


Capulavirus and *Grablovirus*: two new genera in the family *Geminiviridae*

Arvind Varsani^{1,2}  · Philippe Roumagnac³ · Marc Fuchs⁴ · Jesús Navas-Castillo⁵ · Enrique Moriones⁵ · Ali Idris⁶ · Rob W. Briddon⁷ · Rafael Rivera-Bustamante⁸ · F. Murilo Zerbini⁹ · Darren P Martin¹⁰

Received: 20 January 2017 / Accepted: 27 January 2017 / Published online: 17 February 2017
© Springer-Verlag Wien 2017

Abstract Geminiviruses are plant-infecting single-stranded DNA viruses that occur in most parts of the world. Currently, there are seven genera within the family *Geminiviridae* (*Becurtovirus*, *Begomovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocovirus* and *Turncurtovirus*). The rate of discovery of new geminiviruses has increased significantly over the last decade as a result of new molecular tools and approaches (rolling-circle amplification and deep sequencing) that allow for high-throughput workflows. Here, we report the establishment of two new genera: *Capulavirus*, with four new species (*Alfalfa leaf curl virus*, *Euphorbia caput-medusae latent virus*, *French bean severe leaf curl virus* and *Plantago lanceolata latent virus*), and *Grablovirus*, with one new species (*Grapevine red blotch virus*). The aphid species *Aphis craccivora* has been shown to be a vector for *Alfalfa leaf curl virus*, and

the treehopper species *Spissistilus festinus* is the likely vector of *Grapevine red blotch virus*. In addition, two highly divergent groups of viruses found infecting citrus and mulberry plants have been assigned to the new species *Citrus chlorotic dwarf associated virus* and *Mulberry mosaic dwarf associated virus*, respectively. These species have been left unassigned to a genus by the ICTV because their particle morphology and insect vectors are unknown.

Introduction

The family *Geminiviridae* includes plant-infecting viruses with circular ssDNA genomes that are encapsidated in virions comprised of 22 pentameric capsomeres arranged in a ~18- to 30-nm twinned icosahedral (or geminate)

✉ Arvind Varsani
arvind.varsani@asu.edu

✉ Darren P Martin
darrenpatrickmartin@gmail.com

¹ The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine and School of Life Sciences, Arizona State University, Tempe 85287-5001, AZ, USA
² Structural Biology Research Unit, Department of Clinical Laboratory Sciences, University of Cape Town, Cape Town 7701, South Africa
³ CIRAD-INRA-SupAgro, UMR BGPI, Campus International de Montpellier-Baillarguet, 34398 Montpellier Cedex-5, France
⁴ Section of Plant Pathology and Plant-Microbe Biology, School of Integrative Plant Science, Cornell University, New York State Agricultural Experiment Station, Geneva, NY 14456, USA

⁵ Instituto de Hortofruticultura Subtropical y Mediterránea “La Mayora”, Universidad de Málaga-Consejo Superior de Investigaciones Científicas (IHSM-UMA-CSIC), 29750 Algarrobo-Costa, Málaga, Spain

⁶ School of Plant Sciences, University of Arizona, Tucson, AZ 85721-0107, USA

⁷ National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan

⁸ Departamento de Ingeniería Genética, Centro de Investigación y de Estudios Avanzados del IPN (Cinvestav), Unidad Irapuato, 36821 Irapuato, GTO, Mexico

⁹ Dep. de Fitopatología/Bioagro, Universidade Federal de Viçosa, Viçosa, Minas Gerais 36570-900, Brazil

¹⁰ Computational Biology Group, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town 7925, South Africa

configuration [7, 41]. Geminiviruses have a near-global distribution and infect both monocotyledonous and dicotyledonous plants. Geminiviruses are often responsible for serious yield losses in economically important crops, including tomato, maize, cotton, chickpea and cassava [25]. Symptoms of geminivirus infections can include foliar crinkling, curling, yellowing, stunting, mosaic and/or striations.

The *Geminiviridae* Study Group of the International Committee on Taxonomy of Viruses (ICTV) has provided a series of guidelines for the classification of viruses within the seven currently recognised genera of the family: *Becurtovirus*, *Begomovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus* and *Turncurtovirus*. The criteria include host range, insect vector, genome organisation and genome-wide pairwise sequence identities [1, 11, 12, 26, 38, 39].

Over the last decade, the development of new molecular tools such as rolling-circle amplification (RCA) and high-throughput sequencing have greatly facilitated the discovery of novel geminiviruses in a range of cultivated and non-cultivated host species [29]. Accordingly, these techniques have accelerated the discovery of highly divergent geminivirus-like viruses. For some of these viruses, a large number of closely related full genome sequences have been determined. Analyses of these sequences (pairwise identities and phylogenetics), coupled with the identification of insect vectors and particle morphology, have allowed us to determine with great confidence that some of these divergent viruses should be classified as members of two new genera in the family *Geminiviridae*, which was approved by the Executive Committee of the ICTV:

- 1) *Capulavirus*, which will for now include four species, *Alfalfa leaf curl virus*, *Euphorbia caput-medusae latent virus*, *French bean severe leaf curl virus* and *Plantago lanceolata latent virus* [5, 6, 30, 36]
- 2) *Grablovirus*, which will for now include one species, *Grapevine red blotch virus* [18, 28, 35]

In addition, two other species of highly divergent viruses that are closely related to geminiviruses have been accepted by the Executive Committee of the ICTV for inclusion in the family *Geminiviridae* and will remain unassigned to a genus, awaiting identification of the vector species and confirmation of particle morphology. These new species are:

- 1) *Citrus chlorotic dwarf associated virus* [17, 21]
- 2) *Mulberry mosaic dwarf associated virus* [23, 24]

Furthermore, recently, a single virus isolated from an apple tree in China (tentatively referred to as apple geminivirus; AGmV) [20] and 13 virus isolates from grapevine

Fig. 1 Genome organisation of viruses in the genera of the family *Geminiviridae*. *Citrus chlorotic dwarf associated virus* and *Mulberry mosaic dwarf associated virus* have been accepted by the Executive Committee of the International Committee on Taxonomy of Viruses (ICTV) as a species within the family *Geminiviridae* but have not been assigned to a genus. AGmV and GGVA have not yet been considered as representatives of a species by ICTV

in Hungary, Israel, Japan and South Korea (tentatively referred to as grapevine geminivirus A; GGVA) sharing >97% pairwise identity have been discovered [2]. These have distinctly geminivirus-like genome organisations and are more closely related to members of some established geminivirus species than they are to citrus chlorotic dwarf associated virus (CCDaV) and mulberry mosaic dwarf associated virus (MMCaV). The taxonomic assignment of AGmV and GGVA still needs to be discussed by the ICTV.

