

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

|  |  |  |
| --- | --- | --- |
| **Code assigned:** | ***2022.007M*** |  |
| **Short title:** Create ten new species in the genus *Cytorhabdovirus*, subfamily *Betarhabdovirinae* (*Mononegavirales*: *Rhabdoviridae*) |
|  |

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

|  |  |
| --- | --- |
|  |  |
| Bejerman N, Debat H, Dietzgen RG, Kondo H, Ramos-González P, Whitfield AE, Walker PJ, Freitas-Astúa J | bejerman.nicolas@inta.gob.ar;debat.humberto@inta.gob.ar;r.dietzgen@uq.edu.au; hkondo@rib.okayama-u.ac.jp; plrg1970@gmail.com;awhitfi@ncsu.edu; peter.walker@uq.edu.au;juliana.astua@embrapa.br |

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

|  |
| --- |
|  |

**Corresponding author**

|  |
| --- |
| Bejerman N |

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

|  |
| --- |
| ICTV *Rhabdoviridae* Study Group |

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

|  |
| --- |
| Minor corrections regarding the creation of new species. |

**ICTV Study Group votes on proposal**

|  |  |
| --- | --- |
| **Study Group** | **Number of members** |
| **Votes support** | **Votes against** | **No vote** |
| ICTV *Rhabdoviridae* Study Group | 13 | 0 | 1 |

**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 |

|  |  |  |
| --- | --- | --- |
| **Taxon name** | **Person from whom the name is derived** | **Permission attached (Y/N)** |
|  |  |  |

**Submission dates**

|  |  |
| --- | --- |
| Date first submitted to SC Chair | May 23, 2022 |
| Date of this revision (if different to above) |  |

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

|  |
| --- |
|  |

**Part 2:** **NON-TAXONOMIC PROPOSAL**

**Text of proposal**

|  |
| --- |
|  |

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

**Name of accompanying Excel module**

|  |
| --- |
| 2022.007M.N.v1.Cytorhabdovirus\_10nsp.xlsx |

**Abstract**

|  |
| --- |
| Viruses classified in the family *Rhabdoviridae* infect vertebrates, invertebrates, and plants. Ten new plant-infecting rhabdoviruses were discovered recently and their complete or coding-complete genome sequences were determined. This proposal aims to classify taxonomically these ten viruses in ten new species in the genus *Cytorhabdovirus*.  |

