

Molecular characterization of two previously undescribed begomovirus-associated alphasatellite molecules infecting malvaceous species in Cameroon

W.N. Leke^{1, 2*}, A. Kvarnhed², S. Avelar³, and J.K. Brown³

¹Institute of Agricultural Research for Development (IRAD), Ekona, P.O. Box 25 Buea, Cameroon

²Department of Plant Biology, Uppsala BioCenter, Swedish University of Agricultural Sciences and Linnean Center of Plant Biology in Uppsala, P.O. Box 7080, SE-750 07 Uppsala, Sweden

³School of Plant Sciences, The University of Arizona, Tucson, AZ 85721, USA

Corresponding Author: walter.leke@fulbrightmail.org

Abstract

Two begomovirus-associated alphasatellites were isolated from okra and a *Malvastrum* plant (Malvaceae) in Cameroon. The complete nucleotide sequences of the okra- and *Malvastrum*-infecting alphasatellites were 1375 and 1416-1418 nucleotides, respectively, and both exhibited features characteristic of other alphasatellites. Based on pairwise sequence comparisons, these previously undescribed alphasatellites are distinct species herein tentatively named pepper yellow vein Mali alphasatellite and cotton leaf curl Gezira alphasatellite3, belonging to the genera *Colecusatellite* and *Gosmusatellite*, respectively. Taken together with previous studies, alphasatellites endemic to Cameroon appear to be more diverse and widespread, and to infect plants of far more species and families than currently recognized.

Key words: *Alphasatellitidae*, *Colecusatellite*, *Geminialphasatellitinae*, *Gosmusatellite*

Begomoviruses (genus *Begomovirus*; family *Geminiviridae*) are the most widespread viruses with a genome of single-stranded (ss) DNA that infect plants. They cause crop losses in the tropics and subtropics where their whitefly vectors, members of the *Bemisia tabaci* (Genn.) sibling species complex (possibly, cryptic species) are endemic [2, 17]. Begomoviruses have either a mono- or bi-partite genome of ~2.8 and ~5.2 kb in size, respectively [3]. Monopartite begomoviruses in Asia and Africa are often associated with non-viral satellite ssDNA molecules, referred to as beta- and alphasatellites, which are approximately of half-unit size with respect to the helper virus genome, at ~1.4 kb.

Alphasatellites encode one protein that shares high amino acid sequence similarity with the begomoviral replication-associated protein (Rep). The protein is similar to the rolling-circle

replication initiator of viruses in the family *Nanoviridae* [4, 11] by which they are capable of replicating autonomously. In addition to Rep, alphasatellites have an A-rich region ~200 nucleotides long, downstream of the region encoding Rep and a conserved nonanucleotide sequence, TAGTATTAC. The helper begomovirus is required for systemic infection of the host plant and for whitefly-mediated transmission. The role of alphasatellites in begomovirus infection is variable and not well understood. However, some alphasatellites are known to influence begomovirus accumulation and also may modulate symptom phenotype in the plant host [7, 12].

Taxonomically, alphasatellites have been assigned to the family *Alphasatellitidae*, containing the two subfamilies *Geminialphasatellitinae* and *Nanoalphasatellitinae*. The latter subfamily contains the nanovirus-associated alphasatellites, whereas the *Geminialphasatellitinae* contains geminivirus-associated alphasatellites belonging to the genera *Ageyesisatellite*, *Clecrusatellite*, *Colecusatellite* or *Gosmusatellite* [1]. In Cameroon, six monopartite begomoviruses have been previously reported from different plant species (ageratum, okra and tomato), exhibiting begomovirus-like disease symptoms and all were associated with one or more alphasatellites such as cotton leaf curl Gezira alphasatellite (CLCuGA), tomato leaf curl Buea alphasatellite (ToLCuBuA), tomato leaf curl Cameroon alphasatellite (ToLCuCMA), all belonging to the genus *Colecusatellite* and okra yellow crinkle Cameroon alphasatellite (OkYCCA), now suggested to be named cotton leaf curl Gezira alphasatellite2 (CLCuGeA2), belonging to the genus *Gosmusatellite* [9, 10, 11]. Here, two previously undescribed alphasatellite species were identified in two malvaceous hosts, together with at least two monopartite begomovirus, respectively, in southwestern Cameroon.