Whereas the arrangements of genes and intergenic regions in the genomes of curtoviruses, topocuviruses, turncurtoviruses, AGmV, GGVA and genomes/DNA A components of begomoviruses are all very similar, the becurtoviruses, capulaviruses, eragroviruses, grabloviruses and mastreviruses all have distinctly unique genome organisations (Fig. 1). Furthermore, the genome organisation of CCDaV is also apparently unique [21]. Particularly noteworthy in the CCDaV genome (3640 nt) is the ~921-nt-long virion-sense ORF, V3, which results in the CCDaV genome being 600-800 nt larger than those of the other known monopartite geminiviruses (Fig. 1).

Although the genomes of the known geminiviruses can have up to eight genes, only two of them, the replication-associated protein gene (*rep*) and the coat protein gene (*cp*) are obviously conserved across viruses in all of the genera, whereas the replication enhancer gene (*ren*), symptom determinant gene (*sd*) and silencing suppressor gene (*ss*) homologues are found in begomoviruses, curtoviruses, topocuviruses and turncurtoviruses (Fig. 1), and transcription activator gene (*trap*) homologues are additionally found in eragroviruses (Fig. 1). The genomes of viruses in all of the established genera also have either proven or putative movement protein genes (*mp*), which, in the monopartite members of the family, are immediately upstream of *cp*. It is currently unknown, however, if these genes in viruses from different genera have a common evolutionary origin, because the amino acid sequences that they encode have such low degrees of similarity that it is not possible to conclude that they are homologous.

The genomes of becurtoviruses, capulaviruses, grabloviruses, eragroviruses, mastreviruses, and CCDaVs all have two intergenic regions (Fig. 1). In general, the longer of these is referred to as the long intergenic region (LIR) and contains a conserved nonanucleotide within a hairpin structure that forms the origin of virion-strand replication [3, 22]. Similarly, the shorter intergenic region is referred



to as the short intergenic region (SIR) and contains transcription termination signals and, for mastreviruses at least, the origin of complementary-strand replication.

Introns are present in several geminivirus genes. Rep and RepA proteins in becurtoviruses, capulaviruses, grabloviruses, mastreviruses, MMCAV and CCDaV are either proven to be, or inferred to be, expressed from alternatively spliced complementary-sense transcripts [8, 13, 40] (Fig. 1). Further, an intron occurs in the *mp* of mastreviruses [40], and an intron has been inferred to occur in the *mp* of the capulavirus EcmLV (Fig. 1) [5].

Recombination has played a significant role in the evolution of geminiviruses, and it can therefore be difficult to accurately infer the evolutionary relationships of viruses in different genera. Whereas non-homologous recombination has potentially contributed to the loss or gain of large genome regions in particular genera, homologous recombination has also likely contributed to the exchange of genome regions between viruses in different genera [6, 9, 31, 33].

Nonetheless, we inferred the approximate genome-wide evolutionary relationships between representative members of the various established geminivirus genera, CCDaV, MMCAV, AGmV and GGVA (Fig. 1). The full genome sequences of these representative viruses were aligned using MUSCLE [15] and used to infer a neighbour-joining phylogenetic tree with the Jukes-Cantor nucleotide substitution model and 1000 bootstrap replicates. Branches with <60% support were collapsed using TreeGraph2 [34], and the tree was rooted at the midpoint (Fig. 2). From this phylogenetic tree, it is apparent that the capulaviruses and grabloviruses are closely related to one another and that CCDaV and MMCAV are closely related to one another.

We also analysed the evolutionary relationships of the inferred CP and Rep amino acid sequences of the same set of representative geminiviruses. The CP and Rep sequence datasets were aligned using MUSCLE [15] and used to infer maximum-likelihood phylogenetic trees using PHYML3.0 [16], applying the WAG+I+G and LG+I+G substitution model, respectively (selected as the best-fit models by ProtTest [14]; Fig. 2). Branch support in these trees was determined using approximate likelihood ratio tests (aLRT). The Rep phylogenetic tree was rooted with sequences from the family *Genomoviridae* [19], whereas the CP tree was rooted at the midpoint. Branches with <80% aLRT support were collapsed using TreeGraph2 [34]. Pairwise CP and Rep amino acid sequence identities were determined using SDT v1.2 [27] (Fig. 3).

Discordance in the phylogenetic placement of curtoviruses, becurtoviruses, topocoviruses, eragroviruses, turncurtoviruses and CCDaV in the CP and Rep phylogenetic trees likely reflects the fact that the most recent common ancestors of the known viruses in these genera are

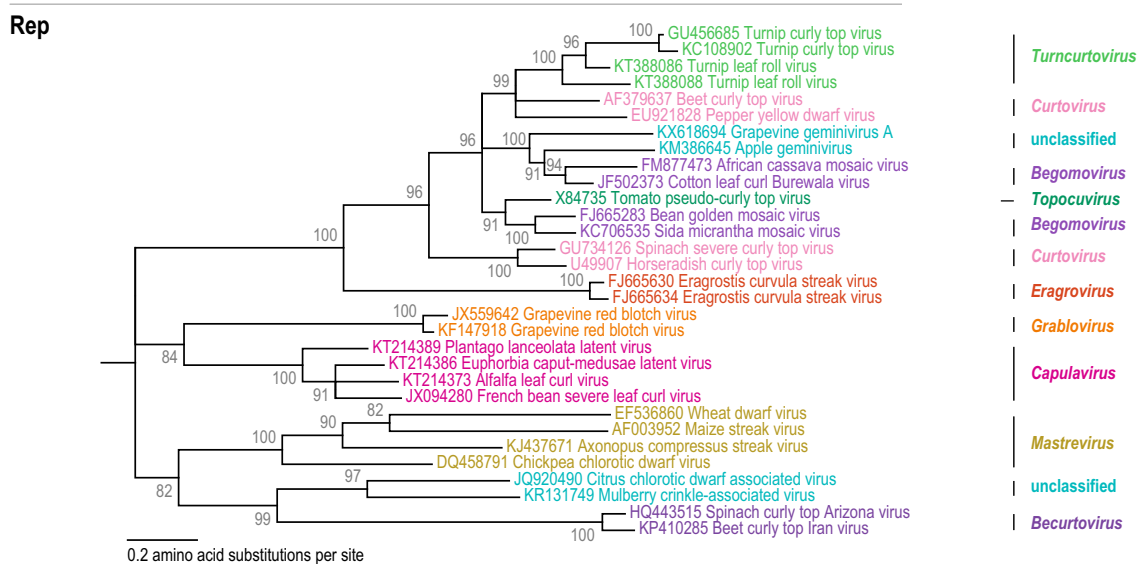
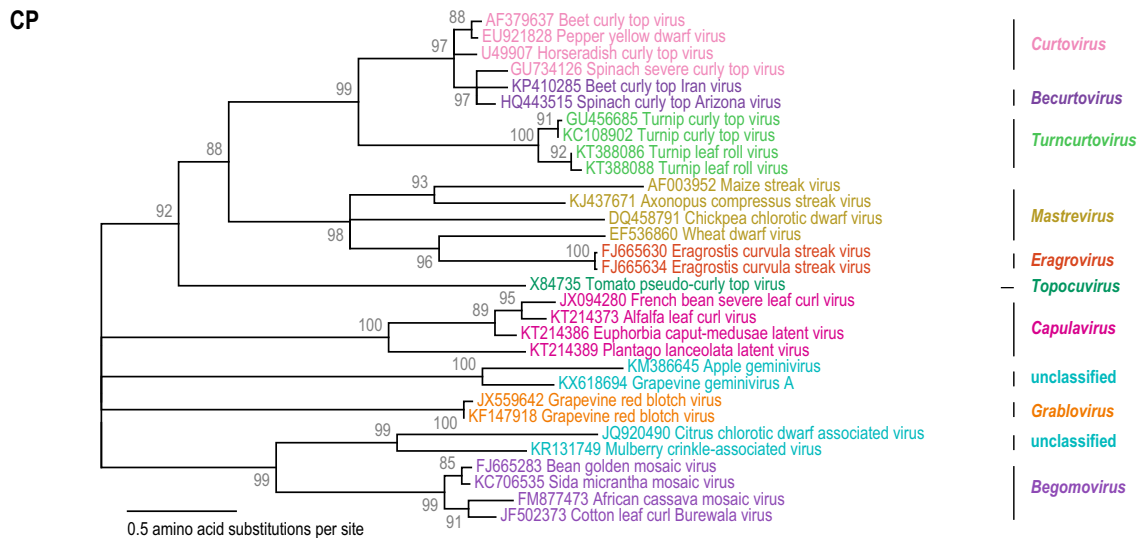
Fig. 2 **A.** Unrooted neighbour-joining tree inferred from aligned full-genome sequences of representative isolates from various genera in the family *Geminiviridae*. Numbers associated with branches indicate percentage bootstrap support for these branches (determined with 1000 bootstrap replicates). **B.** Maximum-likelihood phylogenetic tree (applying the WAG+I+G amino acid substitution model) inferred from aligned CP sequences of representative isolates from various genera in the family *Geminiviridae*. Numbers associated with branches indicate percentage aLRT support for these branches. The CP phylogenetic tree is rooted at the midpoint. Branches with less than 80% aLRT support have been collapsed. **C.** Maximum-likelihood phylogenetic tree (applying the LG+I+G amino acid substitution model) inferred from aligned Rep sequences of representative isolates from various genera in the family *Geminiviridae*. Numbers associated with branches indicate percentage aLRT support for these branches. The Rep phylogenetic tree is rooted with the Rep sequences of members of the family *Genomoviridae* (not included here). Branches with less than 80% aLRT support have been collapsed