**Text of proposal**

|  |  |
| --- | --- |
|

|  |
| --- |
| We propose the taxonomic classification of ten novel plant-infecting rhabdoviruses in ten new species in the established genus *Cytorhabdovirus* in the *subfamily Betarhabdovirinae,* family *Rhabdoviridae*: **1) Actinidia virus D (AcVD)** was identified in kiwifruit (*Actinidia* spp.) plants collected in Jiangxi, China. The complete genome (CG) sequence of AcVD has 13,589 nucleotides (isolate JS27, GenBank #MW550041) and contains seven ORFs in the order 3’-N-P-P3-M-G-P6-L-5’ [1] representing the five canonical rhabdovirus structural protein genes as well as the putative cell-to-cell movement protein gene P3 in the conserved location between P and M genes, and the accessory protein P6 between the G and L genes (**Figure 1**). The CG nucleotide sequence of AcVD has the highest sequence identity with that of Wuhan insect virus 4 (WhIV4; 51.5%), whereas the AcVD large protein (L) amino acid sequence has the highest sequence identity with that of WhIV4 (44.5%) [1]. Based on ML trees generated from complete L sequences, AcVD is placed within a subclade of cytorhabdoviruses, with Asclepias syriaca virus 1, Bacopa monnieri virus 1, and WhIV4 (**Figure 2**).**2)** **Anthurium amnicola virus 1 (AntAmV1)** was identified during an *in silico* analysis of transcriptome data of flamingo lily (*Anthurium amnicola* Dressler.) tissues from Hawaii, USA. The coding-complete genome (CCG) sequence (partial 3’ leader and 5’ trailer) of AntAmV1 has 12,480 nucleotides (GenBank #BK014302) and contains six ORFs in the order 3’-N-P-P3-M-G-L-5’ [2] representing the five canonical rhabdovirus structural protein genes and the putative cell-to-cell movement protein gene P3 in the conserved location between P and M genes (**Figure 1**). The CCG nucleotide sequence of AntAmV1 has the highest sequence identity with that of paper mulberry mosaic-associated virus (PMuMaV; 51.8%). AntAmV1 L amino acid sequence has the highest sequence identity with that of PMuMaV (45.5%) [2]. Based on ML trees generated from complete L sequences, AntAmV1 forms a well-supported clade with the cytorhabdovirus PMuMaV (**Figure 2**).**3) Asclepias syriaca virus 1 (AscSyV1)** was identified during an *in silico* analysis of transcriptome data of common milkweed (*Asclepias syriaca* L.) tissues from Illinois, USA. The coding-complete genome (CCG) sequence (partial 3’ leader and 5’ trailer) of AscSyV1 has 13,392 nucleotides (GenBank #BK014298), and contains seven ORFs in the order 3’-N-P-P3-M-G-P6-L-5’ [2] representing the five canonical rhabdovirus structural protein genes as well as the putative cell-to-cell movement protein gene P3 in the conserved location between P and M genes, and the accessory protein P6 between the G and L genes (**Figure 1**). The CCG nucleotide sequence of AscSyV1 has the highest sequence identity with that of Bacopa monnieri virus 1 (BmV1; 54.8%), whereas the AscSyV1 L protein amino acid sequence has the highest sequence identity with that of BmV1 (53.9%) [2]. Based on ML trees generated from complete L sequences, AscSyV1 is placed within a subclade of cytorhabdoviruses, with Actinidia virus D, BmV1, and Wuhan insect virus 4 (**Figure 2**).**4) Bemisia tabaci-associated virus 1 (BeTaV1)** was identified during an *in silico* analysis of transcriptome data of silverleaf whiteflies (*Bemisia tabaci* Gennadius) collected in India. The plant host of this virus is unknown, but soybean was speculated as its potential host [2]. The coding-complete genome (CCG) sequence (partial 3’ leader and 5’ trailer) of BeTaV1 has 13,025 nucleotides (GenBank #BK014303) and contains seven ORFs in the order 3’-N-P-P3-P4-M-G-L-5’ [2] representing the five canonical rhabdovirus structural protein genes as well as the putative cell-to-cell movement protein gene P3 in the conserved location between P and M genes, and the accessory protein P4 between the P3 and M genes (**Figure 1**). The CCG nucleotide sequence of BeTaV1 has the highest sequence identity with that of cucurbit cytorhabdovirus 1 (CuCV1; 61.8%), whereas the BeTaV1 L amino acid sequence has the highest sequence identity with that of CuCV1 (72.5%) [2]. Based on ML trees generated from complete L protein sequences, BeTaV1 forms a well-supported clade with the cytorhabdovirus CuCV1 (**Figure 2**).**5) Glehnia littoralis virus 1 (GlLV1)** was identified during an *in silico* analysis of transcriptome data of beach silvertop (*Glehnia littoralis* F.Schmidt ex Miq.) tissues from Fujian, China. The coding-complete genome (CCG) sequence (partial 3’ leader and 5’ trailer) of GlLV1 has 12,193 nucleotides (GenBank #BK014304), and contains seven ORFs in the order 3’-N-P-P3-M-G-P6-L-5’ [2] representing the five canonical rhabdovirus structural protein genes as well as the putative cell-to-cell movement protein gene P3 in the conserved location between P and M genes, and the accessory protein P6 between the G and L genes (**Figure 1**). The CCG nucleotide sequence of GlLV1 has the highest sequence identity with that of Trifolium pratense virus A (TpVA; 65.6%), whereas the GlLV1 L protein amino acid sequence has the highest sequence identity with that of TpVA (76.1%) [2]. Based on ML trees generated from complete L sequences, GlLV1 forms a well-supported clade with the cytorhabdovirus TpVA (**Figure 2**).