Virus Research 238 (2017) 171–178

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Molecular diversity, geographic distribution and host range of monocot-infecting mastreviruses in Africa and surrounding islands

Simona Kraberger[a,b], Salem Saumtally[c], Daniel Pande[d,e], Michel H.R. Khoodoo[c,1], Sonalall Dhayan[c,1], Asha Dookun-Saumtally[c], Dionne N. Shepherd[f], Penelope Hartnady[g], Richard Atkinson[f], Francisco M. Lakay[f], Britt Hanson[f], Devasha Redhi[f], Adérito L. Monjane[f,h], Oliver P. Windram[i], Matthew Walters[b], Sunday Oluwafemi[j], Jean Michel-Lett[k], Pierre Lefeuvre[k], Darren P. Martin[g,**], Arvind Varsani[a,b,l,*]

[a] The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life sciences, Arizona State University, Tempe, AZ 85287-5001, USA
[b] School of Biological Sciences, University of Canterbury, Christchurch, 8140, New Zealand
[c] Mauritius Sugarcane Industry Research Institute, Réduit, Mauritius
[d] Department of Botany, Maseno University, P.O. Box 333, Maseno, Kenya
[e] Department of Biological and Biomedical Science and Technology, Laikipia University, P.O. Box 1100-20300, Nyahururu, Kenya
[f] Department of Molecular and Cell Biology, University of Cape Town, Rondebosch, 7701, Cape Town, South Africa
[g] Computational Biology Group, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, 7925, South Africa
[h] Department of Immunology, Norwegian Veterinary Institute, Pb 750 Sentrum, N-0106 Oslo, Norway
[i] Grand Challenges in Ecosystems & the Environment, Imperial College London, Silwood Park Campus, Buckhurst Road, SL5 7PY Ascot, Berks, UK
[j] Department of Crop Production, Soil and Environmental Management, Bowen University, P.M.B. 284, Iwo, Osun State, Nigeria
[k] CIRAD, UMR PVBMT, Pôle de Protection des Plantes, 7 Chemin de l'IRAT, 97410 Saint-Pierre, Ile de La Réunion, France
[l] Structural Biology Research Unit, Department of Clinical Laboratory Sciences, University of Cape Town, Observatory, Cape Town, South Africa

ARTICLE INFO

Keywords:
Mastrevirus
Geographic distribution
Natural host range
Poaceae

ABSTRACT

Maize streak virus (MSV), an important pathogen of maize in Africa, is the most extensively studied member of the *Mastrevirus* genus in the family *Geminiviridae*. Comparatively little is known about other monocot-infecting African mastreviruses, most of which infect uncultivated grasses. Here we determine the complete sequences of 134 full African mastrevirus genomes from predominantly uncultivated Poaceae species. Based on established taxonomic guidelines for the genus *Mastrevirus*, these genomes could be classified as belonging to the species *Maize streak virus*, *Eragrostis minor streak virus*, *Maize streak Reunion virus*, *Panicum streak virus*, *Sugarcane streak Reunion virus* and *Sugarcane streak virus*. Together with all other publicly available African monocot-infecting mastreviruses, the 134 new isolates extend the known geographical distributions of many of these species, including MSV which we found infecting *Digitaria* sp. on the island of Grand Canaria: the first definitive discovery of any African monocot-infecting mastreviruses north-west of the Saharan desert. These new isolates also extend the known host ranges of both African mastrevirus species and the strains within these. Most notable was the discovery of MSV-C isolates infecting maize which suggests that this MSV strain, which had previously only ever been found infecting uncultivated species, may be in the process of becoming adapted to this important staple crop.

1. Introduction

Members of the genus *Mastrevirus* (family *Geminiviridae*) are circular single-stranded (ss) DNA viruses (~2.5–2.7 kb) which infect a wide range of plants. Within this genus, species can be grouped into those which infect dicotyledonous (dicot) plants and those which infect monocotyledonous (monocot) plants. The most extensively characterised of the monocot-infecting mastreviruses are those found in

* Corresponding author at: The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life sciences, Arizona State University, Tempe, AZ, 85287-5001, USA.
** Corresponding author at: Computational Biology Group, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, 7925, South Africa.
E-mail addresses: darrenpatrickmartin@gmail.com (D.P. Martin), Arvind.varsani@asu.edu (A. Varsani).
[1] Previously at Mauritius Sugar Industry Research Institute.

http://dx.doi.org/10.1016/j.virusres.2017.07.001
Received 17 May 2017; Received in revised form 28 June 2017; Accepted 1 July 2017
Available online 04 July 2017
0168-1702/

Africa. Often referred to as the African streak viruses (AfSV), this group currently comprises twelve species which, throughout Africa and the adjacent Indian Ocean islands of La Réunion and Mayotte, are known both to infect various plant species in the family Poaceae, and transmitted by various leafhopper species in the genus *Cicadulina* (Martin and Shepherd, 2009; Rose, 1974; Rose, 1978; Shepherd et al., 2010).

The most extensively sampled and studied mastrevirus species is *Maize streak virus*: the causal agent of maize streak disease (MSD), the most significant viral disease of maize (*Zea mays*) in Africa (Harkins et al., 2009; Martin et al., 2001; Monjane et al., 2011; Mullineaux et al., 1984; Oluwafemi et al., 2014; Oppong et al., 2015; Owor et al., 2007a; Pande et al., 2017; Varsani et al., 2009; Varsani et al., 2008b). Whereas isolates of ten of the currently identified maize streak virus (MSV) strains, MSV-B, −C, −D, −E, −F, −G, −H, −I, −J and −K, have primarily been found in the field infecting uncultivated grass species, the only strain that is known to be associated with economically significant MSD is MSV-A (Martin et al., 2001). Besides MSV, one other mastrevirus has so-far been isolated from maize plants displaying severe MSD symptoms is maize streak Reunion virus (MSRV); a virus that been found infecting maize on La Réunion Island (Pande et al., 2012) and in China (Chen et al., 2015), and infecting uncultivated grasses (*Rottboellia* sp. and *Setaria barbata*) growing within maize fields in Nigeria (Oluwafemi et al., 2014).

The second most extensively sampled and studied AfSV species is *Panicum streak virus*, isolates of which have been sampled throughout much of Africa and on the islands of the south-west Indian Ocean (Varsani et al., 2009; Varsani et al., 2008a). Panicum streak virus (PanSV), as its name suggests, has largely been isolated from *Panicum* sp. However, the natural host range of PanSV also includes several other uncultivated grasses including *Brachiaria deflexa*, *Brachiaria plantaginea* (synonym: *Urochloa plantaginea*), *Ehrharta calycina*, *Panicum maximum* (synonym: *Urochloa maximum*) and *Panicum trichocladum* (Varsani et al., 2009; Varsani et al., 2008a). PanSVs are apparently as genetically diverse as MSVs and nine genetically distinct strains (PanSV-A through −K) have been documented so far (Varsani et al., 2008a).