possibly the products of inter-genus recombination events [10, 37]. For example, the becurtovirus CPs are most closely related to those of curtoviruses, whereas the becurtovirus Reps are most closely related to those of CCDaV and MMCAV (Fig. 2 and Fig. 3).

Capulavirus

A group of closely related viruses isolated from *Euphorbia caput-medusae* (South Africa), *Medicago sativa* (France and Spain), *Phaseolus vulgaris* (India) and *Plantago lanceolata* (Finland) have been assigned to the new genus *Capulavirus*. The genus name *Capulavirus* was derived from the type member of the genus: euphorbia *caput-medusae* *latent virus*. As with viruses in the genera *Mastrevirus* and *Becurtovirus*, capulaviruses have two intergenic regions and express Rep from a spliced complementary-strand transcript. In common with begomoviruses and curtoviruses, they have a large complementary-sense ORF (C3) that is completely embedded within *rep*. A unique feature of capulavirus genomes is a complex arrangement of possible MP-encoding ORFs located in the 5' direction from *cp* (Fig. 1). Two or more of these ORFs may constitute an intron-containing *mp* (Fig. 1). Geminate particles have been observed in purified preparations of EcmLV by transmission electron microscopy [30]. Aphids of the species *Aphis craccivora* have been shown to transmit ALCV [30]. All known capulaviruses have the nonanucleotide motif "TAATATTAC" at their presumed origins of virion-strand replication.

An analysis of the distribution of pairwise identities (one minus Hamming distances of pairwise aligned sequences with pairwise deletion of gaps) of known capulavirus genomes using SDT v1.2 [27] ($n = 47$; Table 1) indicates that a pairwise-identity-based species demarcation criterion that would minimize conflicts (i.e., possible assignments of



individual isolates to two or more species) could be placed either in the 72 to 84% identity interval, or in the 87–94% interval (Fig. 4). Whereas a species demarcation threshold within the 87–94% range would yield five species, placing it within the 72–84% range would yield four species. In the case of a threshold within the 87–94% range, the alfalfa leaf curl virus (ALCV) sequences, which have all been isolated from alfalfa plants, would be split into two species. Given that there seems to be no good biological reason to split the ALCVs into two species, we have opted to tentatively place the species demarcation threshold in the 72–84% range. To bring the capulavirus species demarcation threshold in line with that of the genus *Mastrevirus* (the recognized genus of geminiviruses that the capulaviruses are most closely related to; Fig. 4) we have opted to tentatively select 78% as the pairwise identity value above which two sequences should be considered isolates of the same species. The four proposed capulavirus species to which the 47 known capulavirus full genome sequences (Table 1) have been assigned are *Alfalfa leaf curl virus* (n = 26), *Euphorbia caput-medusae latent virus* (n = 17), *French bean severe leaf curl virus* (n = 2) and *Plantago lanceolata latent virus* (n = 2). Based on pairwise identity comparisons, all isolates (ALCV; EcmLV; french bean severe leaf curl virus, FbSLCV; plantago lanceolata latent virus, PILV) within each of these species share between 63% and 73% genome-wide sequence identity with all isolates that have been assigned to other proposed capulavirus species (Fig. 2).

Regardless of whether the full-genome nucleotide sequence, the inferred Rep amino acid sequence, or the inferred CP amino acid sequence is considered, all of the viruses belonging to the proposed capulavirus species group with 100% aLRT support with other proposed capulaviruses (Fig. 2). The viruses assigned to the genus *Capulavirus* share <22% CP amino acid sequence identity and <45% Rep amino acid sequence identity with viruses in other established genera of the family *Geminiviridae* (Fig. 3).

Grablovirus

A group of closely related viruses discovered infecting grapevines in Canada, South Korea and the USA have been assigned to the genus *Grablovirus*. The genus name is based on the name of the type member: grapevine red blotch virus (GRBV), the first and currently the only member of this genus [18]. These viruses have the virion-strand origin of replication nonanucleotide motif TAA-TATTAC and a unique genome arrangement (Fig. 1; Table 2). The natural vector is likely the three-cornered alfalfa treehopper (*Spissistilus festinus* Say) [4].

Analysis of the 27 known grablovirus full-genome sequences indicate that they all share >91% genome-wide nucleotide sequence identity (Fig. 5) and are therefore all assigned to the species *Grapevine red blotch virus*. All of the GRBV isolates share <45% Rep amino acid sequence identity with all viruses in other geminivirus genera (Fig. 3). Regardless of whether the full-genome nucleotide sequence, Rep amino acid sequence or CP amino acid is considered, the GRBV isolates cluster together with 100% aLRT support (Fig. 2).