**6) Gymnadenia densiflora virus 1 (GymDenV1)** was identified during an *in silico* analysis of transcriptome data of marsh fragrant orchid [*Gymnadenia densiflora* (Wahlenb.) Dietrich.] tissues from Switzerland. Interestingly the intergenic sequence of GymDenV1 gene junction starts with an adenine instead of the typical guanine residue observed in every other known plant rhabdovirus [2]. The coding-complete genome (CCG) sequence (partial 3’ leader and 5’ trailer) of GymDenV1 has 9,887 nucleotides (GenBank #BK014305) and contains four ORFs in the order 3’-N-P-M-L-5’ [2] (**Figure 1**). The CCG nucleotide sequence of GymDenV1 has the highest sequence identity with that of Trachyspermum ammi virus 1 (TrAV1; 53.8%), whereas its L amino acid sequence has the highest sequence identity with that of TrAV1 (56.3%) [2]. Based on ML trees generated from complete L sequences, GymDenV1 forms a well-supported clade with TrAV1 (**Figure 2**).**7) Kenyan potato cytorhabdovirus (KePCyV)** was identified in potato (*Solanum tuberosum* L.) plants collected in Kenya. The coding complete genome (CCG) sequence of KePCyV has 13,131 nucleotides (isolate 18-1062, GenBank #MN689395) and contains seven ORFs in the order 3’-N-P-P3-M-G-P6-L-5’ [unpublished] representing the five canonical rhabdovirus structural protein genes as well as the putative cell-to-cell movement protein gene P3 in the conserved location between P and M genes, and the accessory protein P6 between the G and L genes (**Figure 1**). The CCG nucleotide sequence of KePCyV has the highest sequence identity with that of strawberry virus 1 (StrV1; 55.5%), whereas the KePCyV L protein amino acid sequence has the highest sequence identity with that of StrV1 (56.22%) [unpublished]. Based on ML trees generated from complete L sequences KePCyV forms a well-supported clade with the cytorhabdoviruses StrV1 and tomato yellow mottle-associated virus (**Figure 2**).**8) Nymphaea alba virus 1 (NymAV1)** was identified during an *in silico* analysis of transcriptome data of European white waterlily (*Nymphaea alba* L.) tissues from Turkey. The coding-complete genome (CCG) sequence (partial 3’ leader and 5’ trailer) of NymAV1 has 12,886 nucleotides (GenBank #BK014307), and contains seven ORFs in the order 3’-N-P-P3-M-G-P6-L-5’ [2] representing the five canonical rhabdovirus structural protein genes as well as the putative cell-to-cell movement protein gene P3 in the conserved location between P and M genes, and the accessory protein P6 between the G and L genes (**Figure 1**). The CCG nucleotide sequence of NymAV1 has the highest sequence identity with that of Trifolium pretense virus A (TpVA; 53.8%), whereas the NymAV1 L protein amino acid sequence has the highest sequence identity with that of TpVA (50.1%) [2]. Based on ML trees generated from complete L sequences, NymAV1 is placed within a subclade of the cytorhabdoviruses, with Glehnia littoralis virus 1, Kenyan potato cytorhabdovirus, strawberry virus 1, tomato yellow mottle-associated virus, and TpVA (**Figure 2**).**9) Tagetes erecta virus 1 (TaEV1)** was identified during an *in silico* analysis of transcriptome data of African marigold (*Tagetes erecta* L.) tissues from Anhui, China. The coding-complete genome (CCG) sequence (partial 3’ leader and 5’ trailer) of TaEV1 has 11,707 nucleotides (GenBank #BK014308), and contains five ORFs in the order 3’-N-P-P3-M-L-5’ [2] (**Figure 1**). The CCG nucleotide sequence of TaEV1 has the highest sequence identity with that of maize yellow striate virus (MYSV; 50.3%), whereas the TaEV1 L amino acid sequence has the highest sequence identity with that of MYSV (41.7%) [2]. Based on ML trees generated from complete L sequences, TaEV1 forms a well-supported clade with other cytorhabdoviruses (**Figure 2**).**10) Trachyspermum ammi virus 1 (TrAV1)** was identified during an *in silico* analysis of transcriptome data of Ajowan (*Trachyspermum ammi* L.) tissues from Tehran, Iran. The coding-complete genome (CCG) sequence (partial 3’ leader and 5’ trailer) of TrAV1 has 10,920 nucleotides (GenBank #BK014309), and contains five ORFs in the order 3’-N-P-P3-M-L-5’ [2] (**Figure 1**). Interestingly the intergenic sequence of TrAV1 gene junction starts with an adenine instead of the typical guanine residue observed in every other plant rhabdovirus [2]. The CCG nucleotide sequence of TrAV1 has the highest sequence identity with that of Gymnadenia densiflora virus 1 (GymDenV1; 53.8%), whereas its L amino acid sequence has the highest sequence identity with that of GymDenV1 (56.3%) [2]. Based on ML trees generated from complete L sequences, TrAV1 forms a well-supported clade with GymDenV1 (**Figure 2**).Viruses assigned to different species within the genus *Cytorhabdovirus*have several of the following characteristics: 1. nucleotide sequence identity less than 75% for the complete genome sequence;
2. amino acid sequence identity less than 80% in proteins encoded by all the cognate open reading frames;
3. occupy different ecological niches as evidenced by differences in hosts and/or arthropod vectors.