Other mastreviruses which have been recovered from uncultivated grasses in Africa are assigned to the species *Axonopus compressus streak virus* (Oluwafemi et al., 2014), *Eragrostis minor streak virus* (Martin et al., 2011), *Eragrostis streak virus* (Shepherd et al., 2008b) and *Urochloa streak virus* (Oluwafemi et al., 2008).

Isolates of five AfSV species, *Saccharum streak virus* (Lawry et al., 2009), *Sugarcane streak virus* (Hughes et al., 1993; Shepherd et al., 2008b), *Sugarcane streak Egypt virus* (Bigarre et al., 1999), *Sugarcane streak Reunion virus* (Bigarre et al., 1999; Shepherd et al., 2008b), and *Sugarcane white streak virus* (Candresse et al., 2014) have predominantly been found infecting sugarcane. SSRV and SSV have, however, also been recovered from uncultivated grasses. Additionally, sugarcane chlorotic streak virus (Yahaya et al., 2017), which is currently unclassified but is a proposed new mastrevirus species, has recently been identified infecting sugarcane in Nigeria.

Geographically, MSRV, PanSV and MSV have the broadest known geographical ranges of the various AfSV species. In the case of PanSV and MSV, however, these viruses have been more extensively sampled than other AfSV species and, possibly as a consequence of this, isolates have been identified infecting a variety of host species in several countries throughout Africa and on various Indian Ocean islands. Possibly due to both insufficient sampling, and a bias towards the sampling of cultivated Poaceae species, most of the other AfSV species have only ever been found infecting particular hosts in a single country.

Here we undertook a large-scale study to better elucidate the diversity, host ranges and geographical distributions of Poaceae-infecting mastreviruses in Africa and its adjacent Indian Ocean and Atlantic islands.

2. Material and methods

2.1. Sampling, DNA extraction and recovery of mastrevirus genomes

Poaceae samples displaying foliar striation/streak symptoms that are typical of monocot-infecting mastrevirus infections were collected from Gran Canaria (n = 34), Zimbabwe (n = 5), Namibia (n = 18), Nigeria (n = 4), La Réunion (n = 10), Mauritius (n = 40), South Africa (n = 109) and Kenya (n = 94). Total genomic DNA was extracted from dried leaf material of each sample using either an Extract-N-Amp™ Plant kit (Sigma-Aldrich, USA) as described in Shepherd et al. (2008a) or using a GF-1 nucleic acid extraction kit (Vivantis Technologies, Malaysia) according to the manufacturer's instructions. Circular viral DNA was amplified by rolling circle amplification (RCA) from total genomic DNA using an Illustra TempliPhi Amplification Kit (GE Healthcare, USA). Full viral genomes were isolated using either restriction digest or polymerase chain reaction (PCR). For each restriction digestion reaction 1.5 µl of RCA DNA was digested using either *Bam*HI, *Kpn*I or *Hind*III to yield unit length ~2.7 kb genomes. For those samples where restriction digests were unsuccessful, PCR amplification was performed using 0.5 µl RCA DNA, KAPA HiFi hotstart DNA polymerase (Kapa Biosystems, USA) and the degenerate primer pair: forward 5′-GAN TTG GTC CGC AGT GTA GA-3′, reverse 5′-GTA CCG GWA AGA CMW CYT GG-3′ (Hadfield et al., 2012). The PCR was carried out under the following thermal cycling conditions: 94 °C for 3 min, 25x [98 °C (3 min), 52 °C (30 s), 72 °C (2.45 min)] and 72 °C for 3 min. Resulting PCR products and restricted products were purified using the quick-spin PCR Product Purification Kit (iNtRON Biotechnology, Korea) and respectively ligated to the vectors pJET1.2 (Fermentas, USA) and pGEM3Zf(+) (Promega, USA). The resulting mastrevirus genome clones were Sanger sequenced by primer walking by Macrogen Inc. (Korea). Full mastrevirus genomes were assembled from overlapping Sanger sequencing reads using DNA Baser sequence assembler V4 (Heracle BioSoft, Romania) with further manual editing carried out using MEGA 5.2 (Tamura et al., 2011).

2.2. Molecular identification on host species

Host species were identified for each sample using polymerase chain reaction (PCR) amplification, sequencing and phylogenetic analysis of a portion of the chloroplast *ndh*F gene (~1.1 kb) using the primer pair ndhF972F (5′-GT CTCAATTGGGTTATATGAT-3′) ndhF2110R (5′-CCCCCTAYATATTTGATACCTT-3′) as described in Giussani et al. (2001) and Olmstead and Reeves (1995). 4 µl of genomic DNA together with Kapa HIFI hotstart DNA polymerase were subjected to the following thermal cycling conditions: 94 °C for 3 min, 25x [98 °C (20 s), 50 °C (15 s), 72 °C (1 min)], final extension of 72 °C for 3 min. PCR products were purified using PCR quick-spin Purification Kit (iNtRON Biotechnology Inc, Korea) Sanger sequenced by Macrogen Inc. (Korea).

2.3. Analysis of mastrevirus genomes

Full genome sequence identities were calculated using SDT1.2 (Muhire et al., 2014). A full genome monocot-infecting mastrevirus dataset was assembled which included genome sequences from this study together with all those available in GenBank (downloaded 01/4/2017). These sequences were all linearized at the start of the movement protein gene and then aligned using MUSCLE (Edgar, 2004), as implemented in MEGA 5.2 (Tamura et al., 2011). The alignments were used to infer Neighbor joining trees in MEGA 5.2 (Tamura et al., 2011) with Jukes-Cantor substitution model and 1000 bootstrap iterations. Branches with < 60% support were collapsed with TreeGraph 2 (Stover and Muller, 2010). Bipartite plots were generated using R bipartite package (Dormann et al., 2009).