In the absence of additional related sequences in the genus, we suggest that, based on the species demarcation criteria used for most other geminivirus genera, grablovirus isolates sharing <80% genome-wide pairwise identity with members of accepted grablovirus species should be classified as members of distinct grablovirus species. This is a tentative species demarcation criterion that will need to be refined as more full-genome sequences of grabloviruses are determined.

New geminivirus species that remain unassigned to a genus

Citrus chlorotic dwarf associated virus and Mulberry mosaic dwarf associated virus

Two complete CCDaV genomes, isolated from citrus trees in Turkey (JQ920490) [21] and China (KF561253), and eight complete MMCaV genomes, isolated from mulberry bushes in China (KP303687, KP699128–KP699132, KP728254, KR131749) [23, 24], are clearly closely related to those of geminiviruses, based on both their geminivirus-like genome organization and the fact that their inferred Rep and CP amino acid sequences cluster phylogenetically with those of the known geminiviruses (Fig. 1 and Fig. 2). All known CCDaV and MMDaV isolates also have the nonanucleotide TAATATTAC, which is characteristic of the virion-strand origin of replication of the majority of geminiviruses.

The two CCDaV sequences share 99.5% genome-wide pairwise identity, whereas the eight MMDaV sequences share >97% identity. The predicted Rep proteins of these CCDaV and MMDaV isolates share ~58% amino acid sequence similarity with one another, but only 33–42% similarity with those of other geminiviruses (Fig. 3). Regardless of whether full genome nucleotide sequences, inferred Rep amino acid sequences, or inferred CP amino acid sequences are considered, the two proposed species group separately from all other geminiviruses, with 100% aLRT branch support (Fig. 2).

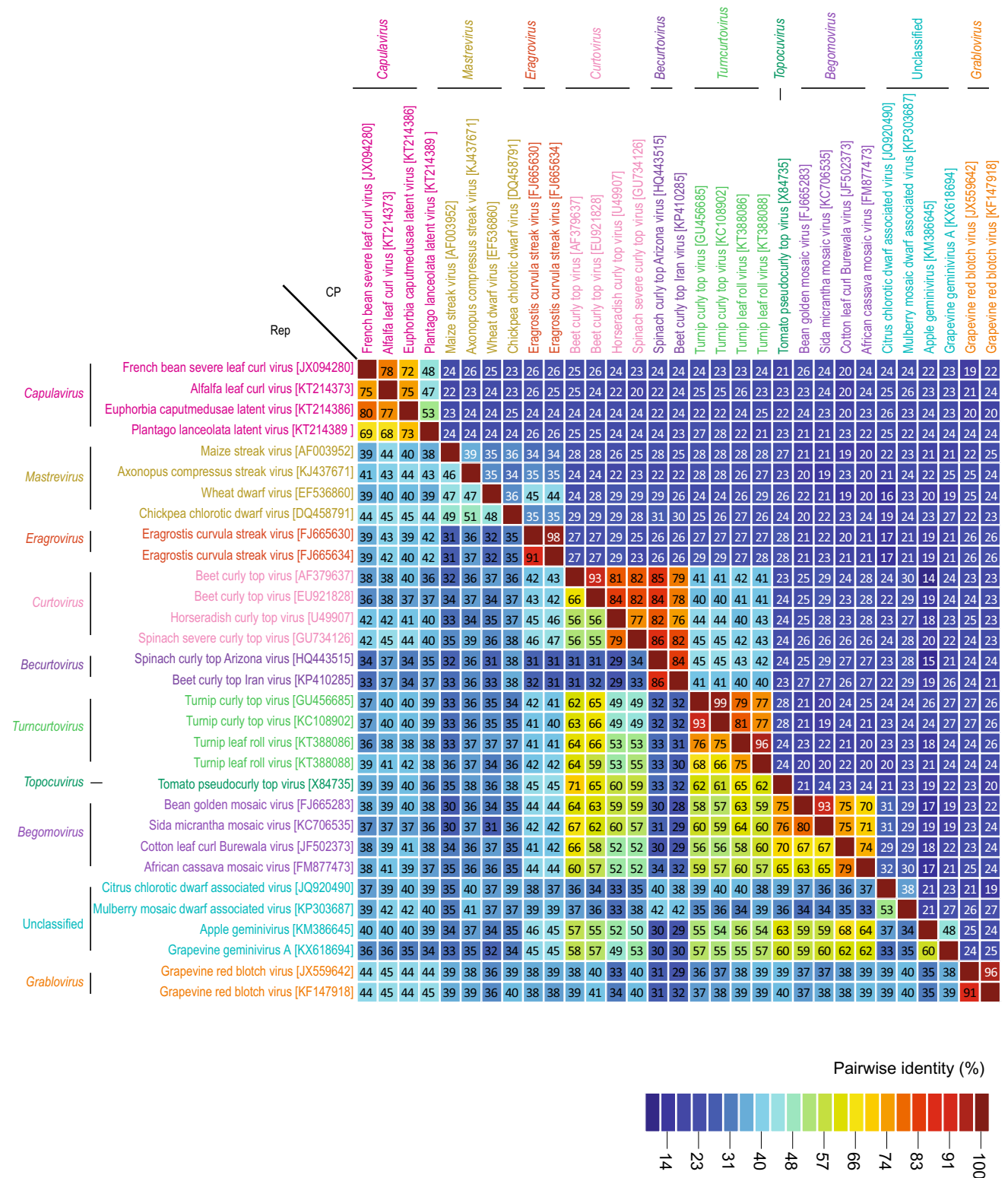


Fig. 3 Pairwise identities of the Rep and CP sequences of representative isolates from various genera in the family *Geminiviridae* as determined using SDT v1.2 (Muhire et al., 2014)

Table 1 Details of virus isolates that have been assigned to the four proposed species within the genus *Capulavirus*