Leaf samples of plants of *Malvastrum* sp. (n = 3) from Mundemba in southwestern Cameroon (4°58'24.39"N, 8°54'28.72"E) and okra (*Abelmoschus esculentus* (L.) Moench) (n = 3),

from Barombi Nkang, southwestern Cameroon, (4°35'33.97"N, 9°27'52.26"E), collected from symptomatic plants (foliar crinkling, leaf curling, yellow mosaic) (Fig. 1) between 2007 and 2008, were applied to FTA Classic Cards [16] or pressed and dried between notebook pages and the total DNA was later extracted as previously described [15]. For initial begomovirus identification and identification of restriction sites for cloning, the core coat protein gene (*cp*) region was amplified as previously described [20]. Out of the three samples respectively collected from each plant species, only one tested positive. Subsequently, full-length genomes of the begomoviruses from okra and *Malvastrum* were obtained by rolling circle amplification (RCA) [8] and restriction digestion with the enzymes *Nde*I and *Hind*III, respectively. The digestion products were ligated into the corresponding sites of the pGEM-5zf (+) and pGEM-3zf (+) plasmid vectors, respectively (Promega).

The alphasatellites were obtained by digest of RCA products with *Bam*HI or PCR amplification with the alphasatellite primers DNA101/DNA102 [5]. Amplicons were cloned using either *Bam*HI-digested pBluescript II-KS+ (Invitrogen) or pGEM-T Easy (Promega) plasmid vector, according to the manufacturer's instructions. The ligated begomovirus and alphasatellite amplicons were transformed into competent cells of *Escherichia coli* DH5 α (Invitrogen) and clones with the expected insert sizes were subjected to bi-directional capillary Sanger sequencing (Macrogen, South Korea) and (BecHA Hub, Kenya). Pairwise nucleotide sequence identities were calculated using the Sequence Demarcation Tool (SDT), version 1.2 [14]. All attempts made to amplify betasatellites using universal betasatellite primers as well as specific betasatellite primers based on sequences of betasatellites already identified in Africa were negative.

Analyses of full-length amplification products showed that the okra sample OBKG was infected by an isolate of *Cotton leaf curl Gezira virus* (CLCuGeV, MN372225) and an alphasatellite (FM164739) which is 1375 nt in size and had the features characteristic of other alphasatellites. Sequence comparisons using SDT showed that the alphasatellite shared highest nt identity with two isolates of cotton leaf curl Gezira alphasatellite 2 (CLCuGA2, MK032294) from Burkina Faso and (CLCuGA2, FN675287) from Cameroon at 80.5%, Mesta yellow vein mosaic alphasatellite (MeYVMA, JX183090) from India at 77.5%, hollyhock yellow vein alphasatellite (HoYVA, FR772086) from Pakistan at 74.0% and an isolate of CLCuGA2 (FN554583), at 79.9%, previously reported from okra in Burkina Faso [19]. The Rep of the alphasatellite of 302 AA was 13 AA residues longer than that of its closest relatives. Based on the International Committee on the Taxonomy of Viruses (ICTV)-established species cut-off for alphasatellite species, of <88% [1], the alphasatellite from okra represents a new species for which the name cotton leaf curl Gezira alphasatellite 3 (CLCuGeA3) is proposed. Previously, two other alphasatellites have been identified together with CLCuGeV and for those we suggest that the name be changed to CLCuGeA1 (presently CLCuGeA) and CLCuGeA2 (presently OkYCCA). In this way, all alphasatellites found associated with CLCuGeV are named after their helper virus. Interestingly, the CLCuGeA satellites belong to both the genus *Colecusatellite* (CLCuGeA1/CLCuGeA) and the genus *Gosmusatellite* (CLCuGeA2/OkYCCA and CLCuGeA3).

Similar analyses of the sample Mun11 from *Malvastrum* revealed the presence of *Pepper yellow vein Mali virus* (PepYVMLV, MN372224) and an alphasatellite (FN675303, FN675297 and FN675298). The cloned alphasatellite sequences (MUN11a, MUN11b and MUN11c) from the *Malvastrum* plant ranged in size from 1416-1418 nt, which is longer than the regular size of 1.3

kb for most identified alphasatellites, but similar to a long sequence of CLCuGeA reported to infect tomato in Sudan [6]. Its Rep is 317 aa, similar to that of its closest relative CLCuGeA1 and had features characteristic of other known alphasatellites. They shared their highest nt identity, at 72%, with two CLCuGeA1 isolates (FN675289 and HE858192) from okra in Cameroon [9], while also sharing <88% identity with all other known alphasatellites, indicating that the sequences from the *Malvastrum* sp. plant represent a new species tentatively termed pepper yellow vein Mali alphasatellite. The identification of PepYVMLV from *Malvastrum* sp. in Cameroon (this study) and from *Sida acuta* in Burkina Faso (MH778662) increases its known host range to include plants of the family Malvaceae.

