustering patterns (Supplementary Fig. 1). All 12 SSV-like sequences shared lower levels of full-genome sequence similarity with several AfSVs, including eragrostis streak virus (ESV) (66-67%), sugarcane streak virus (SSV) (64-67%), sugarcane streak virus Reunion (64-66%), urochloa streak virus (USV) (62-64%), sugarcane streak Egypt virus (62-63%), and saccharum streak virus (61%) (Supplementary Fig. 2). They were also distantly related to two other streak viruses recently reported from Nigeria, ACSV (42%) and MSRV (45%), on a whole-genome basis (Supplementary Fig. 2). Based on these results, and considering the <78% genome-wide nucleotide sequence identity species demarcation threshold established for members of the genus *Mastrevirus* [10], it can be concluded that the sequences in the MSV-like group are isolates of the A-strain of maize streak virus, while those in the SSV-like group represent isolates of a novel mastrevirus. The tentative name "sugarcane chlorotic streak virus" (SCSV) is proposed for this virus.

The genomic features of SCSV are similar to those described for members of the genus *Mastrevirus*, including the presence of the conserved "TAATATT↓AC" nonanucleotide sequence described for known geminiviruses, the partitioning of the single genome component into two virion-sense proteins and two variants of complementary-sense replication-associated proteins, and the presence of two (short and long) intergenic region sequences demarcating the virion- and complementary-sense proteins [10]. Further analysis revealed that whereas SCSV was more phylogenetically related to, and shared maximum identity (66-67%) with, ESV on a full-genome basis (Fig. 2; Supplementary Fig. 2), its CP amino acid sequences were most similar (89-90% identical) and phylogenetically related (Supplementary Fig. 1A) to the corresponding sequences of isolates of USV. In contrast, SCSV clustered more closely and shared maximum sequence identity (78-79%) with ESV and the A-strain of SSV, based on analysis of the Rep amino acid sequences (Supplementary Fig. 1B). These results indicate incongruence in the comparative phylogenies of full-genome versus gene-specific sequences of SCSV, suggesting that the novel mastrevirus might have



◀**Fig. 2** Unrooted neighbor-joining (NJ) tree depicting the phylogenetic relationships among 84 complete genome sequences of strains of monocot-infecting mastreviruses. The analysis involved 19 complete genome sequences determined in this study and 65 sequences obtained from GenBank. The NJ tree was generated using the MEGA6 program [15]

arisen due to recombination. To test this hypothesis, the aligned sequences derived above were scanned using the RDP4.36 program [9] with default parameters, with the exception that the default option of 'linear' was replaced with 'circular' genome architecture. A putative recombination event was considered authentic based on a threshold of its prediction by at least four of the seven RDP-implemented programs and strong statistical support ($P > 0.05$) for the detected events. Based on these criteria, SCSV was determined to be a putative recombinant, with the putative major and minor parental sequences being ESV and USV, respectively (Supplementary Table 3). Interestingly, although ESV has not yet been observed in Nigeria, USV was recently characterized from *Urochloa* sp. and *Eleucine coracana* in northern Nigeria [12], suggesting that weed host plants might provide an environment for the evolution of agriculturally relevant recombinant mastreviruses.

In conclusion, we report here a novel mastrevirus, tentatively named SCSV, from field-collected sugarcane samples in northern Nigeria. Our analysis indicated that SCSV is a recombinant virus associated with chlorotic streak symptoms in sugarcane. We also present evidence for the occurrence of the economically important maize-adapted MSV-A in sugarcane under field conditions in Nigeria. MSV-A has been detected previously in maize crops and weed species in Nigeria [12], but not from sugarcane. Therefore, the results of this study suggest that the perennially grown sugarcane crop in Nigeria may serve as reservoir host of cereal-infecting viruses such as MSV to maintain the 'green bridge' required for survival of the virus when maize is off season. The results also point to a greater diversity of mastreviruses and their reservoir hosts under field conditions in Nigeria. Future studies are necessary to determine the prevalence of SCSV, identify its leafhopper vector(s) under experimental and field conditions, and to determine its host range among cultivated and non-cultivated plant species growing in sugarcane and cereal landscapes in the northern Guinea savannah region of Nigeria.