<i>Capulavirus</i>	Accession no.	Virus acronym	Isolate	Country	Host	
<i>Alfalfa leaf curl virus</i>	KP732474	ALCV	44-1E	France	<i>Medicago sativa</i>	
	KT214350	ALCV	TDV14_44-16	France	<i>Medicago sativa</i>	
	KT214351	ALCV	PB14_LUZ166	France	<i>Medicago sativa</i>	
	KT214352	ALCV	ASS14_Assas2	France	<i>Medicago sativa</i>	
	KT214353	ALCV	BON14_LUZ075	France	<i>Medicago sativa</i>	
	KT214354	ALCV	SSL14_Toul1	France	<i>Medicago sativa</i>	
	KT214355	ALCV	ALB14_LUZ147	France	<i>Medicago sativa</i>	
	KT214356	ALCV	ALB14_LUZ163	France	<i>Medicago sativa</i>	
	KT214357	ALCV	SSL14_Toul7	France	<i>Medicago sativa</i>	
	KT214358	ALCV	TDV12_48-2A	France	<i>Medicago sativa</i>	
	KT214359	ALCV	PB14_LUZ184	France	<i>Medicago sativa</i>	
	KT214360	ALCV	ASS14_Assas1	France	<i>Medicago sativa</i>	
	KT214361	ALCV	VAU14_LUZ136	France	<i>Medicago sativa</i>	
	KT214362	ALCV	GAG14_LUZ193	France	<i>Medicago sativa</i>	
	KT214363	ALCV	PB14_LUZ182	France	<i>Medicago sativa</i>	
	KT214364	ALCV	ALB14_LUZ148	France	<i>Medicago sativa</i>	
	KT214365	ALCV	PB14_LUZ-GS4	France	<i>Medicago sativa</i>	
	KT214366	ALCV	ALB14_LUZ164	France	<i>Medicago sativa</i>	
	KT214367	ALCV	PB14_LUZ171	France	<i>Medicago sativa</i>	
	KT214368	ALCV	PB14_LUZ188	France	<i>Medicago sativa</i>	
	KT214369	ALCV	PB14_LUZ-GS6	France	<i>Medicago sativa</i>	
	KT214370	ALCV	BON14_LUZ076	France	<i>Medicago sativa</i>	
	KT214371	ALCV	PB14_LUZ165	France	<i>Medicago sativa</i>	
	KT214372	ALCV	PB14_LUZ178	France	<i>Medicago sativa</i>	
	KT214373	ALCV	VAU14_LUZ142	France	<i>Medicago sativa</i>	
	KT214374	ALCV	PB14_GS8	France	<i>Medicago sativa</i>	
	KT214375	ALCV	PB14_LUZ179	France	<i>Medicago sativa</i>	
	<i>Euphorbia caput-medusae latent virus</i>	HF921459	EcmLV	Dar10	South Africa	<i>Euphorbia caput-medusae</i>
		HF921460	EcmLV	Dar11	South Africa	<i>Euphorbia caput-medusae</i>
		HF921477	EcmLV	Lap11	South Africa	<i>Euphorbia caput-medusae</i>
		KT214376	EcmLV	DAR12_CM243	South Africa	<i>Euphorbia caput-medusae</i>
		KT214377	EcmLV	DAR12_CM26	South Africa	<i>Euphorbia caput-medusae</i>
KT214378		EcmLV	DAR12_CM176	South Africa	<i>Euphorbia caput-medusae</i>	
KT214379		EcmLV	DAR12_CM27	South Africa	<i>Euphorbia caput-medusae</i>	
KT214380		EcmLV	DAR12_CM192	South Africa	<i>Euphorbia caput-medusae</i>	
KT214381		EcmLV	DAR12_CM186	South Africa	<i>Euphorbia caput-medusae</i>	
KT214382		EcmLV	DAR12_CM240	South Africa	<i>Euphorbia caput-medusae</i>	
KT214383		EcmLV	DAR12_CM162	South Africa	<i>Euphorbia caput-medusae</i>	
KT214384		EcmLV	PAT12_CM96	South Africa	<i>Euphorbia caput-medusae</i>	
KT214385		EcmLV	PAT12_CM99	South Africa	<i>Euphorbia caput-medusae</i>	
KT214386		EcmLV	PAT12_CM101	South Africa	<i>Euphorbia caput-medusae</i>	
KT214387		EcmLV	DAR12_CM217	South Africa	<i>Euphorbia caput-medusae</i>	
KT214388	EcmLV	DAR12_CM251	South Africa	<i>Euphorbia caput-medusae</i>		
<i>French bean severe leaf curl virus</i>	JX094280	FbSLCV	FbSLCV-1	India	<i>Phaseolus vulgaris</i>	
	JX094281	FbSLCV	FbSLCV-2	India	<i>Phaseolus vulgaris</i>	
<i>Plantago lanceolata latent virus</i>	KT214389	PILV	ALA13_PI11	Finland	<i>Plantago lanceolata</i>	
	KT214390	PILV	ALA13_PI5	Finland	<i>Plantago lanceolata</i>	