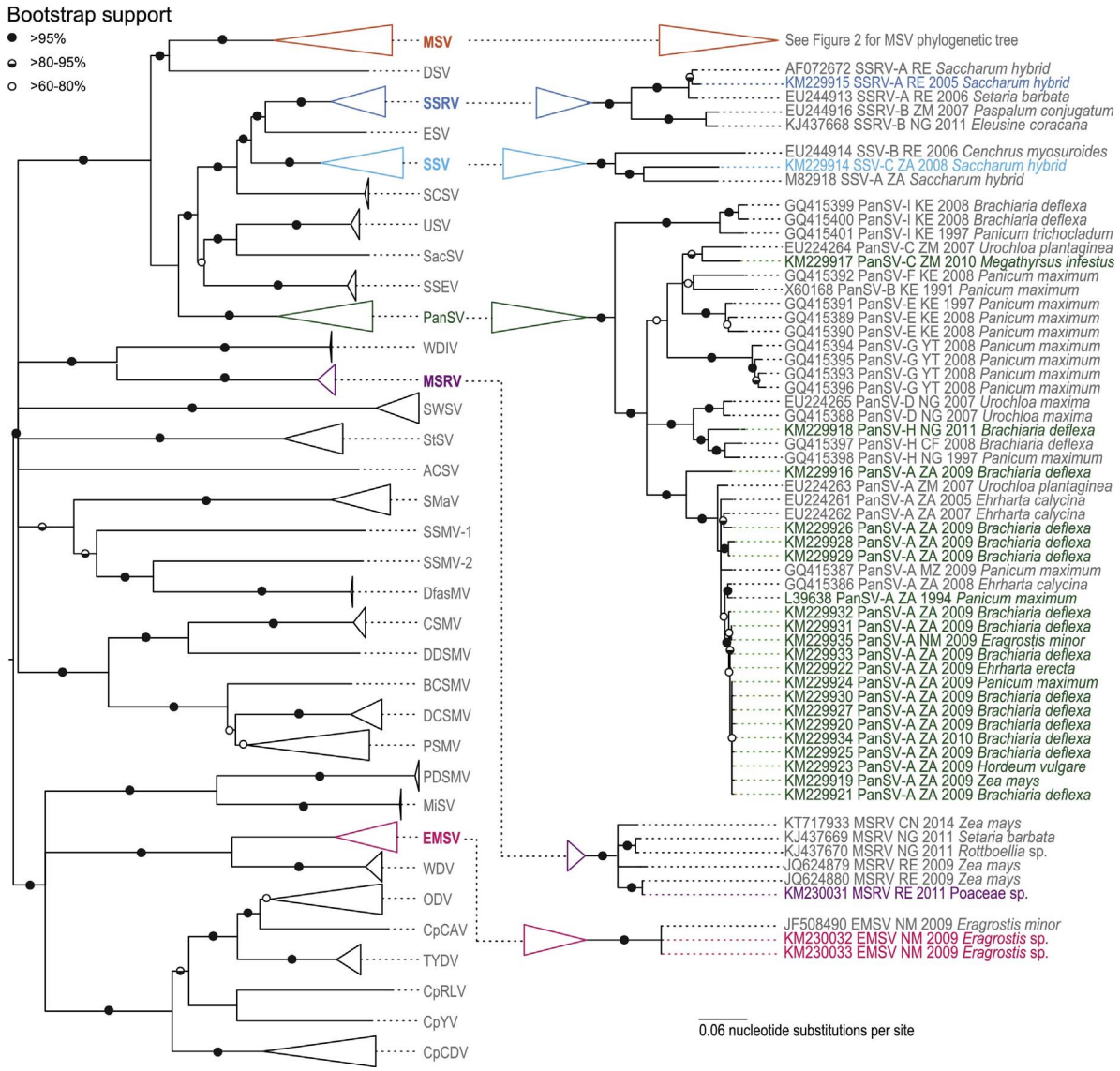

Fig. 1. Condensed Neighbor-joining phylogenetic tree of the mastreviruses (left) and expanded clades showing taxa sampled as part of this study (coloured font) together with those available in GenBank (taxa with grey font). The expanded clade for MSV is presented in Fig. 2. All branches with < 60% bootstrap support have been collapsed. Branches with > 60–80% support are shown with open circles, > 80–95% with half-filled circles and > 95% with filled circles.

3. Results and discussion

3.1. Identification of mastrevirus sequences

A total of 134 full AfSV genomes were recovered from symptomatic uncultivated and cultivated grasses representing 22 Poaceae genera (*Bambusa, Brachiaria, Bromus, Cenchrus, Chlorocalymma, Digitaria, Ehrharta, Eleusine, Eragrostis, Hordeum, Hyparrhenia, Lolium, Megathyrsus, Oplismenus, Panicum, Paspalum, Polypogon, Saccharum, Sclerochloa, Setaria, Urochloa* and *Zea*) (Supplementary Table 1). These 134 AfSVs were recovered from grasses sampled from five African countries (Kenya, Namibia, Nigeria, South Africa and Zimbabwe) and three islands (Gran Canaria, Mauritius and La Réunion) (Supplementary Table 1). In accordance with recommendations for mastrevirus classification (Muhire et al. 2013), full genome pairwise comparisons using SDT 1.2 (Muhire et al., 2014) indicated that the 134 genomes each belonged to one of five established species: *Eragrostis minor streak virus* (n = 2), *Maize streak Reunion virus* (n = 1), *Maize streak virus* (n = 109), *Panicum streak virus* (n = 20), *Sugarcane streak Reunion virus* (n = 1) and *Sugarcane streak virus* (n = 1) (Figs. 1 and 2). Further, isolates which shared > 94% identity with previously described strains within these species were assigned to these previously established strains. Accordingly, the new MSV isolates from this study were assigned to the established strains MSV-B (n = 33), MSV-C (n = 33), MSV-D (n = 5), MSV-E (n = 3), MSV-F (n = 31), MSV-G (n = 1), MSV-J (n = 1) and MSV-K (n = 2) (Fig. 2). The new PanSV isolates were assigned to the established strains PanSV-A (n = 18), PanSV-C (n = 1) and PanSV-H (n = 1) (Fig. 1). The single recovered sugarcane streak Reunion virus (SSRV) isolate was assigned to the established strain, SSRV-A (Fig. 1). The single isolate of sugarcane streak virus (SSV) that was recovered shares < 94% identity with other known SSV isolates and was therefore tentatively assigned to a new strain, SSV-C (Fig. 1). Considering that to-date only three SSV isolates have been characterised, and that each of these belong to a different strain, it is apparent that SSV displays a relatively high degree of diversity. The SSV-A and −C isolates were both from sugarcane plants sampled in South Africa whereas the SSV-B isolate was from a *Cenchrus echinatus* plant sampled in La Réunion (Supplementary Table 2).

