

**Part 1:** **TITLE, AUTHORS, APPROVALS, etc**

|  |  |  |
| --- | --- | --- |
| **Code assigned:** | **2022.009P** |  |
| **Short title:** Create a new family, *Amesuviridae*, in the order *Mulpavirales*, including two genera, *Temfrudevirus* and *Yermavirus*, each with one species | | |
|  | | |

**Author(s) and email address(es)**

|  |  |
| --- | --- |
| Silva JPH, Silva JCF, Bejerman N, Zerbini, FM | joao.silva2@ufv.br, cleydson@ufv.br, bejerman.nicolas@inta.gob.ar, zerbini@ufv.br |

**Author(s) institutional address(es) (optional)**

|  |
| --- |
| Universidade Federal de Viçosa, Viçosa, Brazil [JPHS, JCFS, FMZ]; Instituto Nacional de Tecnología Agropecuaria, Córdoba, Argentina [NB] |

**Corresponding author**

|  |
| --- |
| FM Zerbini (zerbini@ufv.br) |

**List the ICTV Study Group(s) that have seen this proposal**

|  |
| --- |
| *Geminiviridae*  *Nanoviridae* |

**ICTV Study Group comments and response of proposer**

|  |
| --- |
|  |

**ICTV Study Group votes on proposal**

|  |  |  |  |
| --- | --- | --- | --- |
| **Study Group** | **Number of members** | | |
| **Votes support** | **Votes against** | **No vote** |
| *Geminiviridae* | 6 | 0 | 0 |

**Authority to use the name of a living person**

|  |  |
| --- | --- |
| **Is any taxon name used here derived from that of a living person (Y/N)** | N |

**Submission dates**

|  |  |
| --- | --- |
| Date first submitted to SC Chairs | May 27, 2022 |
| Date of this revision (if different to above) |  |

**ICTV-EC comments and response of the proposer**

|  |
| --- |
|  |

**Part 3:** **TAXONOMIC PROPOSAL**

**Name of accompanying Excel module**

|  |
| --- |
| 2022.009P.N.v1.Amesuviridae\_nf.xlsx |

**Abstract**

|  |
| --- |
| The creation of a new family of plant-infecting viruses is proposed. The new family is named *Amesuviridae*, and includes two new genera: *Temfrudevirus*, including the species *Temfrudevirus temperatum* (with temperate fruit decay-associated virus as a member),and *Yermavirus*, including the species *Yermavirus ilicis* (withyerba mate-associated circular DNA virus as a member)*.* Both viruses have circular, single-stranded DNA genomes and encode replication-associated proteins with an HUH endonuclease domain and a SH3 superfamily helicase domain. The new family *Amesuviridae* should be classified in the order *Mulpavirales* (which currently includes the families *Metaxyviridae* and *Nanoviridae*), class *Arfiviricetes*, phylum *Cressdnaviricota*, kingdom *Shotokuvirae*, realm *Monodnaviria*. |