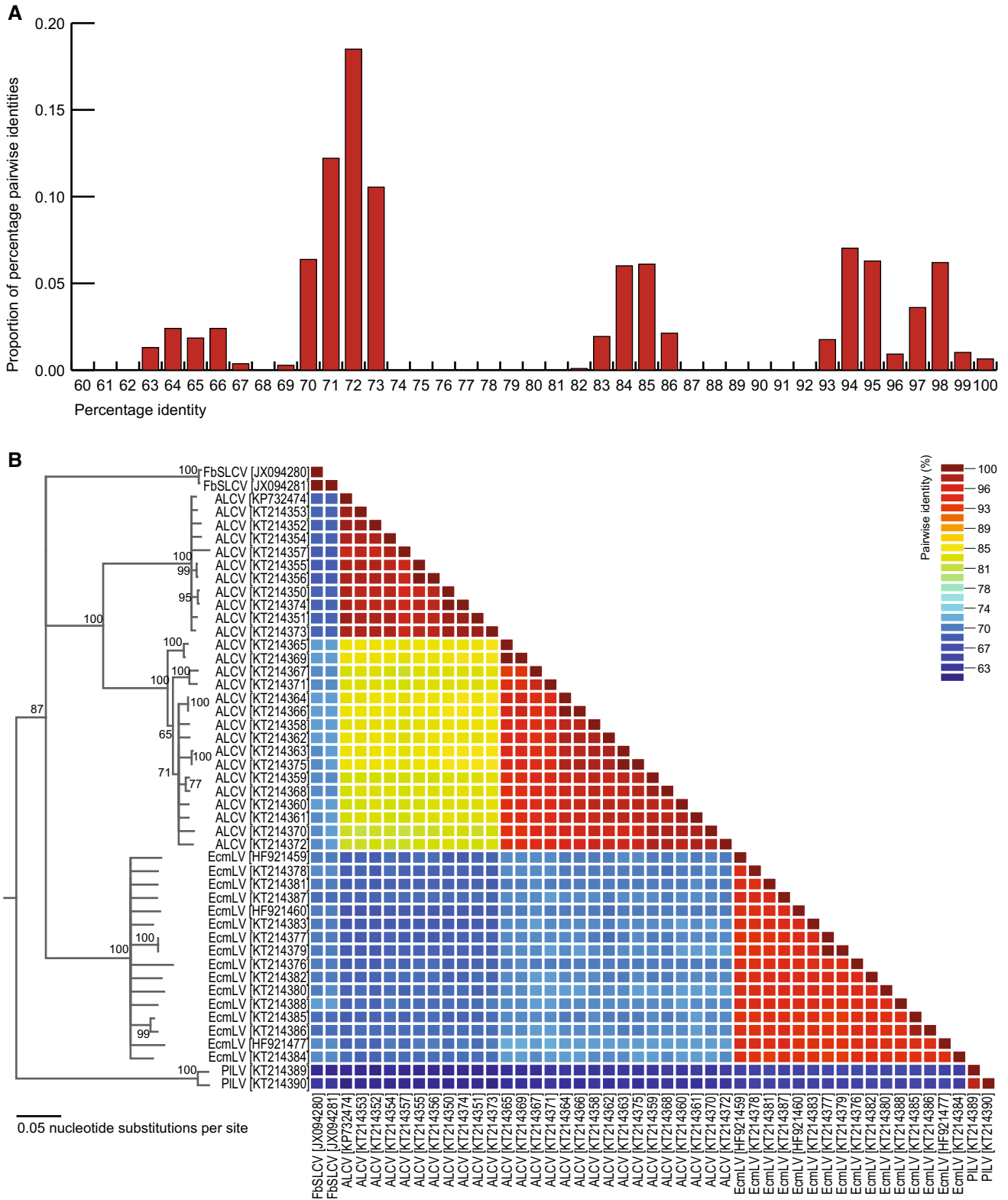


Fig. 4 **A.** Distribution of pairwise nucleotide sequence identities of 47 known capulavirus genomes. **B.** Genome-wide pairwise identities determined using SDT v1.2 [27] and a neighbor-joining phylogenetic

tree of capulavirus isolates. The tree is rooted with mastrevirus genome sequences (not shown)

Table 2 Details of virus isolates that have been assigned to the proposed species *Grapevine red blotch virus* within the genus *Grablovirus*

<i>Grablovirus</i>	Accession no.	Virus acronym	Isolate	Country	Host
<i>Grapevine red blotch virus</i>	KF147916	GRBV	NY175	USA ¹	<i>Vitis vinifera</i>
	KU564249	GRBV	NY701	USA ²	<i>Vitis vinifera</i> cv. Cabernet Sauvignon
	JQ901105	GRBV	JRT[456]17NOV10	USA ³	<i>Vitis vinifera</i>
	KC896623	GRBV	CF214-1	USA	<i>Vitis vinifera</i>
	KC896625	GRBV	Z1A-1	USA	<i>Vitis vinifera</i>
	KF137562	GRBV	OR1a	USA ⁴	<i>Vitis vinifera</i> cv. Pinot noir
	KF147917	GRBV	NY135	USA ⁵	<i>Vitis vinifera</i>
	KF147918	GRBV	NY137	USA ⁵	<i>Vitis vinifera</i>
	KF751708	GRBV	NY147	USA ³	<i>Vitis vinifera</i> cv. Pinot Noir
	KP221559	GRBV	1	USA	<i>Vitis vinifera</i> cv. Early Burgundy
	KU564247	GRBV	NY699	USA ²	<i>Vitis californica</i> x <i>Vitis vinifera</i> hybrid
	KU564248	GRBV	NY700	USA ²	<i>Vitis californica</i> x <i>Vitis vinifera</i> hybrid
	KU564250	GRBV	NY702	USA ²	<i>Vitis vinifera</i> cv. Merlot
	KU564251	GRBV	NY703	USA ²	<i>Vitis vinifera</i> cv. Cabernet franc
	KU564252	GRBV	NY704	USA ²	<i>Vitis vinifera</i> cv. Cabernet franc
	KU564253	GRBV	NY705	USA ²	<i>Vitis vinifera</i> cv. Cabernet franc
	KU564254	GRBV	NY913	USA ²	<i>Vitis californica</i> x <i>Vitis vinifera</i> hybrid
	KU564255	GRBV	NY921	USA ²	<i>Vitis vinifera</i> cv. Cabernet franc
	KU564256	GRBV	NY926	USA ²	<i>Vitis vinifera</i> cv. Cabernet franc
	JX559642	GRBV	3138-03	Canada	<i>Vitis vinifera</i>
	KC896624	GRBV	CS337-1	USA	<i>Vitis vinifera</i>
	KF147915	GRBV	NY271	USA	<i>Vitis vinifera</i>
	KU821056	GRBV	SW6	South Korea	<i>Vitis vinifera</i>
	KC427993	GRBV	GiGV-WA-RS	USA	<i>Vitis vinifera</i>
	KC427994	GRBV	GiGV-WA-DS	USA	<i>Vitis vinifera</i>
	KC427995	GRBV	GiGV-WA-MR	USA	<i>Vitis vinifera</i> cv. Merlot
	KC427996	GRBV	GiGV-WA-CF	USA	<i>Vitis vinifera</i> cv. Cabernet franc

US state: ¹MD, ²CA, ³NY, ⁴OR, ⁵PA

CCDaV and MMDaV will need to remain unassigned to a genus until their respective vector species are identified and/or it is confirmed that they form geminate particles.

New geminiviruses identified recently that need to be reviewed by ICTV

Apple geminivirus and grapevine geminivirus A

One AGmV complete genome isolated from an apple tree in China (accession no. KM386645) [20] and thirteen GGVA complete genomes isolated from grapevine in China (KX570611), Hungary (KX570618), Israel (KX618694), Japan (KX570610, KX570612, KX570613, KX570615, KX570616, KX570617) and South Korea

(KX570607, KX570609, KX570614) group together, with 94% bootstrap support (Fig. 2). These viruses isolated from woody plants contain the typical geminivirus-like nanonucleotide motif TAATATTAC and display a distinctly geminivirus-like genome organisation. In addition, these viruses are clearly recombinant, with a begomovirus-like Rep (sharing 59–68% amino acid sequence identity) and a divergent CP that is unlike that of any known geminivirus.

The thirteen GGVA sequences share >97% identity. The unique AGmV sequence and the thirteen GGVA sequences share ~63% genome-wide pairwise identity. The Rep and CP sequences of the AGmV and GGVA isolates share ~60% and 48% amino acid sequence identity, respectively, with one another. The Rep proteins are most closely related to those of begomoviruses, curtoviruses, topocovirus and turncurtoviruses, sharing 50–68% identity (Fig. 3). Although, based