3.2. Identification of mastrevirus hosts

The genus and/or species of all but fifteen uncultivated grasses were

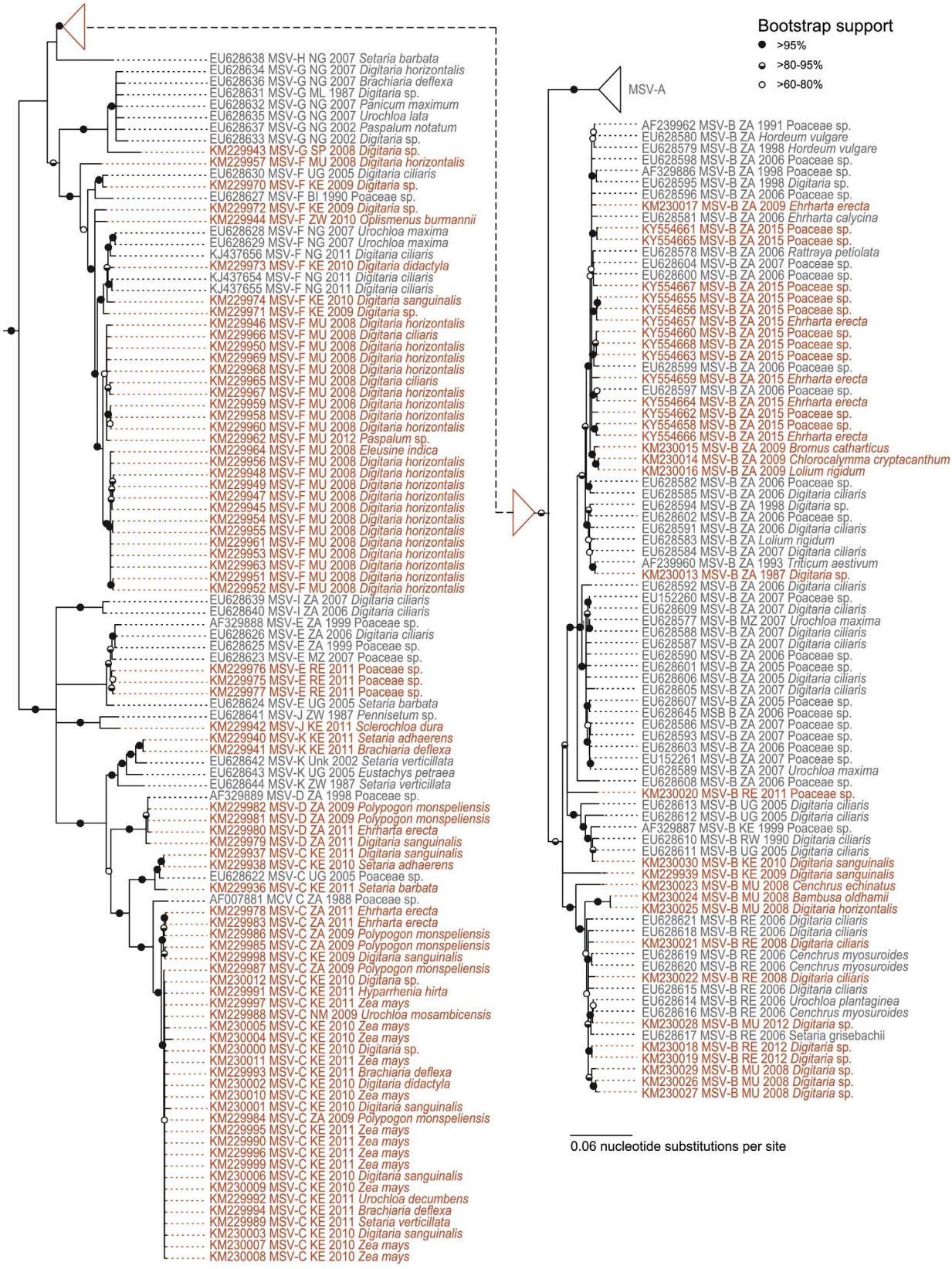

Fig. 2. Neighbor-joining phylogenetic trees of MSV taxa sampled as part of this study (orange font) together with those available in GenBank (taxa with grey font). All branches with < 60% bootstrap support have been collapsed. Branches with > 60–80% support are shown with open circles, > 80–95% with half-filled circles and > 95% with filled circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

identified for the mastrevirus positive samples via phylogenetic analysis of choloroplast *ndh*F gene sequences (Supplementary Table 1). This revealed several previously unknown hosts both for the various detected mastrevirus species, and for the individual strains within these species. Six new species (*Brachiaria plantaginea, Ehrharta erecta, Eragrostis minor, Hordeum vulgare, Panicum infestum* (synonym:

Megathyrsus infestus) and *Zea mays*) from four subfamilies (Chloridoideae, Ehrhartoideae, Panicoideae and Pooideae) were identified as PanSV hosts. The grass species from which the PanSV strain has most commonly been isolated is *B. deflexa* (n = 17) (Fig. 1; Supplementary Table 2).

3.3. Host and geographical range of mastreviruses

Thirteen new host species in four different subfamilies (Bambusoideae, Chloridoideae, Panicoideae and Pooideae) were identified for MSV (Fig. 2; Supplementary Table 2). Focusing on individual MSV strains, 12 new hosts (*B. deflexa, Digitaria didactyla, D. sanguinalis, E. erecta, Hyparrhenia hirta, Polypogon monspeliensis, Setaria verticillata* (synonym: *Setaria adhaerens*), *S. barbata, S. verticillata, Brachiaria decumbens* (synonym: *Urochloa decumbens*), *U. mosambicensis* and *Z. mays*) were identified for MSV-C (Fig. 2; Supplementary Table 2), seven (*D. horizontalis*, *D. didactyla*, *D. sanguinalis*, *Eleusine indica*, *Oplismenus burmannii* and a *Paspalum* sp.) for MSV-F (previously it had only been found in *D. ciliaris* and *Urochola maxima*), three for MSV-K (*Sclerochloa dura, S. adhaerens* and *B. deflexa*), three for MSV-D (*D. sanguinalis, E. erecta* and *P. monspeliensis*) and eight for MSV-B (*Bambusa oldhamii, Bromus catharticus, Cenchrus echinatus, Chlorocalymma cryptacanthum, D. horizontalis, D. sanguinalis* and *E. erecta*) (Fig. 2; Supplementary Table 2). EMSV was recovered from an *Eragrostis* sp. sampled in Namibia, MSRV from an unidentified grass from La Réunion, SSRV-A from a *Saccharum* hybrid on La Réunion and SSV-C from a *Saccharum* hybrid from South Africa (Fig. 1; Supplementary Table 1).

It has been suggested that MSV-A likely has a broader natural host range than MSV-B since MSV-A isolates have been recovered from grass species spanning eight genera and MSV-B isolates have only ever been isolated from grass species spanning six genera, (Varsani et al. 2008b). However, analysis of a significantly larger dataset here − which includes all of the MSV-A (n = 695) and B (n = 85) isolates recovered to-date from cultivated and uncultivated grass species − suggests that MSV-A has a host range of at least ten Poacece genera whereas that of MSV-B encompasses at least fourteen genera: implying that MSV-B may in fact have a broader natural host range than MSV-A (Fig. 3A; Supplementary Table 2).

Monocot-infecting mastreviruses have been sampled extensively and from a wide range of locations throughout Africa over the last decade. This large dataset offers the opportunity to investigate potential patterns in the geographic distributions of the AfSVs. To gain an overview of the known geographical ranges of the different AfSV species, we mapped all of the fully sequenced AfSV genomes (both those recovered in this study and those available in GenBank) to their countries of origin (Fig. 3B). It is worth reiterating here that, despite the large numbers of genomes analysed, there are inescapable sampling biases in this dataset in terms of both the countries where sampling has been carried out, and in the numbers of samples that have been collected from those countries. For example, sampling undertaken in Burkina Faso, Cameroon and Ghana was solely of maize plants displaying MSD symptoms and therefore it is not surprising that only the maize adapted MSV-A strain has so-far been found in these countries. Although these sampling biases make it impossible to meaningfully compare AfSV diversity between countries, they do not bias our intent to indicate how the new data presented here expands the known geographical ranges of particular AfSV species and strains.