**Text of proposal**

|  |  |
| --- | --- |
| |  | | --- | | *Cressdnaviricota* is a phylum including viruses that have small, circular, single-stranded DNA (ssDNA) genomes encoding replication-associated proteins with an HUH endonuclease domain and a SH3 superfamily helicase domain [1, 2]. Currently, *Cressdnaviricota* includes 11 families, three of which, *Geminiviridae*, *Metaxyviridae* and *Nanoviridae*, contain plant-infecting viruses [2]. A large number of viruses that harbor typical features of cressdnaviricots remain unclassified.  Here we propose the creation of a new family of plant-infecting cressdnaviricots named *Amesuviridae*, including two new genera: *Temfrudevirus*, including the species *Temfrudevirus temperatum* (with temperate fruit decay-associated virus as a member),and *Yermavirus*, including the species *Yermavirus ilicis* (withyerba mate-associated circular DNA virus as a member)*.*  Temperate fruit decay-associated virus(TFDaV) was described as infecting apple, pear and grapevine plants in Brazil [3]. The viral genome (Figure 1A) is approximately 3.4 kb in size and contains five ORFs, three in the viral strand (V1, V2 and V3) and two in the complementary strand (C1 and C2). *In silico* analysis indicated that V1 is a putative movement protein (MP), V2 encodes a viral coat protein (CP), and C1 is a replication-associated protein (Rep). The two remaining ORFs (V3 and C2) have no similarity with any other proteins in the databases [3]. The viral origin of replication is located in the minor intergenic region, and is comprised of a hairpin structure with a nonanucleotide sequence (5'-TAGTATTAC-3') identical to that of nanovirids, circovirids and alphasatellitids.  Yerba mate-associated circular DNA virus(YMaCV) was reported to infect yerba mate (*Ilex paraguariensis*) in Argentina [4]. Its genome has a length of 2.7 kb, with two intergenic regions, one of which harbors the putative origin of replication with the nonanucleotide 5'-CATTATTAC-3' (Figure 1B). The genome contains five ORFs, three in the viral strand (V1, V2 and V3) and two in the complementary strand (C1 and C2). The V2 ORF encodes a putative protein with a transmembrane domain similar to that of TFDaV, and is supposed to be the viral movement protein (MP). The V1 ORF encodes the putative coat protein (CP). The product of the C1 ORF is the replication-associated protein (Rep). The two remaining ORFs (V3 and C2) have no similarity with any other proteins in the databases [4].  To determine the taxonomical position and evolutionary relationship between TFDaV, YMaCV and the other cressdnaviricots, a data set of Rep amino acid sequences of representatives of all CRESS-DNA virus groups (including those not yet classified in the phylum *Cressdnaviricota*) [2] was constructed. A multiple sequence alignment was obtained using MAFFT v. 7 [5], with an open gap penalty of 1.53 and a gap extension penalty of 0.1. The alignment was inspected and regions with more than 50% gaps were removed