AcVD, AntAmV1, AscSyV1, BeTaV1, GlLV1, GymDenV1, KePCyV, NymAV1, TaEV1, and TrAV1 meet criteria A and B. Consequently, we propose to classify AcVD, AntAmV1, AscSyV1, BeTaV1, GlLV1, GymDenV1, KePCyV, NymAV1, TaEV1, and TrAV1 into ten new cytorhabdovirus species. As the viruses have been detected only by metagenomic sequencing, their natural ecology is uncertain but they have been detected in different hosts and geographic regions. |

 |

**Supporting evidence**

****

**Figure 1**. Genome graphs depicting architecture and gene products of viruses to be included in proposed species within the genus C*ytorhabdovirus*.Abbreviations: N: nucleoprotein; P: phosphoprotein; P3: putative cell-to-cell movement protein; P4: protein 4; M: matrix protein; G: glycoprotein; P6: Protein 6; L: large protein. Virus name abbreviations: Actinidia virus D (AcVD), Anthurium amnicola virus 1 (AntAmV1), Asclepias syriaca virus 1 (AscSyV1), Bemisia tabaci-associated virus 1 (BeTaV1), Glehnia littoralis virus 1 (GlLV1), Gymnadenia densiflora virus 1 (GymDenV1), Kenyan potato cytorhabdovirus (KePCyV), Nymphaea alba virus 1 (NymAV1), Tagetes erecta virus (TaEV1) and Trachyspermum ammi virus 1 (TrAV1)

****

**Figure 2.** AMaximum Likelihood (ML) phylogenetic tree of plant-infecting rhabdovirus L protein sequences. Amino acid sequences were aligned using MUSCLE. The resulting alignment was used to generate a phylogenetic tree using MegaX with the best-fit model LG + G + I +F. Ten viruses potentially belonging to the new species are indicated with green squares. Numbers at the nodes indicate bootstrap support (1000 replicates).

**References**

[1] Wang Y, Wang G, Bai J, Zhang Y, Wang Y, et al. (2021). A novel Actinidia cytorhabdovirus characterized using genomic and viral protein interaction features. Mol Plant Pathol 22:1271-1287. PMID: 34288324, doi: 10.1111/mpp.13110.

[2] Bejerman N, Dietzgen RG, Debat H (2021). Illuminating the plant rhabdovirus landscape through metatranscriptomics data. Viruses 13:1303. PMID: 34372509, doi: 10.3390/v13071304.