For the first time MSV (MSV-G strain) has been definitively identified on the Atlantic island of Gran Canaria (Fig. 2). This is also the first record of an AfSV having been found north-west of the Sahara desert. MSV-G has previously only been recorded in Mali and Nigeria (Varsani et al., 2008b) despite the host species (*Digitaria* sp., *Panicum* sp., *Paspalum* sp., and *Brachiaria* sp.) having been sampled multiple times elsewhere in Africa. This strain therefore appears to have a distribution which is restricted to West Africa and the island of Gran Canaria.

An additional 37 MSV isolates (Fig. 1; Supplementary Table 1) were recovered from uncultivated grasses and maize sampled in Kenya. Among these the MSV strains MSV-C, −F, −J and −K were identified for the first time in Kenya.

Kenya is the only region where MSV-C has been found infecting maize. This is particularly noteworthy because, while maize is known to be sensitive to MSV-C infection under laboratory conditions (Schnippenkoetter et al., 2001), MSV-C has never before been detected in any of the over 600 MSD-infected maize samples (Supplementary Table 2) analysed in previous studies (Harkins et al., 2009; Martin et al., 2001; Monjane et al., 2011; Oluwafemi et al., 2014; Oppong et al., 2015; Owor et al., 2007a; Owor et al., 2007b; Pande et al., 2012; Pande et al., 2017; Schnippenkoetter et al., 2001; Shepherd et al., 2008a; Yahaya et al., 2017). All twelve of the Kenyan MSV-C isolates from maize cluster phylogenetically within a clade of isolates all sharing > 98% nucleotide sequence identity with isolates found in uncultivated grass species from Kenya (26 isolates in this clade) (Fig. 2), Namibia (a single isolate, KM229988, from *U. mosambicensis*) and South Africa (four isolates, KM229984 − KM229987, from *P. monspeliensis*). It is of some concern that this widespread MSV-C lineage is apparently adapted to infecting maize as this could mean that it is in the process of emerging as a new pathogen in this important African crop species.

The islands off the south-west Indian Ocean islands of Africa have previously been shown to be hotspots of geminivirus diversity (Lefeuvre et al., 2007; Peterschmitt et al., 1996; Shepherd et al., 2008b). Mastreviruses from uncultivated grasses have previously been sampled from the islands of La Réunion and Mayotte (Shepherd et al., 2008b; Varsani et al., 2009; Varsani et al., 2008a; Varsani et al., 2008b). La Réunion has proven to be a hub of diversity with four mastrevirus species identified there (MSRV, MSV, SSRV and SSV) and we therefore undertook sampling on the neighbouring island of Mauritius. Here, from 32 grass samples spanning five Poaceae genera, we recovered MSV-B and −F isolates but no other AfSV species (Fig. 2; Supplementary Tables 1 and 2).

PanSV isolates were previously shown to display a higher degree of geographical structure than that displayed by MSV isolates sampled from uncultivated grasses (Varsani et al., 2009), indicating that MSV likely spreads throughout continental Africa more readily than does PanSV. With the addition here of more PanSV sequences from four additional countries it remains apparent that the geographical clustering of PanSV isolates is stronger than that of MSV isolates.

4. Conclusion

The AfSVs are the best characterised group of mastreviruses, with over 950 genomic sequences (Supplementary Table 2) having been recovered from infected grass samples from 21 countries in Africa and from four islands off the coast of Africa. This large set of sequences indicates that the two most sampled AfSV species, MSV (n = 885) and PanSV (n = 43), are widely distributed throughout Africa and its neighbouring islands, with MSV having a distribution which extends at least as far north and west as Gran Canaria Island, as far east as Mauritius island and as far south as South Africa (Fig. 3B). It possible, therefore, that additional sampling will reveal that the distributions of many of the other AfSV species are similarly extensive. It is also highly likely that there are asymptomatic grasses infected with mastreviruses and have not be sampled with the current symptom-biased approaches. For example, a study by Prendeville et al. (2012) demonstrated (using probe based assays) that of a variety of samples they tested for potyviruses, 80% of the hosts were asymptomatic. Thus fine scale sampling of all grasses within an ecosystem coupled with plant viral metagenomics (Roossinck et al., 2015) approaches will yield robust data on the true host ranges of these viruses and their population dynamics for better informed management strategies.

Nonetheless, besides better identification of the geographical ranges of MSV and PanSV, the greater degree of MSV and PanSV sampling in this study revealed that both species also have very broad and largely

A

Grass Subfamily	Grass Genus	ACSV	EMSV	ESV	MSRV	MSV A	MSV B	MSV C	MSV D	MSV E	MSV F	MSV G	MSV H	MSV I	MSV J	MSV K	PanSV A	PanSV B	PanSV C	PanSV D	PanSV E	PanSV F	PanSV G	PanSV H	PanSV I	SacSV	SCSV	SSEV	SSRV A	SSRV B	SSV A	SSV B	SSV C	SWSV A	SWSV B	SWSV C	USV	Total
Bambusoideae	*Bambusa*						1																															1
Chloridoideae	*Danthoniopsis*						1																															1
	Eleusine										1																			1							1	3
	Eragrostis		3	1		1											1																					6
	Eustachys					1										1																							2
Ehrhartoideae	*Ehrharta*					1	6	2	1								4																							14
Panicoideae	*Axonopus*	1				1																																2
	Brachiaria					1	1	3				2				1	14		1				2	2														8	35
	Cenchrus					1	4																									1						6
	Chlorocalymma						1																															1
	Digitaria					25	30	8	1	2	32	4		2																									104
	Hyparrhenia							1																															1
	Oplismenus										1											1																	2
	Panicum						2				2	1					3	1	1	2	3		4	1	1														22
	Paspalum										1	1																		1									3
	Pennisetum														1																								1
	Rottboellia				1	1																																	2
	Saccharum					30																				1	12	5	2		1		1	2	3	1			58
	Setaria				1		1	4		1			1			3													1										12
	Urochloa							1	1																													2	4
	Zea				3	627	2	12									1																							645
Pooideae	*Bromus*						1																																1
	Hordeum						2										1																							3
	Lolium						2																																2
	Polypogon							4	2																														6
	Sclerochloa														1																								1
	Triticum						1																																1
	Unknown host				1	5	30	1		5	1																												43
	Total	1	3	1	6	695	85	36	5	8	38	8	1	2	2	5	24	1	2	2	3	1	4	3	3	1	12	5	3	2	1	1	1	2	3	1	11	