using trimAL v. 1.2 [6], preserving at least 60% of the total length of the alignment. The best fitting substitution model (RTREV+G4+F) was determined using ModelTest-NG [7]. Maximum likelihood phylogenetic analysis was performed using RaxML-NG [8], with 1,000 transfer bootstrap expectation (TBE). Phylogenetic trees were viewed in FigTree v. 1.4.4. TFDaV and YMaCV formed a well-supported monophyletic clade close to nanovirids and alphasatellitids (Figure 2). To clarify the relationships between TFDaV and YMaCV, a data set including Rep amino acid sequences of all isolates of these two viruses plus those of nanovirids and alphasatellitids was constructed. TFDaV and YMaCV isolates formed two distinct, highly supported clades (Figure 3). Pairwise identity comparisons using Sequence Demarcation Tool (SDT) v. 1.2 [9] also indicated two highly divergent groups corrresponding to TFDaV and YMaCV isolates, which share <50% identity (Figure 4).  Phylogenetic analysis was also performed for the CP. A dataset of CP amino acid sequences was constructed including all TFDaV and YMaCV isolates plus representatives of all 14 genera in the *Geminiviridae*. The most suitable amino acid substitution model was LG+I+G4+F. Although the CPs of both TFDaV and YMaCV are related to those of geminivirids, they showed distinct clustering patterns. While the YMaCV CP is more related to those of mastreviruses, the TFDaV CP forms a more external group with no obvious relationship with the CPs of viruses in any genus (Figure 5).  The incongruence between the Rep and CP phylogenetic trees suggests that the evolutionary history of TFDaV and YMaCV includes a recombination event involving ancestors of a nanovirid and a geminivirid. However, no recombination events were reported for either virus [3, 4]. This could be the result of long divergence times that erased the evidence of the event, due to the fact that the common ancestor is unknown, because it has not yet been sampled or because it is extinct.  Together, these analyses indicate that TFDaV and YMaCV should be classified as members of two genera, with the suggested names *Temfrudevirus* and *Yermavirus*, respectively, in a new family, to which we propose to name *Amesuviridae*. The threshold for genus demarcation is proposed to be set at 55% pairwise identity for the Rep amino acid sequence. We also propose that the new family *Amesuviridae* should be classified in the order *Mulpavirales* (which currently includes the families *Metaxyviridae* and *Nanoviridae*), class *Arfiviricetes*, phylum *Cressdnaviricota*, kingdom *Shotokuvirae*, realm *Monodnaviria*.  **Derivation of names:**  *Amesuviridae*: from **Amé**rica do **Su**l or **Ame**rica del **Su**r (South America in Portuguese and Spanish, respectively), the continent in which the two viruses were described.  *Temfrudevirus*: from ***Tem****perate* ***fru****it* ***de****cay-associated* ***virus***  *Yermavirus*: from ***Yer****ba* ***ma****te-associated circular DNA* ***virus*** | |