The determined alphasatellite sequences and selected alphasatellite reference sequences were aligned using MUSCLE [5], followed by phylogenetic analysis using Maximum Likelihood (ML) (1000 bootstrap iterations), implemented in MEGA6.0 [18]. The ML tree placed each of the alphasatellites into two well-supported clades (100% bootstrap values) corresponding to the genera *Colecusatellite* and *Gosmusatellite* (Fig. 2). The colecusatellite clade contained alphasatellites previously identified in Asia, Burkina Faso, Cameroon, Egypt, Kenya and Mali as well as the alphasatellite from *Malvastrum* (labeled A in Fig. 2) and other alphasatellites previously identified from Cameroon in okra (B), tomato (C) and ageratum (D) [11, 12], while the gosmusatellite clade contained alphasatellites from Burkina Faso, Cameroon and Pakistan as well as the newly identified CLCuGeA3 from okra in Cameroon (labeled E in Fig. 2) and CLCuGeA2/OkYCCA (F) also isolated from okra in Cameroon [12]. Recombination Detection Program (RDP4) [13] analysis of the alphasatellite sequences subjected to phylogenetic analysis did not provide significant support for recombination for either of the newly discovered alphasatellite species, based on a significant *p*-value of <10⁻³ for a minimum of three RDP methods.

While alphasatellites of the genus *Colecusatellite* are known to have a wider host range thus far, those of the genus *Gosmusatellite* have been found mostly in malvaceous hosts (cotton, okra, hollyhock etc), and in West Africa, only in okra plants. In Cameroon, alphasatellites have been identified in hosts representing three different plant families (Fig. 1): Asteraceae (*Ageratum conyzoides*), Solanaceae (tomato), and Malvaceae (*Malvastrum* sp., okra). This observation suggests that alphasatellites may be more widespread and infect a greater number of host species representing even more diverse plant families than previously recognized.

Acknowledgements

This research was supported by the following organizations or foundations to which we express our sincere gratitude: a grant to WN Leke from the International Foundation of Science; Rothamsted International-African Fellows program; The Swedish Institute; USAID IPM-CRSP *Insect-transmitted viruses*-Global Theme Project, and The Nilsson-Ehle Foundation, Sweden.

Conflict of Interest

The authors, Walter N. Leke, Anders Kvarnheden, Sofia Avelar, and Judith K. Brown declare no conflict of interest.

Ethical approval

This research did not involve studies with human participants or other animals.

References

1. Briddon RW, Martin DP, Roumagnac P, Navas-Castillo J, Fiallo-Olive E, Lett JM, Zerbini FM, Varsani A (2018) *Alphasatellitidae*: a new family with two subfamilies for the classification of geminivirus- and nanovirus-associated alphasatellites. *Arch Virol* 163:2587-2600 <https://doi.org/10.1007/s00705-018-3854-2>
2. Brown JK (2010) *Bemisia*: Phylogenetic biology of the *Bemisia tabaci* sibling species group. Pages 31-67 in: *Bemisia*: Bionomics and Management of a Global Pest. Stansly, P.A. and S. E. Naranjo (eds.), Springer, Netherlands, 350pp.
3. Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva JCF, Fiallo-Olive E, Briddon RW, Hernandez-Zepeda C, Idris AM, Malathi VG, Martin DP, Rivera-Bustamante R, Ueda S, Varsani A (2015) Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. *Arch Virol* 160:1593-1619
4. Bull SE, Briddon RW, Markham PG (2003) Universal primers for the PCR-mediated amplification of DNA 1: A satellite-like molecule associated with begomovirus DNA- β complexes. *Mol Biotech* 23:83–86
5. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
6. Fiallo-Olivé E, Hamed A, Navas-Castillo J, Moriones E (2013) Cotton leaf curl Gezira alphasatellite associated with tomato leaf curl Sudan virus approaches the expected upper size limit in the evolution of alphasatellites. *Virus Res* 178:506-510
7. Idris AM, Shahid MS, Briddon RW, Khan AJ, Zhu JK, Brown JK (2011) An unusual alphasatellite associated with monopartite begomoviruses attenuates symptoms and reduces betasatellite accumulation. *J Gen Virol* 92:706-717