**Acknowledgements** Adama Yahaya is a fellow of the Norman E. Borlaug Leadership Enhancement in Agriculture Program (Borlaug-LEAP) funded by the USAID. The authors are grateful to Mallam Ibrahim Bello (Ahmadu Bello University, Zaria, Kaduna State, Nigeria) and Ms. L. Gregg (Texas A&M AgriLife Research & Extension Center, Weslaco, TX, USA) for technical help. This study was funded through financial support from the Borlaug LEAP Program, Texas A&M AgriLife Research & Extension Center, Weslaco, the CGIAR Research Program on MAIZE, and the Amina Ado Foundation.

**Compliance with ethical standards**

**Conflict of interest** All authors declare that they have no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

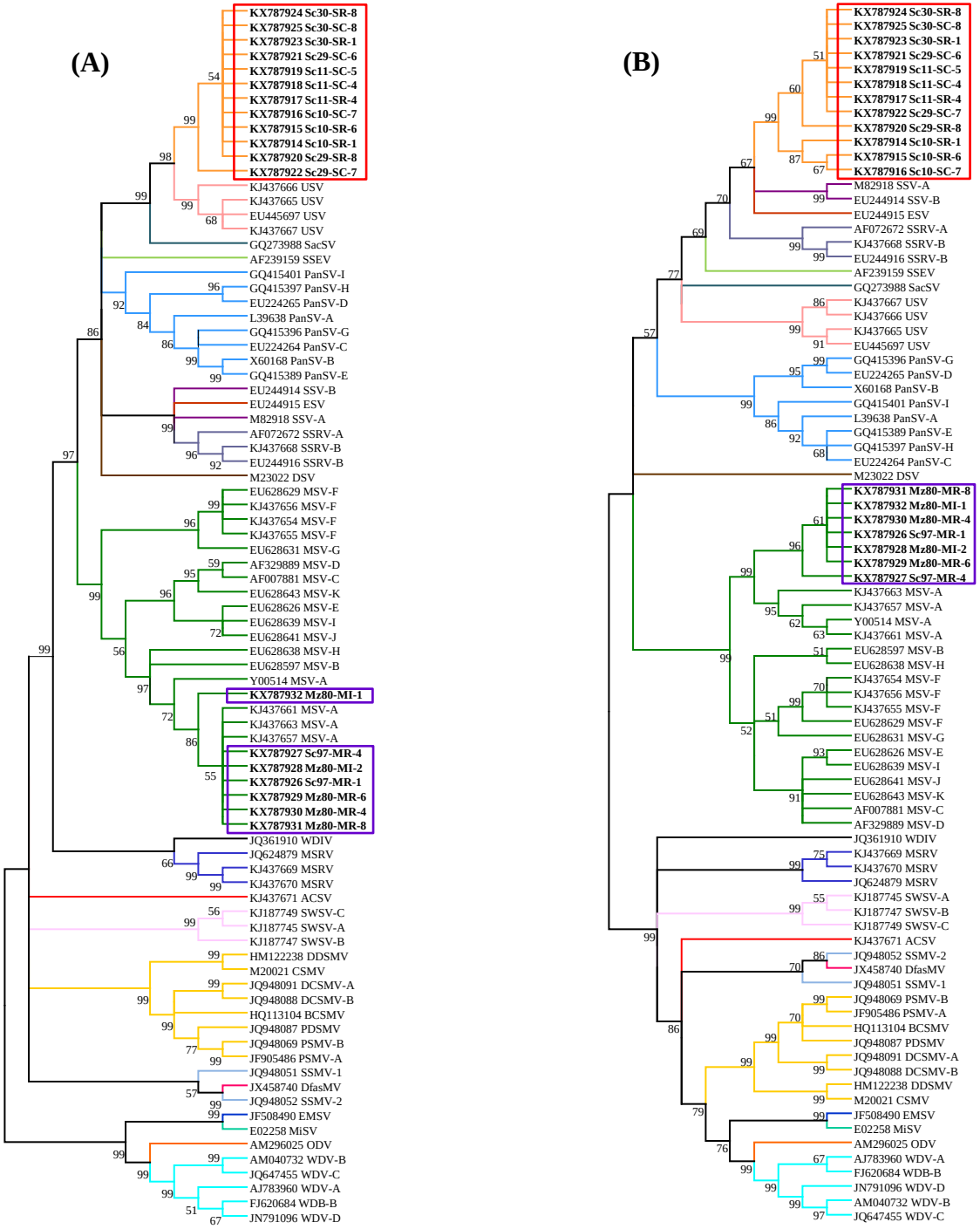
## References

1. Alabi OJ, Villegas C, Gregg L, Murray KD (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
2. Dangora DB, Yahaya A, Khan AU, Zangoma AM (2014) Identification of virus isolate inducing mosaic of sugarcane in Makarfi Local Government areas of Kaduna State. Afr J Biotechnol 13:1351–1357
3. Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep 1:19–21
4. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl Acids Res 32:1792–1797
5. Ha C, Coombs S, Revill PA, Harding RM, Vu M, Dale JL (2008) Design and application of two novel degenerate primer pairs for the detection and complete genomic characterization of potyviruses. Arch Virol 153:25–36
6. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41:95–98
7. Harkins GW, Martin DP, Duffy S, Monjane AL, Shepherd DN, Windram OP, Owor BE, Donaldson L, van Antwerpen T, Sayed RA, Flett B, Ramusi M, Rybicki EP, Peterschmitt M, Varsani A (2009) Dating the origins of the maize-adapted strain of maize streak virus, MSV-A. J Gen Virol 90:3066–3074
8. Knierim D, Deng TC, Tsai WS, Green SK, Kenyon L (2010) Molecular identification of three distinct *Polerovirus* species and a new recombinant Cucurbit aphid-borne yellows virus strain infecting cucurbit crops in Taiwan. Plant Pathology 59:991–1002
9. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evolut 1(1):vev003
10. Muhire B, Martin DP, Brown JK, Navas-Castillo J, Moriones E, Zerbini FM, Rivera-Bustamante R, Malathi VG, Briddon RW, Varsani A (2013) A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus *Mastrevirus* (family *Geminiviridae*). Arch Virol 158:1411–1424
11. Muhire B, Varsani A, Martin DP (2014) SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 9:e108277
12. Oluwafemi S, Kraberger S, Shepherd DN, Martin DP, 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
13. Palmer KE, Rybicki EP (2001) Investigation of the potential of *Maize streak virus* to act as an infectious gene vector in maize plants. Arch Virol 146:1089–1104
14. Rybicki EP (2015) A top ten list for economically important plant viruses. Arch Virol 160:17–20