**References**

1. Kazlauskas D, Varsani A, Koonin EV, Krupovic M (2019) Multiple origins of prokaryotic and eukaryotic single-stranded DNA viruses from bacterial and archaeal plasmids. Nat Commun 10:3425

2. Krupovic M, Varsani A, Kazlauskas D, Breitbart M, Delwart E, Rosario K, Yutin N, Wolf YI, Harrach B, Zerbini FM, Dolja VV, Kuhn JH, Koonin EV (2020) *Cressdnaviricota*: a virus phylum unifying 7 families of Rep-encoding viruses with single-stranded, circular DNA genomes. J Virol 94:e00582-20

3. Basso MF, Silva JC, Fajardo TV, Fontes EP, Zerbini FM (2015) A novel, highly divergent ssDNA virus identified in Brazil infecting apple, pear and grapevine. Virus Res 210:27-33

4. Bejerman N, de Breuil S, Nome C (2018) Identification and molecular characterization of a novel circular single-stranded DNA virus associated with yerba mate in Argentina. Arch Virol 163:2811-2815

5. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772-780

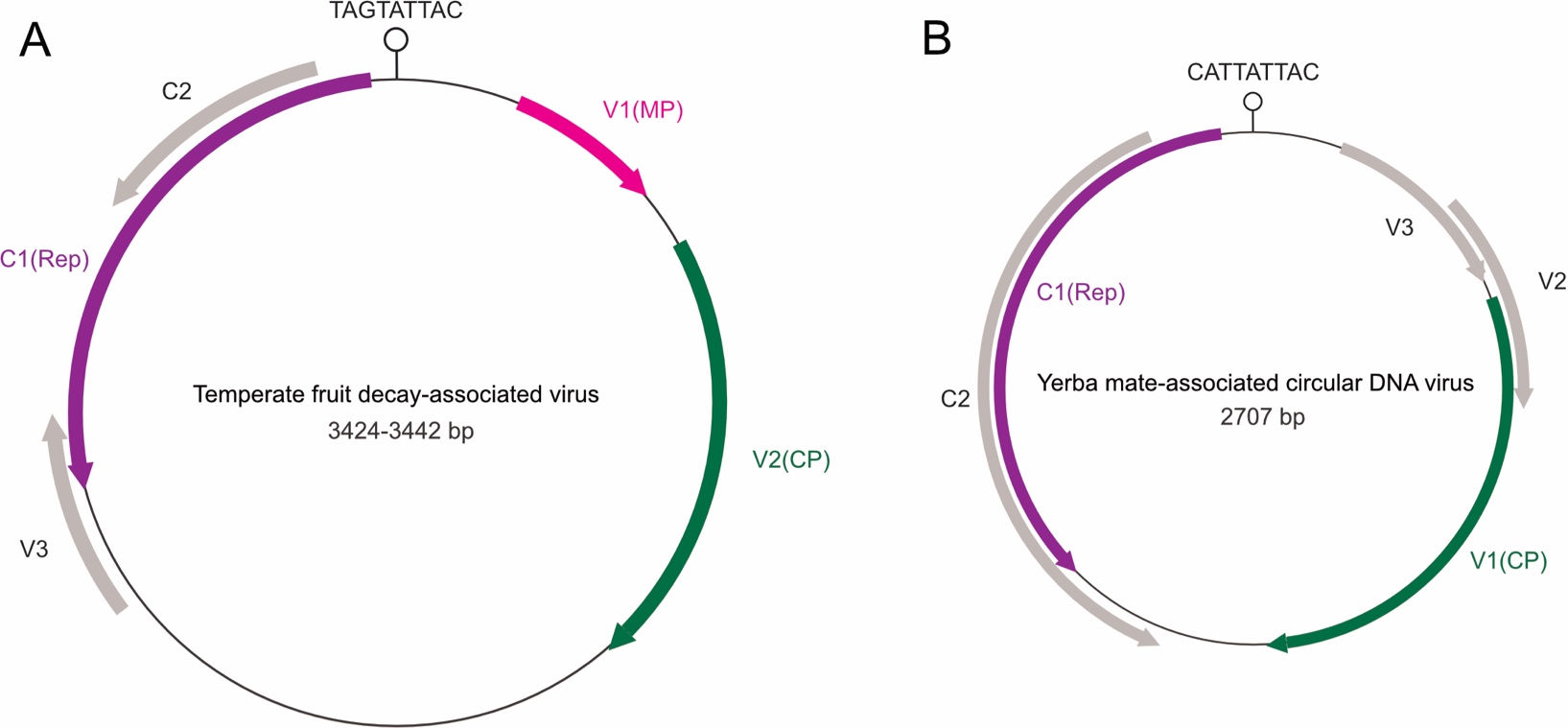
6. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972-1973

7. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T (2020) ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models. Mol Biol Evol 37:291-294

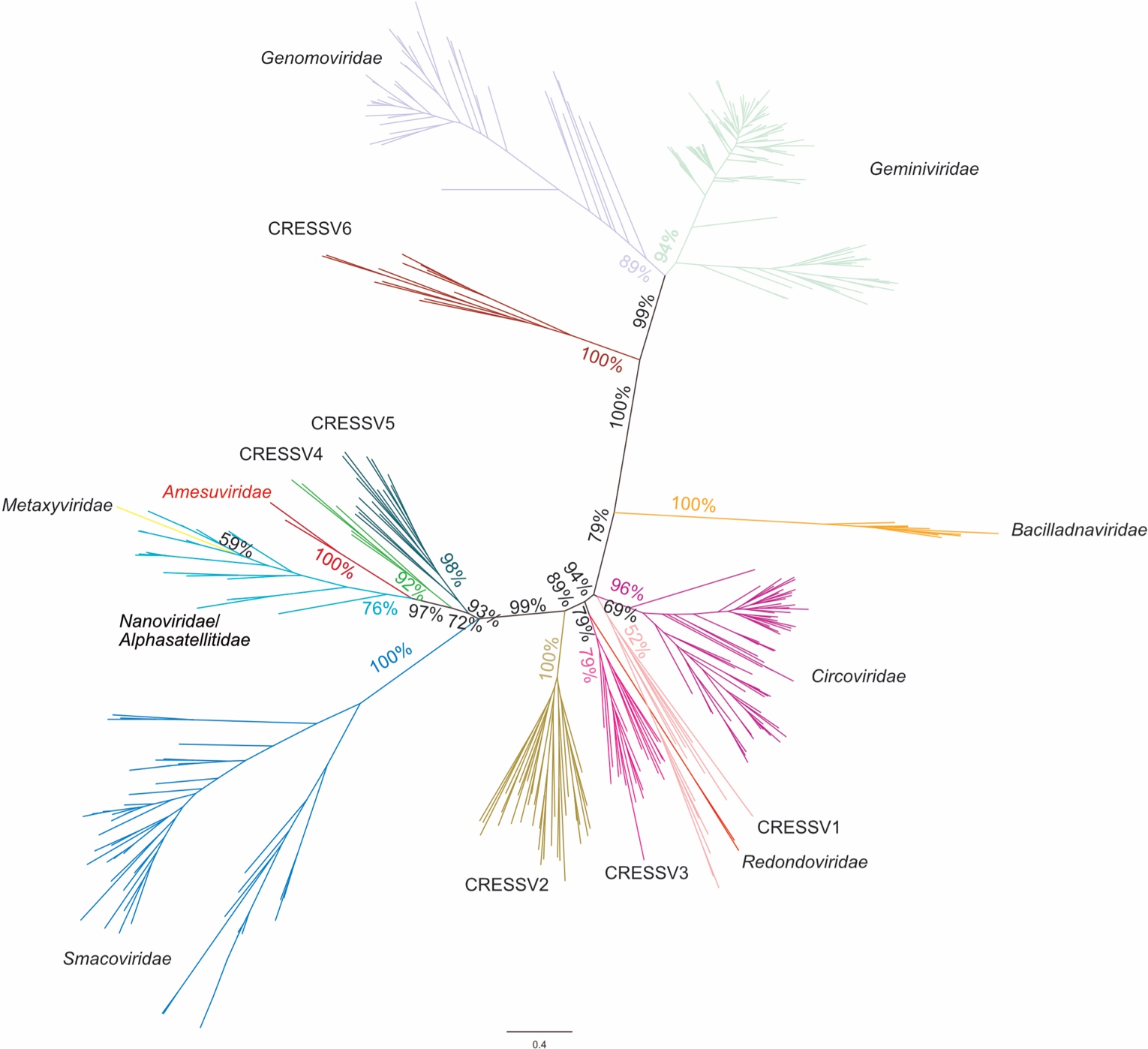
8. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A (2019) RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453-4455