B

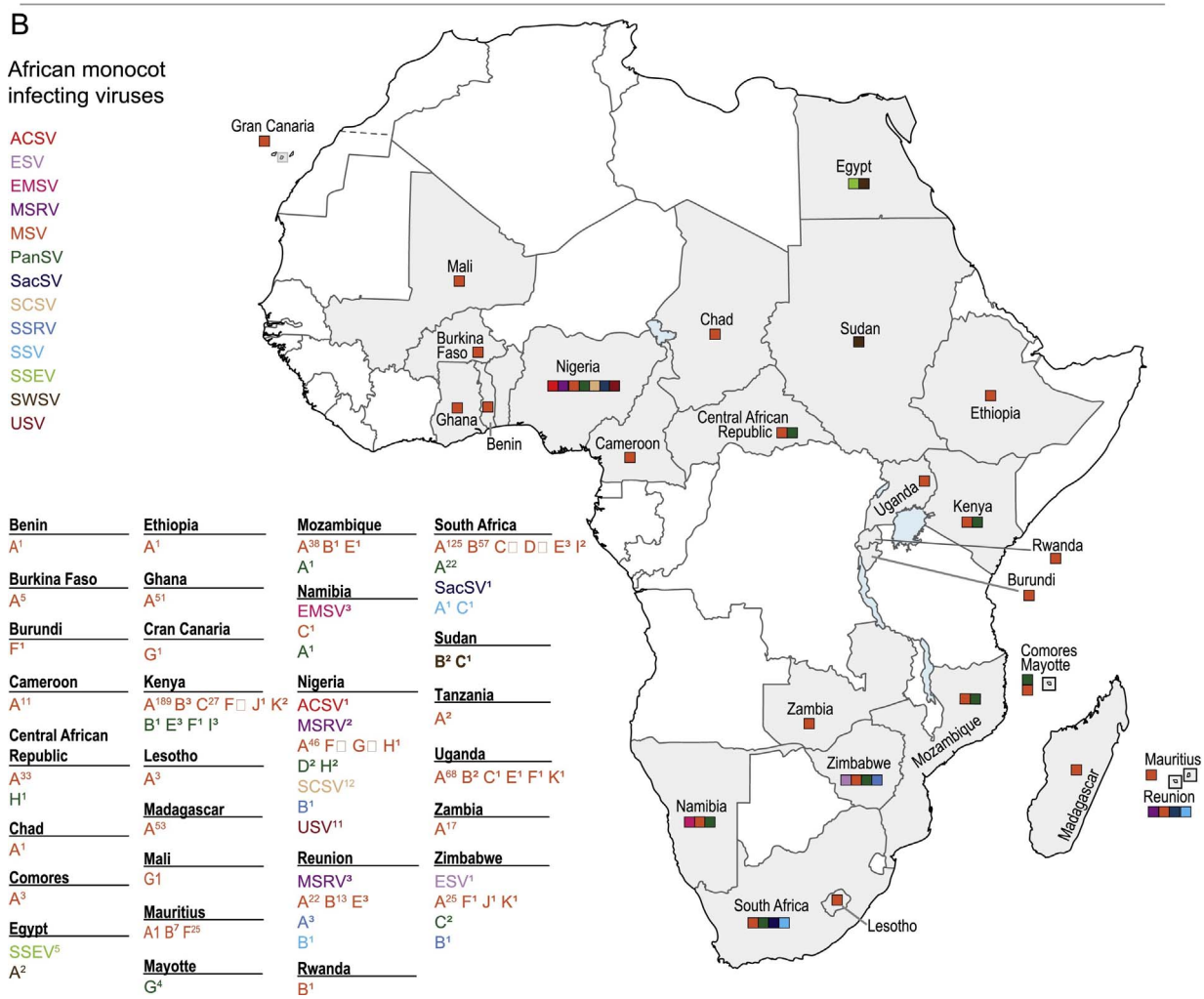

Fig. 3. **(A)** Summary of the genomes recovered of African monocot-infecting mastreviruses and their hosts (including samples from this study and those with associated full genome sequences available in GenBank). Number of genomes belonging to each mastrevirus species/strains is colour coded and total genomes assigned to each mastrevirus species/strains are at the bottom whereas total genomes from particular hosts are provided on the left. **(B)** Map of Africa and surrounding islands indicating countries and territories where AfSVs have been sampled (including samples from this study and those with associated full genome sequences available in GenBank). Different AfSV species are represented by colours in the key. Letters in the country sample list represent the strain of the virus species. The number of samples obtained for each species and strain from the various countries is indicated by the superscript number next to species acronym and/or strain letter.

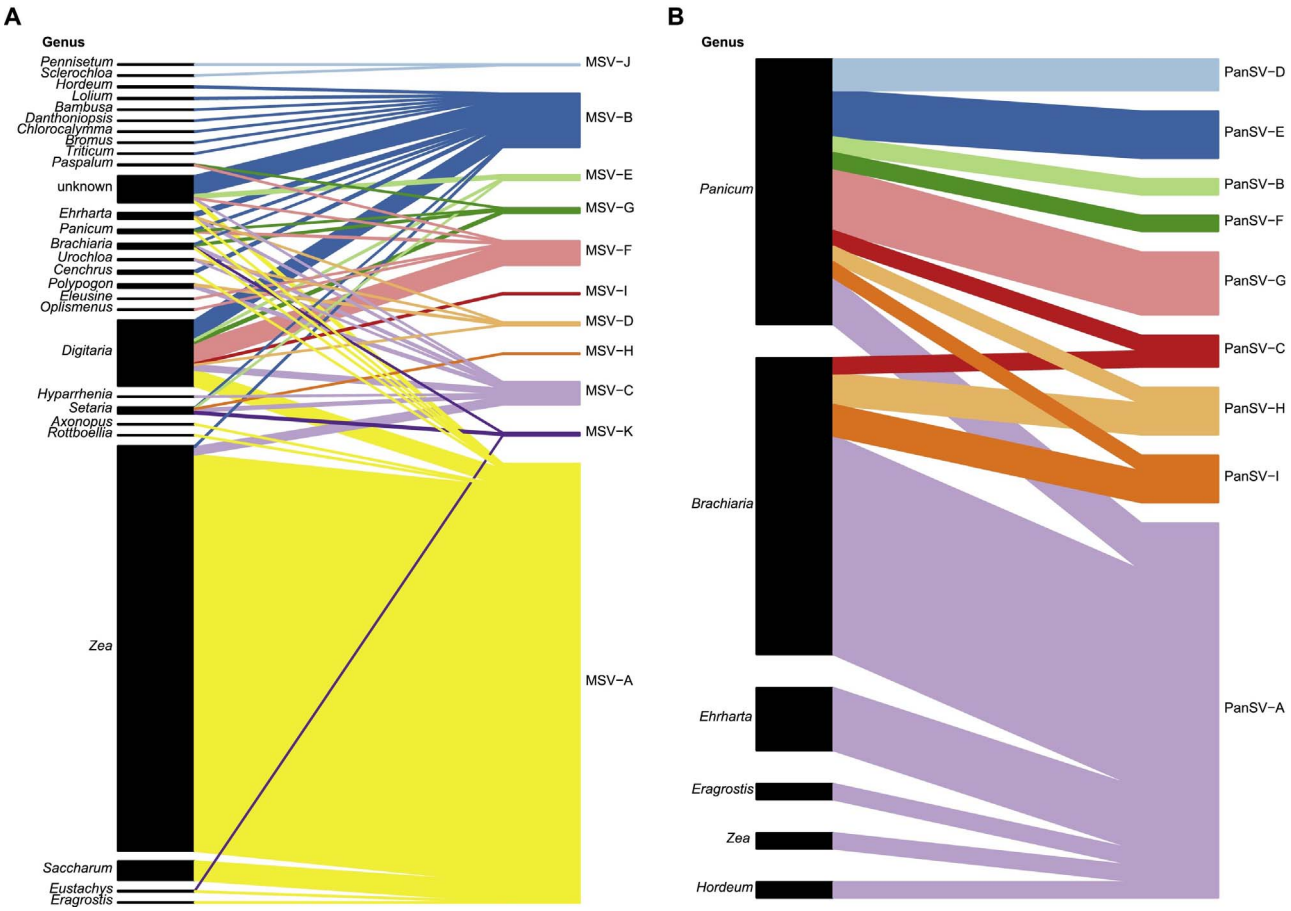

Fig. 4. Tanglegrams illustrating the known natural host ranges (at the genus level) of individual strains of the two most sampled African mastrevirus species: (A) maize streak virus (B) panicum streak virus.