8. Inoue-Nagata AK, Albuquerque LC, Rocha WB, Nagata T (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage phi29 DNA polymerase. *J Virol Methods* 116:209-211
9. Leke WN, Kvarnheden A, Ngane EB, Titanji VPK, Brown JK (2011) Molecular characterization of a new begomovirus and divergent alphasatellite from tomato in Cameroon. *Arch Virol* 156:925-928
10. Leke WN, Brown JK, Ligthart ME, Sattar N, Njualement DK, Kvarnheden A (2012) *Ageratum conyzoides*: a host to a unique begomovirus disease complex in Cameroon. *Virus Res* 163:229-237
11. Leke WN, Sattar MN, Ngane EB, Ngeve JM, Kvarnheden A, Brown JK (2013) Molecular characterization of begomoviruses and DNA satellites associated with okra leaf curl disease in Cameroon. *Virus Res* 174:116-125
12. Mar TB, Mendes IR, Lau D, Fiallo-Olivé E, Navas-Castillo I, Alves MS, Murilo-Zerbini F (2017) Interaction between the New World begomovirus *Euphorbia yellow mosaic virus* and its associated alphasatellite: effects on infection and transmission by the whitefly *Bemisia tabaci*. *J Gen Virol* 98:1552-1562
13. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol* 1:vev003
14. Muhire B, Martin DP, Brown JK, Navas-Castillo Moriones JE, 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

15. Ndunguru J, Taylor NJ, Yadav J, Aly H, Legg JP, Aveling T, Thompson G, Fauquet CM (2005) Application of FTA technology for sampling, recovery and molecular characterization of viral pathogens and virus-derived transgenes from plant tissues. *Virology* 2:45
16. Owor BE, Shepherd DN, Taylor NJ, Edema R, Monjane AL, Thomson JA, Martin DP, Varsani A (2007) Successful application of FTA® Classic Card technology and use of bacteriophage Φ 29 DNA polymerase for large-scale field sampling and cloning of complete maize streak virus genomes. *J Virol Methods* 140:100–105
17. Rojas MR, Macedo MA, Maliano MR, Aguilar MS, Souza JO, Briddon RW, Kenyon L, Bustamante RFR, Zerbini FM, Adkins S, Legg JP, Kvarnheden A, Wintermantel WM, Sudarshana MR, Peterschmitt M, Lapidot M, Martin DP, Moriones E, Inoue-Nagata AK, Gilbertson RL (2018) World management of geminiviruses. *Annu Rev Phytopathol* 56:637–677
18. Tamura K, Glen S, Daniel P, Alan F, Sudhir K (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Bio Evol* 30:2725-2729
19. Tiendrébéogo F, Lefeuvre P, Hoareau M, Villemot J, Konate G, Traoré AS, Barro N, Traoré VS, Reynaud B, Traoré O, Lett JM (2010) Molecular diversity of *Cotton leaf curl Gezira virus* isolates and their satellite DNAs associated with okra leaf curl disease in Burkina Faso. *Virology* 7:48
20. Wyatt SD, Brown JK (1996) Detection of subgroup III geminivirus isolates in leaf extracts by degenerate primers and polymerase chain reaction. *Phytopathology* 86:1288–1293

Figures

Figure 1 Cultivated and wild plant hosts showing begomoviral disease symptoms. A, Okra plant with foliar crinkling symptoms and B, *Malvastrum* sp. exhibiting yellow mosaic symptoms.



Figure 2 Maximum Likelihood analysis using MEGA 6.0 showing the predicted relationships between alphasatellite isolates based on the MUSCLE alignment of complete nucleotide sequences. The alphasatellite isolates from Cameroon are indicated in bold, and those identified in this study are indicated by a square box. The values placed at the major nodes show bootstrap values (1000 iterations) where they exceeded 70 percent. The alphasatellites names are according to Briddon *et al.* [1], and the respective GenBank Accession number for each is indicated.