9. Muhire BM, Varsani A, Martin DP (2014) SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 9:e108277

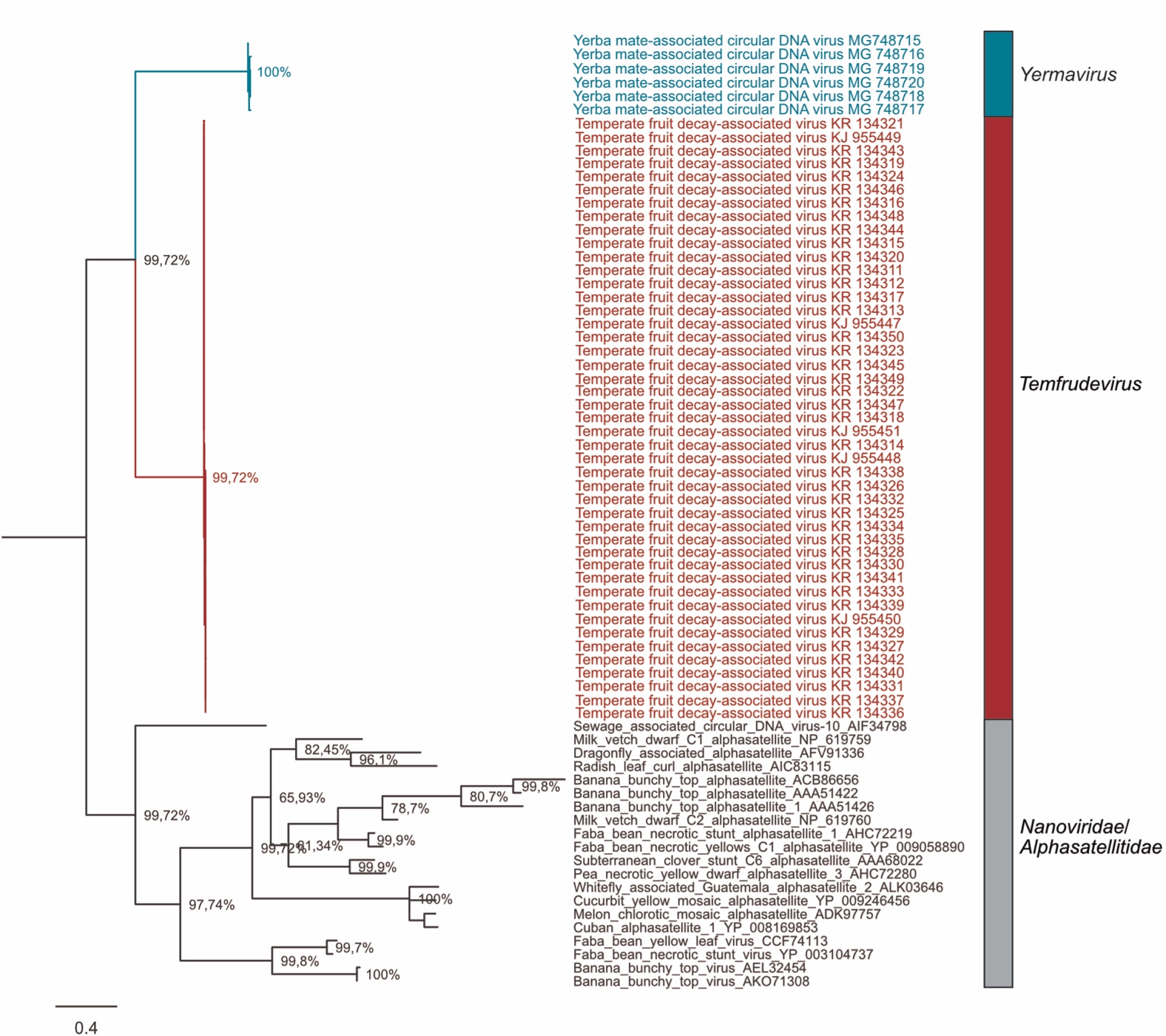
**Supporting evidence**



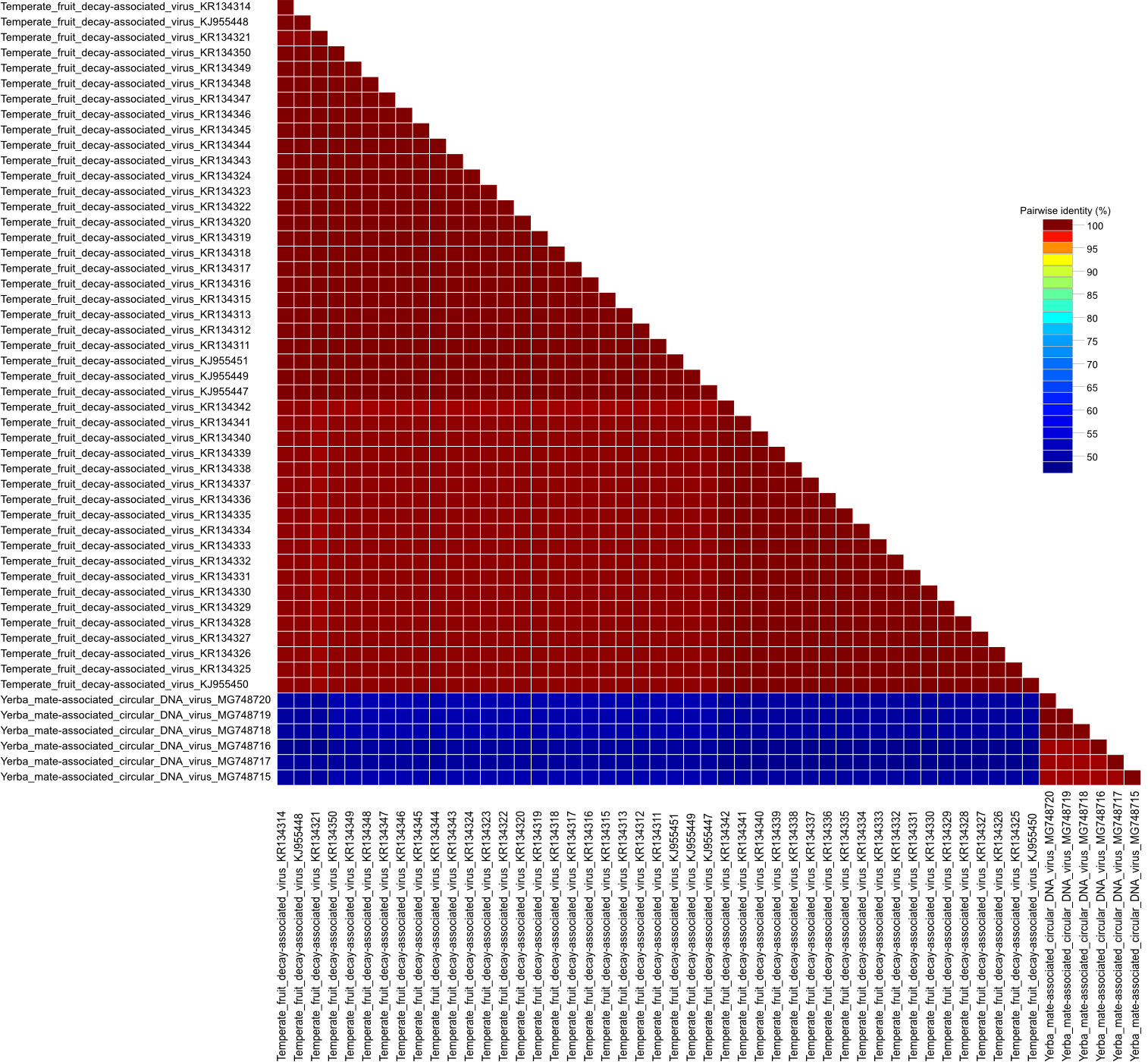
**Figure 1.** Genome organization of **(A)** temperate fruit decay-associated virus (TFDaV) and **(B)** yerba mate-associated circular DNA virus (YMaDV), the two members of the proposed new family *Amesuviridae*. The arrows indicate open reading frames (ORFs), with those encoding products that are homologous to known replication associated proteins indicated in purple, coat proteins in green and movement proteins in pink. ORFs in gray encode putative proteins with no detectable homology with any other proteins in the databases. The hairpin structure containing the putative origin of virion strand replication is also indicated.