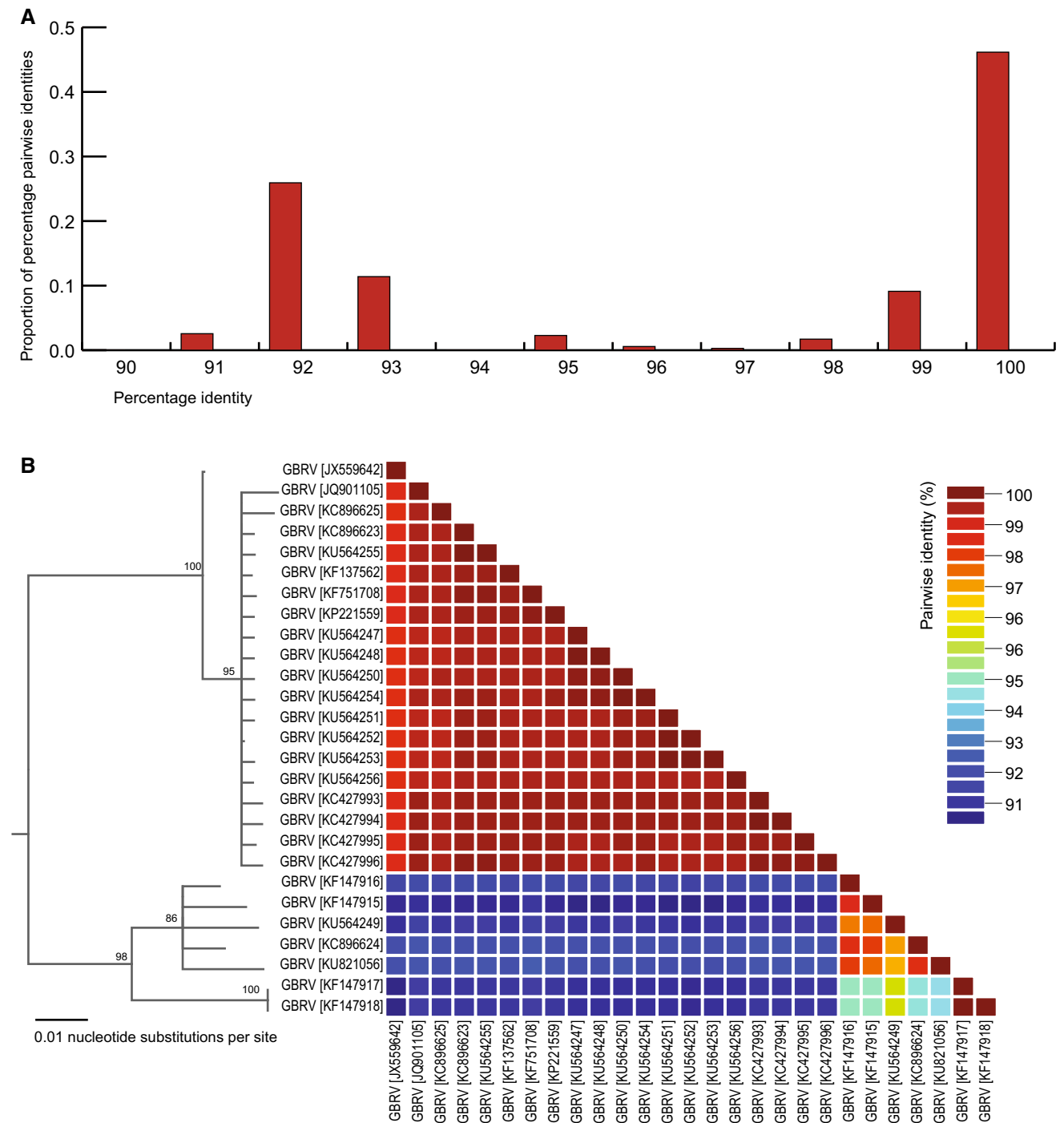


Fig. 5 **A.** Distribution of pairwise nucleotide sequence identities of 27 grablovirus genomes. **B.** Genome-wide pairwise identities determined using SDT v1.2 [27] and a neighbor-joining phylogenetic tree

of grablovirus isolates. The tree is rooted with mastrevirus genome sequences (not shown)

on inferred Rep, AGmV and GGVA branch with begomoviruses, both viruses group separately from all other geminiviruses, with 100% aLRT branch support (Fig. 2), based on the inferred CP amino acid sequences.

Concluding remarks

AGmV and GGVA must also remain unassigned to a genus until the vectors are identified and/or geminate particles are observed.

The past ten years have seen the establishment of five new geminivirus genera, and the next ten will likely see many more being established as metagenomics-based approaches begin enabling the convenient and cost-effective discovery of viruses in greater numbers from cultivated and non-

cultivated plant host species [29, 32]. As well as illuminating the actual diversity of geminiviruses and the role played by recombination in their evolutionary history, such approaches also have the potential to elucidate the full range of geminivirus host and vector species. Knowing where geminiviruses are found (especially in non-cultivated hosts where they may cause no obvious disease symptoms) and how they are transmitted will not only enable prediction of which geminiviruses could pose serious threats to agriculture but also provide a detailed view of the actual role of geminiviruses in natural and disturbed terrestrial ecosystems.

Disclaimer This article is based on the taxonomic proposal 2016.014a-iP.A.v2.Geminiviridae_2gen.pdf, which has been considered and approved by the Executive Committee (EC) of the ICTV. A.V. and F.M.Z. are elected members of the ICTV EC. A.V., D.P.M., P.R., J.N.-C., E.M., A.I., R.W.B., R.R.-B. and F.M.Z. are current members of the *Geminiviridae* study group of the ICTV.

Compliance with ethical standards

There are no conflicts of interest; the research did not involve human participants or animals. The data used for the analyses in this manuscript are publicly available in GenBank.