overlapping host ranges (Fig. 4). Further, individual well sampled PanSV and MSV strains were found to have broader host ranges in the field than was previously appreciated, with a total of 20 new host genera identified for the various strains (Fig. 3A). It is plausible too that the host ranges of the other less thoroughly sampled AfSVs could be similarly broad but additional sampling will be required to verify whether this is the case.

The overwhelming majority of sampled AfSV sequences have been classified as belonging to the maize adapted MSV-A strain: a strain which evidently emerged as a serious maize pathogen ~150 years ago as a result of recombination events among MSV strains that were likely adapted to infecting uncultivated grasses (Harkins et al., 2009; Monjane et al., 2011; Varsani et al., 2008b). Given the extent of recombination detected in MSVs (Harkins et al., 2009; Monjane et al., 2011; Owor et al., 2007a; Pande et al., 2017; Varsani et al., 2008b), it may be prudent to monitor these and other AfSVs in uncultivated host species in order to identify new strains, or new lineages of old strains, that may pose a significant future threat to crop cultivation in Africa.

In this regard it is significant that, prior to this study, only MSRV and the MSV-A strain had ever been recovered from maize displaying severe MSD symptoms. Here we recovered twelve MSV-C genomes from MSD afflicted maize plants sampled in Kenya that all fell phylogenetically within a clade of closely related MSV-C variants. The variants within this clade have a geographical distribution ranging at the very least from Kenya to Namibia and South Africa. Determining the actual geographical distributions of the members of this lineage, and determining their degree of pathogenicity in maize is a priority as this lineage may represent a serious emerging threat to maize cultivation throughout much of Africa.

GenBank accession numbers

KM229914–KM230033; KY554655–KY554668.

Acknowledgements

DP was supported by the International Centre for Genetic Engineering and Biotechnology, The World Academy of Sciences and the Kenya National Commission for Science, Technology and Innovation (Grant No. CRP.BMB/ZAF09-02). PL and JML were supported by the European Union (FEDER) and the *Région Réunion* (Grant No. GURDT I2016-1731-0006632). AV and DPM were supported by South African National Research Foundation. The funding sources had no role in the study design; the collection, analysis and interpretation of the data; the writing of the report or the decision to submit the article for publication.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.virusres.2017.07.001.

References

Bigarre, L., Salah, M., Granier, M., Frutos, R., Thouvenel, J., Peterschmitt, M., 1999. Nucleotide sequence evidence for three distinct sugarcane streak mastreviruses. Arch. Virol. 144 (12), 2331–2344.

Candresse, T., Filloux, D., Muhire, B., Julian, C., Galzi, S., Fort, G., Bernardo, P., Daugrois, J.H., Fernandez, E., Martin, D.P., Varsani, A., Roumagnac, P., 2014. Appearances can be deceptive: revealing a hidden viral infection with deep sequencing in a plant quarantine context. PLoS One 9 (7), e102945.

Chen, S., Huang, Q., Wu, L., Qian, Y., 2015. Identification and characterization of a