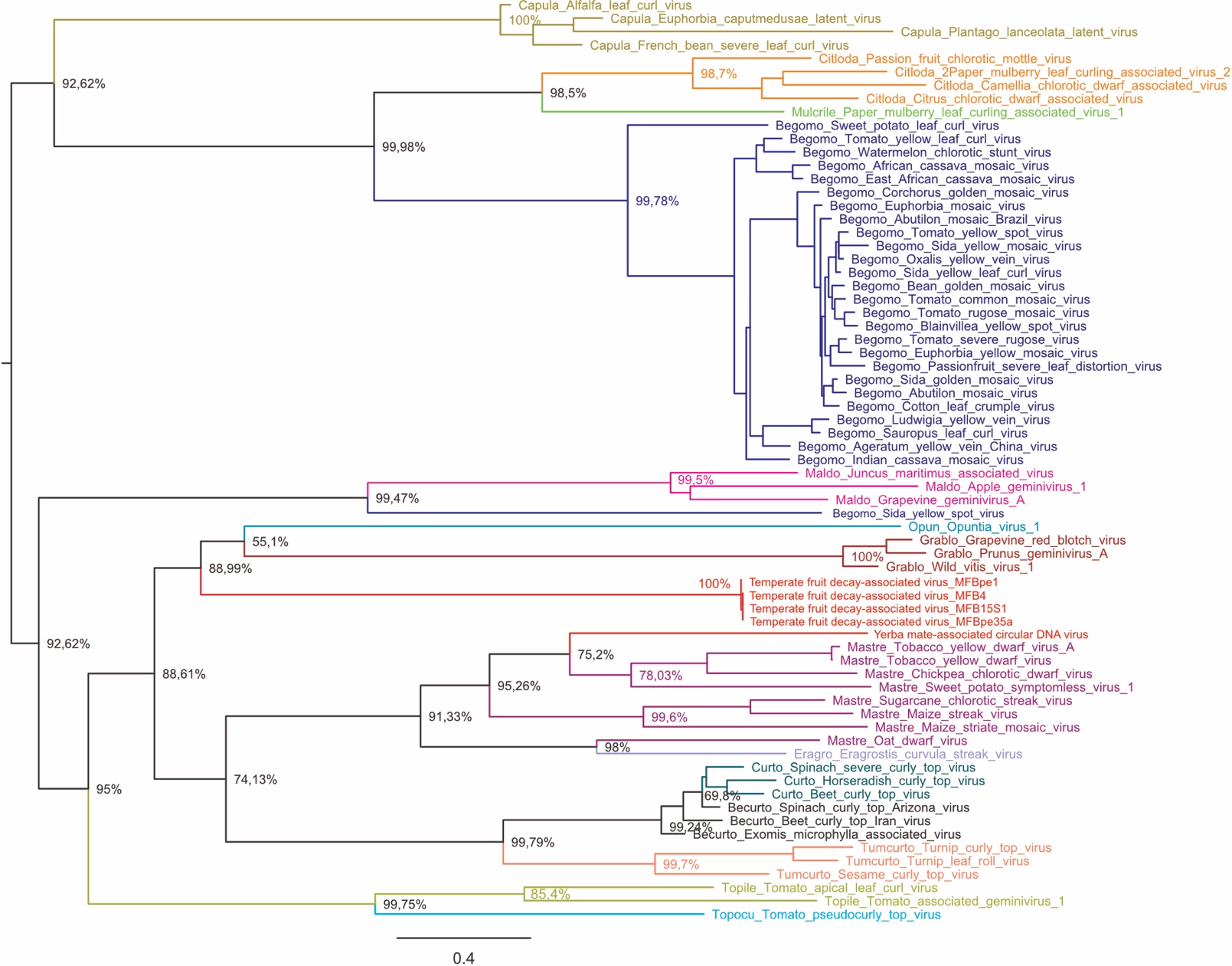
**Figure 2.** Unrooted maximum-likelihood phylogenetic tree based on the Rep amino acid sequences of members of the phylum *Cressdnaviricota* plus non-classified CRESS-DNA viruses (CRESSV1-6) reconstructed with RaxML-NG using the RTREV+G4+F substitution model. Numbers associated with branches indicate degrees of bootstrap support for these branches. The scale bar indicates substitutions per site. The branch representing viruses in the proposed family *Amesuviridae* is indicated in red.



**Figure 3.** Phylogenetic tree based on the Rep amino acid sequences from temperate fruit decay-associated virus (TFDaV) and yerba mate-associated circular DNA virus (YMaCV) isolates, rooted with the Rep sequences of nanovirids and alphasatellitids. Numbers associated with branches indicate degrees of bootstrap support for these branches. The scale bar indicates substitutions per site. The vertical bar at the right indicates the two proposed new genera in the family *Amesuviridae*.



**Figure 4.** Matrix of pairwise amino acid sequence identities of Rep proteins of temperate fruit decay-associated virus (TFDaV) and yerba mate-associated circular DNA virus (YMaCV), determined using SDT v1.2.



**Figure 5.** Mid-point rooted maximum-likelihood phylogenetic tree based on the CP amino acid sequences of temperate fruit decay-associated virus (TFDaV), yerba mate-associated circular DNA virus (YMaCV) isolates, and representatives of all 14 genera of the family *Geminiviridae* (indicated by colours and by the prefix of each genus name in front of the virus names). The tree was constructed using RaxML-NG and the LG+I+G4+F substitution model. Numbers associated with branches indicate degrees of bootstrap support for those branches. The scale bar represents substitutions per site.