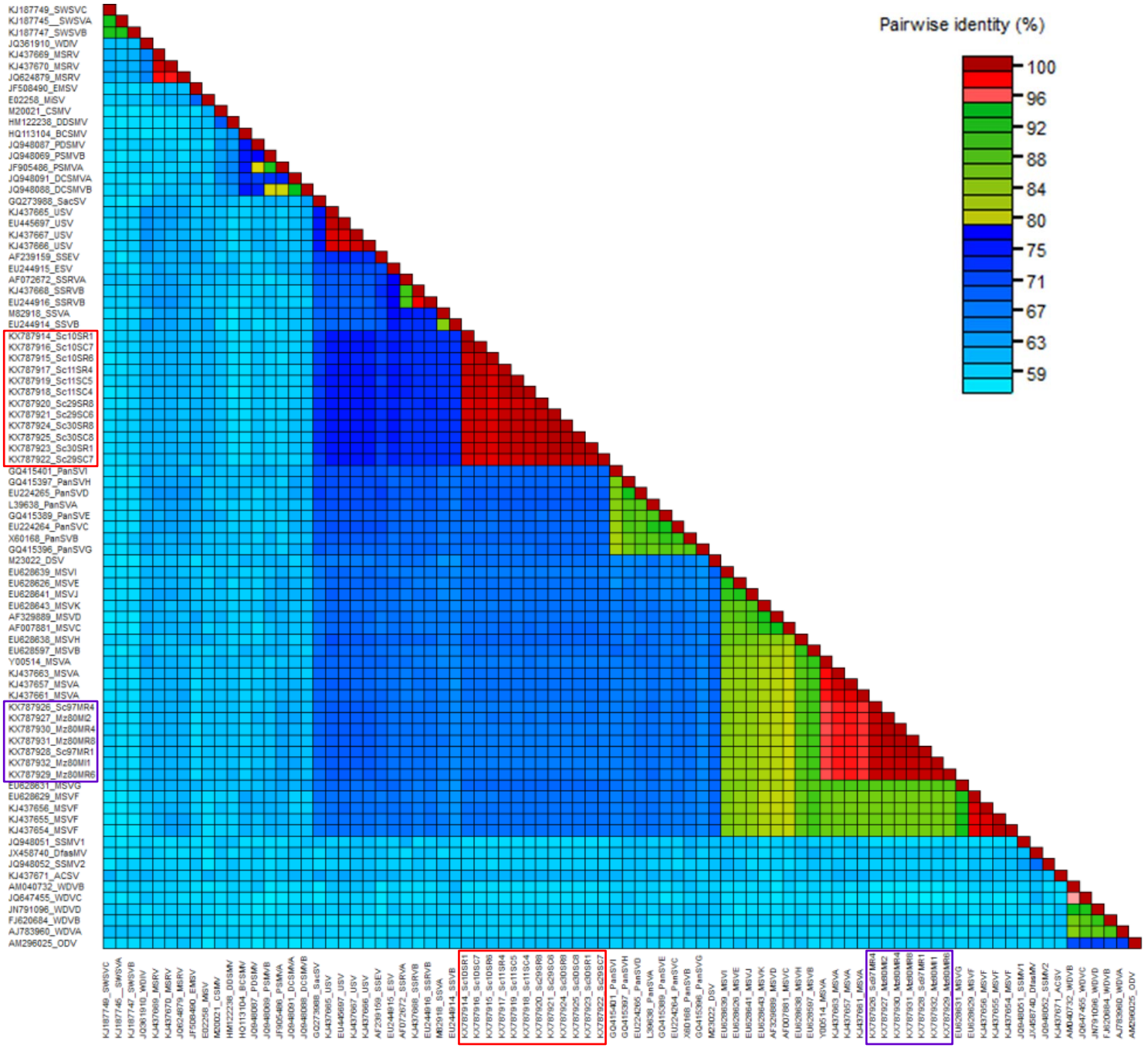
15. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729

16. Zheng L, Rodoni BC, Gibbs MJ, Gibbs AJ (2010) A novel pair of universal primers for the detection of potyviruses. Plant Pathol 59:211–220

**Supplementary Figure 1**



**Supplementary Figure 2**



**Supplementary Table 1.** Provenance of virus isolates analyzed in the study in Nigeria.

| Sample ID | Location | State | GPS Coordinates | | Host plant | Symptoms | Viruses detected[α] |
|---|---|---|---|---|---|---|---|
| Mz5 | Botanical garden | Kaduna | N10°58.080 | E007°39.234 | Maize | Chlorotic streaks and mosaic | MSV, SCMV |
| Sc10 | Kubanni | Kaduna | N11°05.716 | E007°43.270 | Sugarcane | Chlorotic streaks | SCSV |
| Sc11 | Kubanni | Kaduna | N11°05.717 | E007°43.271 | Sugarcane | Chlorotic streaks | SCSV |
| Sc13 | Kubanni | Kaduna | N11°05.718 | E007°43.272 | Sugarcane | Chlorotic streaks | SCSV |
| Sc29 | Damquarters | Kaduna | N11°09.851 | E007°34.909 | Sugarcane | Chlorotic streaks and mosaic | SCSV, SCMV |
| Sc30 | Damquarters | Kaduna | N11°09.852 | E007°34.910 | Sugarcane | Chlorotic streaks and mosaic | SCSV, SCMV |
| Sc57 | Kwangilla | Kaduna | N11°07.617 | E007°42.616 | Sugarcane | Chlorotic streaks and mosaic | MSV |
| Mz80 | Tashar Kargo | Katsina | N11°21.637 | E007°35.664 | Maize | Chlorotic streaks | MSV |
| Sc97 | Kahutu | Katsina | N11°23.175 | E007°41.710 | Sugarcane | Mosaic | MSV, SCMV |