References

- Adams MJ, Lefkowitz EJ, King AM, Harrach B, Harrison RL, Knowles NJ, Kropinski AM, Krupovic M, Kuhn JH, Mushegian AR, Nibert M, Sabanadzovic S, Sanfacon H, Siddell SG, Simmonds P, Varsani A, Zerbini FM, Gorbalenya AE, Davison AJ (2016) Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2016). *Arch Virol* 161:2921–2949
- Al Rwahnih M, Alabi OJ, Westrick NM, Golino D, Rowhani A (2017) Description of a novel monopartite geminivirus and its defective subviral genome in grapevine. *Phytopathology* 107:240–251
- Arguello-Astorga GR, Ruiz-Medrano R (2001) An iteron-related domain is associated to Motif 1 in the replication proteins of geminiviruses: identification of potential interacting amino acid-base pairs by a comparative approach. *Arch Virol* 146:1465–1485
- Bahder BW, Zalom FG, Jayanth M, Sudarshana MR (2016) Phylogeny of geminivirus coat protein sequences and digital PCR aid in identifying *Spissistilus festinus* as a vector of grapevine red blotch-associated virus. *Phytopathology* 106:1223–1230
- Bernardo P, Golden M, Akram M, Naimuddin Nadarajan N, Fernandez E, Granier M, Marais A, Monge Blasco P, Martin DP, Roumagnac P (2013) Identification and characterisation of a highly divergent geminivirus: evolutionary and taxonomic implications. *Virus Res* 177:35–45
- Bernardo P, Muhire B, Francois S, Deshoux M, Hartnady P, Farkas K, Kraberger S, Filloux D, Fernandez E, Galzi S, Ferdinand R, Granier M, Marais A, Monge Blasco P, Candresse T, Escriu F, Varsani A, Harkins GW, Martin DP, Roumagnac P (2016) Molecular characterization and prevalence of two capulaviruses: alfalfa leaf curl virus from France and euphorbia caput-medusae latent virus from South Africa. *Virology* 493:142–153
- Bottcher B, Unsel S, Ceulemans H, Russell RB, Jeske H (2004) Geminate structures of African cassava mosaic virus. *J Virol* 78:6758–6765
- Bozorgi N, Heydarnejad J, Kamali M, Massumi H (2017) Splicing features in the expression of the complementary-sense genes of Beet curly top Iran virus. *Virus Genes*. doi:10.1007/s11262-016-1422-y
- Briddon RW, Bedford ID, Tsai JH, Markham PG (1996) Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, tomato pseudo-curly top virus, suggests a recombinant origin. *Virology* 219:387–394
- Briddon RW, Heydarnejad J, Khosrowfar F, Massumi H, Martin DP, Varsani A (2010) Turnip curly top virus, a highly divergent geminivirus infecting turnip in Iran. *Virus Res* 152:169–175
- Brown JK, Fauquet CM, Briddon RW, Zerbini FM, Moriones E, Navas-Castillo J (2012) Family Geminiviridae. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) *Virus taxonomy*. Ninth report of the international committee on taxonomy of viruses. Elsevier Academic Press, London, pp 351–373
- Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrino R, Silva JC, Fiallo-Olive E, Briddon RW, Hernandez-Zepeda C, Idris A, 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
- Candresse T, Filloux D, Muhire B, Julian C, Galzi S, Fort G, Bernardo P, Daugrois JH, Fernandez E, Martin DP, 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:e102945
- Darriba D, Taboada GL, Doallo R, Posada D (2011) ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 27:1164–1165
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321
- Guo J, Lai XP, Li JX, Yue JQ, Zhang SY, Li YY, Gao JY, Wang ZR, Duan HF, Yang JD (2015) First report on citrus chlorotic dwarf associated virus on lemon in Dehong Prefecture, Yunnan, China. *Plant Dis* 99:1287
- Krenz B, Thompson JR, Fuchs M, Perry KL (2012) Complete genome sequence of a new circular DNA virus from grapevine. *J Virol* 86:7715
- Krupovic M, Ghabrial SA, Jiang D, Varsani A (2016) *Genomoviridae*: a new family of widespread single-stranded DNA viruses. *Arch Virol* 161:2633–2643
- Liang P, Navarro B, Zhang Z, Wang H, Lu M, Xiao H, Wu Q, Zhou X, Di Serio F, Li S (2015) Identification and characterization of a novel geminivirus with a monopartite genome infecting apple trees. *J Gen Virol* 96:2411–2420
- Loconsole G, Saldarelli P, Doddapaneni H, Savino V, Martelli GP, Saponari M (2012) Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family *Geminiviridae*. *Virology* 432:162–172
- Londono A, Riego-Ruiz L, Arguello-Astorga GR (2010) DNA-binding specificity determinants of replication proteins encoded by eukaryotic ssDNA viruses are adjacent to widely separated RCR conserved motifs. *Arch Virol* 155:1033–1046
- Lu QY, Wu ZJ, Xia ZS, Xie LH (2015) Complete genome sequence of a novel monopartite geminivirus identified in mulberry (*Morus alba* L.). *Arch Virol* 160:2135–2138
- Ma Y, Navarro B, Zhang Z, Lu M, Zhou X, Chi S, Di Serio F, Li S (2015) Identification and molecular characterization of a novel

- monopartite geminivirus associated with mulberry mosaic dwarf disease. *J Gen Virol* 96:2421–2434
25. Moffat AS (1999) Plant pathology—geminiviruses emerge as serious crop threat. *Science* 286:1835
 26. 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
 27. Muhire BM, Varsani A, Martin DP (2014) SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* 9:e108277
 28. Poojari S, Alabi OJ, Fofanov VY, Naidu RA (2013) A leafhopper-transmissible DNA virus with novel evolutionary lineage in the family geminiviridae implicated in grapevine redleaf disease by next-generation sequencing. *PLoS One* 8:e64194
 29. Roossinck MJ, Martin DP, Roumagnac P (2015) Plant virus metagenomics: advances in virus discovery. *Phytopathology* 105:716–727
 30. Roumagnac P, Granier M, Bernardo P, Deshoux M, Ferdinand R, Galzi S, Fernandez E, Julian C, Abt I, Filloux D, Mesleard F, Varsani A, Blanc S, Martin DP, Peterschmitt M (2015) Alfalfa leaf curl virus: an aphid-transmitted geminivirus. *J Virol* 89:9683–9688
 31. Rybicki EP (1994) A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. *Arch Virol* 139:49–77
 32. Simmonds P, Adams MJ, Benko M, Breitbart M, Brister JR, Carstens EB, Davison AJ, Delwart E, Gorbalenya AE, Harrach B, Hull R, King AMQ, Koonin EV, Krupovic M, Kuhn JH, Lefkowitz EJ, Nibert ML, Orton R, Roossinck MJ, Sabanadzovic S, Sullivan MB, Suttle CA, Tesh RB, van der Vlugt RA, Varsani A, Zerbini FM (2017) Consensus statement: virus taxonomy in the age of metagenomics. *Nat Rev Microbiol*. doi:10.1038/nrmi.cro.2016.1177
 33. Stanley J, Markham PG, Callis RJ, Pinner MS (1986) The nucleotide sequence of an infectious clone of the geminivirus beet curly top virus. *EMBO J* 5:1761–1767
 34. Stover BC, Muller KF (2010) TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinform* 11:7
 35. Sudarshana MR, Perry KL, Fuchs MF (2015) Grapevine red blotch-associated virus, an emerging threat to the grapevine industry. *Phytopathology* 105:1026–1032
 36. Susi H, Laine A-L, Filloux D, Kraberger S, Farkas K, Bernardo P, Frilander MJ, Martin DP, Varsani A, Roumagnac P (2017) Genome sequences of a capulavirus infecting *Plantago lanceolata* in the Åland archipelago of Finland. *Arch Virol* (in press)
 37. Varsani A, Shepherd DN, Dent K, Monjane AL, Rybicki EP, Martin DP (2009) A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. *Virol J* 6:36
 38. Varsani A, Martin DP, Navas-Castillo J, Moriones E, Hernandez-Zepeda C, Idris A, Murilo Zerbini F, Brown JK (2014) Revisiting the classification of curtoviruses based on genome-wide pairwise identity. *Arch Virol* 159:1873–1882
 39. Varsani A, Navas-Castillo J, Moriones E, Hernandez-Zepeda C, Idris A, Brown JK, Murilo Zerbini F, Martin DP (2014) Establishment of three new genera in the family *Geminiviridae*: *Becurtovirus*, *Eragrovirus* and *Turncurtovirus*. *Arch Virol* 159:2193–2203
 40. Wright EA, Heckel T, Groenendijk J, Davies JW, Boulton MI (1997) Splicing features in maize streak virus virion- and complementary-sense gene expression. *Plant J* 12:1285–1297
 41. Zhang W, Olson NH, Baker TS, Faulkner L, Agbandje-McKenna M, Boulton MI, Davies JW, McKenna R (2001) Structure of the maize streak virus geminate particle. *Virology* 279:471–477