maize-associated mastrevirus in China by deep sequencing small RNA populations. Virol. J. 12, 156.
Dormann, C.F., Fruend, J., Bluethgen, N.B.G., 2009. Indices, graphs and null models: analyzing bipartite ecological networks. Open Ecol. J. 2, 7–24.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32 (5), 1792–1797.
Giussani, L.M., Cota-Sánchez, J.H., Zuloaga, F.O., Kellogg, E.A., 2001. A molecular phylogeny of the grass subfamily Panicoideae (*Poaceae*) shows multiple origins of C4 photosynthesis. Am. J. Bot. 88 (11), 1993–2012.
Hadfield, J., Thomas, J.E., Schwinghamer, M.W., Kraberger, S., Stainton, D., Dayaram, A., Parry, J.N., Pande, D., Martin, D.P., Varsani, A., 2012. Molecular characterisation of dicot-infecting mastreviruses from Australia. Virus Res. 166 (1-2), 13–22.
Harkins, G.W., Martin, D.P., Duffy, S., Monjane, A.L., Shepherd, D.N., Windram, O.P., Owor, B.E., Donaldson, L., van Antwerpen, T., Sayed, R.A., Flett, B., Ramusi, M., Rybicki, E.P., Peterschmitt, M., Varsani, A., 2009. Dating the origins of the maize-adapted strain of maize streak virus, MSV-A. J. Gen. Virol. 90 (Pt 12), 3066–3074.
Hughes, F.L., Rybicki, E.P., Kirby, R., 1993. Complete nucleotide sequence of sugarcane streak Monogeminivirus. Arch. Virol. 132 (1-2), 171–182.
Lawry, R., Martin, D.P., Shepherd, D.N., van Antwerpen, T., Varsani, A., 2009. A novel sugarcane-infecting mastrevirus from South Africa. Arch. Virol. 154 (10), 1699–1703.
Lefeuvre, P., Martin, D.P., Hoareau, M., Naze, F., Delatte, H., Thierry, M., Varsani, A., Becker, N., Reynaud, B., Lett, J.M., 2007. Begomovirus 'melting pot' in the south-west Indian Ocean islands: molecular diversity and evolution through recombination. J. Gen. Virol. 88 (Pt 12), 3458–3468.
Martin, D.P., Shepherd, D.N., 2009. The epidemiology, economic impact and control of maize streak disease. Food Secur. 1 (3), 305–315.
Martin, D.P., Willment, J.A., Billharz, R., Velders, R., Odhiambo, B., Njuguna, J., James, D., Rybicki, E.P., 2001. Sequence diversity and virulence in Zea mays of Maize streak virus isolates. Virology 288 (2), 247–255.
Martin, D.P., Linderme, D., Lefeuvre, P., Shepherd, D.N., Varsani, A., 2011. Eragrostis minor streak virus: an Asian streak virus in Africa. Arch. Virol. 156 (7), 1299–1303.
Monjane, A.L., Harkins, G.W., Martin, D.P., Lemey, P., Lefeuvre, P., Shepherd, D.N., Oluwafemi, S., Simuyandi, M., Zinga, I., Komba, E.K., Lakoutene, D.P., Mandakombo, N., Mboukoulida, J., Semballa, S., Tagne, A., Tiendrebeogo, F., Erdmann, J.B., van Antwerpen, T., Owor, B.E., Flett, B., Ramusi, M., Windram, O.P., Syed, R., Lett, J.M., Briddon, R.W., Markham, P.G., Rybicki, E.P., Varsani, A., 2011. Reconstructing the history of maize streak virus strain a dispersal to reveal diversification hot spots and its origin in southern Africa. J. Virol. 85 (18), 9623–9636.
Muhire, B., Martin, D.P., Brown, J.K., Navas-Castillo, J., Moriones, E., Zerbini, F.M., Rivera-Bustamante, R., Malathi, V.G., Briddon, R.W., Varsani, A., 2013. A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus Mastrevirus (family Geminiviridae). Arch. Virol. 158 (6), 1411–1424.
Muhire, B.M., Varsani, A., Martin, D.P., 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 9 (9), e108277.
Mullineaux, P.M., Donson, J., Morris-Krsinich, B.A., Boulton, M.I., Davies, J.W., 1984. The nucleotide sequence of maize streak virus DNA. EMBO J. 3 (13), 3063–3068.
Olmstead, R.G., Reeves, P.A., 1995. Evidence for the polyphyly of the Scrophulariaceae based on chloroplast rbcL and ndhF sequences. Ann. Mo. Bot. Garden 82 (2), 176–193.
Oluwafemi, S., Varsani, A., Monjane, A.L., Shepherd, D.N., Owor, B.E., Rybicki, E.P., Martin, D.P., 2008. A new African streak virus species from Nigeria. Arch. Virol. 153 (7), 1407–1410.
Oluwafemi, S., Kraberger, S., Shepherd, D.N., Martin, D.P., 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 (10), 2765–2770.
Oppong, A., Offei, S.K., Ofori, K., Adu-Dapaah, H., Lamptey, J.N., Kurenbach, B., Walters, M., Shepherd, D.N., Martin, D.P., Varsani, A., 2015. Mapping the distribution of maize streak virus genotypes across the forest and transition zones of Ghana. Arch. Virol. 160 (2), 483–492.
Owor, B.E., Martin, D.P., Shepherd, D.N., Edema, R., Monjane, A.L., Rybicki, E.P., Thomson, J.A., Varsani, A., 2007a. Genetic analysis of maize streak virus isolates from Uganda reveals widespread distribution of a recombinant variant. J. Gen. Virol. 88 (Pt 11), 3154–3165.
Owor, B.E., Shepherd, D.N., Taylor, N.J., Edema, R., Monjane, A.L., Thomson, J.A., Martin, D.P., Varsani, A., 2007b. Successful application of FTA Classic Card technology and use of bacteriophage phi29 DNA polymerase for large-scale field sampling and cloning of complete maize streak virus genomes. J. Virol. Methods 140 (1-2), 100–105.
Pande, D., Kraberger, S., Lefeuvre, P., Lett, J.M., Shepherd, D.N., Varsani, A., Martin, D.P., 2012. A novel maize-infecting mastrevirus from La Reunion Island. Arch. Virol. 157 (8), 1617–1621.
Pande, D., Madzokere, E., Hartnady, P., Kraberger, S., Hadfield, J., Rosario, K., Jaschke, A., Monjane, A.L., Owor, B.E., Dida, M.M., Shepherd, D.N., Martin, D.P., Varsani, A., Harkins, G.W., 2017. The role of Kenya in the trans-African spread of maize streak virus strain A. Virus Res. 232, 69–76.
Peterschmitt, M., Granier, M., Frutos, R., Reynaud, B., 1996. Infectivity and complete nucleotide sequence of the genome of a genetically distinct strain of maize streak virus from Reunion Island. Arch. Virol. 141 (9), 1637–1650.
Prendeville, H.R., Ye, X.H., Morris, T.J., Pilson, D., 2012. Virus infections in wild plant populations are both frequent and often unapparent. Am. J. Bot. 99 (6), 1033–1042.
Roossinck, M.J., Martin, D.P., Roumagnac, P., 2015. Plant virus metagenomics: advances in virus discovery. Phytopathology 105 (6), 716–727.
Rose, D.J.W., 1974. Epidemiology of maize streak disease in relation to population-densities of cicadulina spp. Ann. Appl. Biol. 76 (2), 199–207.
Rose, D.J.W., 1978. Epidemiology of maize streak disease. Annu. Rev. Entomol. 23, 259–282.
Schnippenkoetter, W.H., Martin, D.P., Hughes, F.L., Fyvie, M., Willment, J.A., James, D., von Wechmar, M.B., Rybicki, E.P., 2001. The relative infectivities and genomic characterisation of three distinct mastreviruses from South Africa. Arch. Virol. 146 (6), 1075–1088.
Shepherd, D.N., Martin, D.P., Lefeuvre, P., Monjane, A.L., Owor, B.E., Rybicki, E.P., Varsani, A., 2008a. A protocol for the rapid isolation of full geminivirus genomes from dried plant tissue. J. Virol. Methods 149 (1), 97–102.
Shepherd, D.N., Varsani, A., Windram, O.P., Lefeuvre, P., Monjane, A.L., Owor, B.E., Martin, D.P., 2008b. Novel sugarcane streak and Sugarcane streak Reunion mastreviruses from southern Africa and La Reunion. Arch. Virol. 153 (3), 605–609.
Shepherd, D.N., Martin, D.P., Van Der Walt, E., Dent, K., Varsani, A., Rybicki, E.P., 2010. Maize streak virus: an old and complex 'emerging' pathogen. Mol. Plant Pathol. 11 (1), 1–12.
Stover, B.C., Muller, K.F., 2010. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. BMC Bioinf. 11, 7.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28 (10), 2731–2739.
Varsani, A., Oluwafemi, S., Windram, O.P., Shepherd, D.N., Monjane, A.L., Owor, B.E., Rybicki, E.P., Lefeuvre, P., Martin, D.P., 2008a. Panicum streak virus diversity is similar to that observed for maize streak virus. Arch. Virol. 153 (3), 601–604.
Varsani, A., Shepherd, D.N., Monjane, A.L., Owor, B.E., Erdmann, J.B., Rybicki, E.P., Peterschmitt, M., Briddon, R.W., Markham, P.G., Oluwafemi, S., Windram, O.P., Lefeuvre, P., Lett, J.M., Martin, D.P., 2008b. Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. J. Gen. Virol. 89 (9), 2063–2074.
Varsani, A., Monjane, A.L., Donaldson, L., Oluwafemi, S., Zinga, I., Komba, E.K., Plakoutene, D., Mandakombo, N., Mboukoulida, J., Semballa, S., Briddon, R.W., Markham, P.G., Lett, J.M., Lefeuvre, P., Rybicki, E.P., Martin, D.P., 2009. Comparative analysis of Panicum streak virus and Maize streak virus diversity, recombination patterns and phylogeography. Virol. J. 6 (1), 194.
Yahaya, A., Dangora, D.B., Alegbejo, M.D., Kumar, P.L., 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 (2), 597–602.