[α] MSV, maize streak virus; SCMV, sugarcane mosaic virus

**Supplementary Table 2.** Primers used for amplification of partial and complete genome fragments of mastreviruses characterized in the study.

| Primer name[α] | Target[β] | | Sequences (5'-3')[δ] | Orientation | Expected size of amplicon (bp) | Reference |
|---|---|---|---|---|---|---|
| | Virus | Region | | | | |
| MSV1770-1792 | AfSVs | RepB | TTggVCCgMVgATgTASAg | Sense | 1400 | Palmer and |
| MSV215-234 | | MP | CCAAAKDTCAgCTCCTCCg | Antisense | | Rybicki, 2001 |
| SSV2481-2502[a] | SSV- | RepA | CATAgTTCCAAACCTACAgCCT | Sense | 2700 | This Study |
| SSV2499-2520[a] | like | RepA | ACTgCAACTTAgAgCCAgAggC | Antisense | | |
| SSV542-561[a]* | SSV- | CP | gTAAgAAgCAgCCgAgAgTg | Sense | 1145 | This Study |
| SSV1667-1686[a]* | like | RepB | ATACTgTCCTTgCTggAAgC | Antisense | | |
| SSV93-113[b] | SSV- | IR | ggCgCCAAggACTATAAgATg | Sense | 2700 | This Study |
| SSV110-131[b] | like | IR | TgCgATCCCACATACAAgCATC | Antisense | | |
| SSV869-889[b]* | SSV- | CP | TATTYCCgCATCCggATTCTC | Sense | 1188 | This Study |
| SSV2036-2056[b]* | like | RepA | CTgTTCCCTgAgACACATgAA | Antisense | | |
| MSV1945-1964[c] | MSV- | RepA | TggATTgCggATgAggATTg | Sense | 2700 | This Study |
| MSV1961-1982[c] | like | RepA | gATTCAggAAgAgTTCACCAAT | Antisense | | |
| MSV2622-2643[c]* | MSV- | IR | TTATAgTggTTgTgAATgggCC | Sense | 1181 | This Study |
| MSV1096-1115[c]* | like | CP | TACAACgCTCCTCTCTggAT | Antisense | | |
| MSV296-316[d] | MSV- | MP | gCTgAgAgACCTTATCTTAgT | Sense | 2700 | This Study |
| MSV313-332[d] | like | MP | TTgCCgAgCCTTCAgAACTA | Antisense | | |
| MSV2573-2594[d]* | MSV- | IR | CTCAACTCTATACCAACCggTg | Sense | 1775 | This Study |
| MSV1639-1660[d]* | like | RepB | AAgACgCAATCTACAACATCgT | Antisense | | |

[α] Primers with the same alphabet represent combinations of abutting oligonucleotides used to obtain complete virus genomes; and primers used for genome-walking are indicated with *. The genome positions of the primers are also indicated with position number 1 on the virus genome corresponding to the nicking site on the conserved nonanucleotide sequence "TAATATT↓AC"

[β] AfSVs, African streak viruses; SSV-like, sugarcane streak-like viruses; MSV-like, maize streak-like viruses; MP, movement protein; IR, intergenic region; RepA, C1 protein; RepB, C2 protein

[δ] The following IUB group codes are used to identify redundancies: V = A + C + G, M = A + C, S = G + C, K = G + T, D = A + G + T, Y = C + T.

**Supplementary Table 3.** Putative recombination events involving genome sequences of sugarcane chlorotic streak virus (SCSV).

| **Recombinant clone** | **Breakpoints**[α] | **Putative Parentals [% Similarity]**[β] | | **Methods**[δ] | ***p*-value range**[†] |
|---|---|---|---|---|---|
| | | Major | Minor | | |
| Sc10-SR-6 (KX787915) | 380-1186 | ESV (EU244915) [73.5%] | USV(KJ437665) [80.6%] | R, B, M, C, S, 3S | $1.237 \times 10^{-14}$ - $8.932 \times 10^{-04}$ |
| Sc10-SR-1 (KX787914) | 380-1186 | ESV (EU244915) [73.8%] | USV(KJ437665) [80.6%] | R, B, M, C, S, 3S | $1.237 \times 10^{-14}$ - $8.932 \times 10^{-04}$ |
| Sc10-SC-7 (KX787916) | 380-1186 | ESV (EU244915) [73.9%] | USV(KJ437665) [80.6%] | R, B, M, C, S, 3S | $1.237 \times 10^{-14}$ - $8.932 \times 10^{-04}$ |
| Sc11-SC-5 (KX787919) | 380-1080 | ESV (EU244915) [74.5%] | USV(KJ437665) [80.2%] | R, B, M, C, S, 3S | $1.237 \times 10^{-14}$ - $8.932 \times 10^{-04}$ |
| Sc11-SR-4 (KX787917) | 380-1080 | ESV (EU244915) [74.0%] | USV(KJ437665) [80.6%] | R, B, M, C, S, 3S | $1.237 \times 10^{-14}$ - $8.932 \times 10^{-04}$ |
| Sc11-SC-4 (KX787918) | 380-1080 | ESV (EU244915) [74.0%] | USV(KJ437665) [80.6%] | R, B, M, C, S, 3S | $1.237 \times 10^{-14}$ - $8.932 \times 10^{-04}$ |
| Sc29-SC-7 (KX787922) | 370-1225 | ESV (EU244915) [73.5%] | USV(KJ437665) [80.7%] | R, B, M, C, S, 3S | $1.237 \times 10^{-14}$ - $8.932 \times 10^{-04}$ |
| Sc29-SR-8 (KX787920) | 370-1225 | ESV (EU244915) [73.6%] | USV(KJ437665) [80.7%] | R, B, M, C, S, 3S | $1.237 \times 10^{-14}$ - $8.932 \times 10^{-04}$ |
| Sc29-SC-6 (KX787921) | 380-1186 | ESV (EU244915) [73.9%] | USV(KJ437665) [80.6%] | R, B, M, C, S, 3S | $1.237 \times 10^{-14}$ - $8.932 \times 10^{-04}$ |
| Sc30-SR-8 | 380-1186 | ESV (EU244915) | USV(KJ437665) | R, B, M, C, S, | $1.237 \times 10^{-14}$ - $8.932 \times$ |

| | | | | | |
|---|---|---|---|---|---|
| (KX787924) | | [73.9%] | [80.6%] | 3S | $10^{-04}$ |
| Sc30-SC-8 | 380-1186 | ESV (EU244915) | USV(KJ437665) | R, B, M, C, S, | $1.237 \times 10^{-14}$ - $8.932 \times 10^{-04}$ |
| (KX787925) | | [73.9%] | [80.6%] | 3S | |
| Sc30-SR-1 | 380-1186 | ESV (EU244915) | USV(KJ437665) | R, B, M, C, S, | $1.237 \times 10^{-14}$ - $8.932 \times 10^{-04}$ |
| (KX787923) | | [73.9%] | [80.6%] | 3S | |

[α] Position number 1 on the virus genome corresponds to the nicking site on the conserved nonanucleotide sequence "TAATATT↓AC"

[β] Virus abbreviations: ESV, eragrostis streak virus isolate ESV[ZmGur]; USV, urochloa streak virus isolate USV_NG_ng23_Sam1_2011

[δ] The RDP-implemented programs [9] that yielded positive predictions are: R, RDP; B, Bootscan; M, MaxChi; C, CHIMAERA; S, SiScan; 3S, and 3SEQ

[†] The range of p-values corresponding to the most and least statistically supported methods is presented