

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

|  |  |  |
| --- | --- | --- |
| **Code assigned:** | ***2022.021M*** |  |
| **Short title:** Create nine new species in the genus *Varicosavirus*, 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 |

**ICTV Study Group votes on proposal**

|  |  |
| --- | --- |
| **Study Group** | **Number of members** |
| **Votes support** | **Votes against** | **No vote** |
| ICTV *Rhabdoviridae* Study Group | 11 | 0 | 3 |

**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**

|  |
| --- |
| 2022.021M.N.v1.Varicosavirus\_9nsp.xlsx |

**Abstract**

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

**Text of proposal**

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

|  |
| --- |
| We propose the taxonomic classification of nine novel plant-infecting rhabdoviruses in nine new species in the established genus *Varicosavirus* in the *subfamily Betarhabdovirinae,* family *Rhabdoviridae*: **1)** **Allium angulosum virus 1 (AAnV1)** was identified during an *in silico* analysis of transcriptome data of mouse garlic (*Allium angulosum* L.) tissues from Brno, Czech Republic. The AAnV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6679 nucleotides and RNA2 of 4560 nucleotides (GenBank #BK059208 and BK059209) [1]. RNA1 contains one large ORF for the large protein (L), whereas RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of AAnV1 L has the highest sequence identity with that of Brassica rapa virus 1 (BrRV1; 61.61%) (**Table 1**). Based on ML trees generated from complete L sequences, AAnV1 is placed within a subclade of varicosaviruses with BrRV1 and red clover-associated varicosavirus (**Figure 2**).**2)** **Brassica rapa virus 1 (BrRV1)** was identified during an *in silico* analysis of transcriptome data of bok choy (*Brassica rapa* subsp*. chinensis* L.) tissues from Guangdong, China. The BrRV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6397 nucleotides and RNA2 of 4068 nucleotides (GenBank #BK014310 and BK014311) [1]. RNA1 contains one large ORF encoding L, whereas RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of BrRV1 L has the highest sequence identity with that of red clover-associated varicosavirus (RCaVV; 63%) (**Table 1**). Based on ML trees generated from complete L sequences, BrRV1 forms a well-supported clade with the varicosavirus RCaVV (**Figure 2**). **3)** **Lolium perenne virus 1 (LoPV1)** was identified during an *in silico* analysis of transcriptome data of perennial ryegrass (*Lolium perenne* L.) tissues from Store Heddinge, Denmark. The LoPV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6302 nucleotides and RNA2 of 4167 nucleotides (GenBank #BK014312 and BK014313) [1]. RNA1 contains one large ORF encoding L, whereas RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of LoPV1 L has the highest sequence identity with that of Alopecurus myesuroides varicosavirus 1 (AMVV1; 70.07%) (**Table 1**). Based on ML trees generated from complete L sequences, LoPV1 forms a well-supported clade with the varicosavirus AMVV1 (**Figure 2**). **4)** **Melampyrum roseum virus 1 (MelRoV1)** was identified during an *in silico* analysis of transcriptome data of Asian cow-wheat (*Melampyrum roseum* Maxim) tissues from Kanagawa, Japan. The MelRoV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6409 nucleotides and RNA2 of 5288 nucleotides (GenBank #BK014314 and BK014315) [1]. RNA1 contains one large ORF encoding L, whereas RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of MelRoV1 L has the highest sequence identity with that of Zostera-associated varicosavirus 1(ZaVV1; 62.11%) (**Table 1**). Based on ML trees generated from complete L sequences, MelRoV1 is placed within a subclade of varicosaviruses with morning glory varicosavirus, Vitis varicosavirus and ZaVV1 (**Figure 2**). **5) Morning glory varicosavirus (MGVV)** was identified in tall morning glory plants (*Ipomoea purpurea* (L.) Roth) from Henan, China. The MGVV genome is bi-segmented with a complete genome (CG) RNA1 of 6365 nucleotides and RNA2 of 4635 nucleotides (GenBank #MW922438 and MW922439) [2]. RNA1 contains one large ORF encoding L, whereas RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of MGVV L has the highest sequence identity with that of Zostera-associated varicosavirus 1(ZaVV1; 62.05%) (**Table 1**). Based on ML trees generated from complete L sequences, MGVV is placed within a subclade of varicosaviruses with Melampyrum roseum virus 1, Vitis varicosavirus, and ZaVV1 (**Figure 2**).**6)** **Pinus flexilis virus 1 (PiFleV1)** was identified during an *in silico* analysis of transcriptome data of limber pine (*Pinus flexilis* E. James) from Oregon, USA. The coding-complete genome (CCG) sequence of PiFleV1 has 11,740 nucleotides (GenBank #BK014316) [1], and contains five ORFs (**Figure 1**). The PiFleV1 is unique among the varicosaviruses in terms of genome organization, since its genome is unsegmented, a characteristic that differs from all other known varicosaviruses, which have bi-segmented genomes [1]. The nucleotide sequence of PiFleV1 L protein has the highest sequence identity with that of Lolium perenne virus 1 (59.86%) (**Table 1**). Based on ML trees generated from complete L sequences, PiFleV1 forms a well-supported clade with other varicosaviruses (**Figure 2**). **7)** **Vitis varicosavirus (VVV)** was identified in crimson glory vine plants (*Vitis coignetiae* Pulliat ex Planch) from Hokkaido, Japan. The VVV genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6450 nucleotides and RNA2 of 5535 nucleotides (GenBank #LC604719 and LC604720) [3]. RNA1 contains one large ORF encoding L, whereas RNA2 contains five ORFs (**Figure 1**). The nucleotide sequence of VVV has the highest sequence identity with that of Zostera-associated varicosavirus 1 (ZaVV1; 62.57%) (**Table 1**). Based on ML trees generated from complete sequences, VVV is placed within a subclade of varicosaviruses with Melampyrum roseum virus 1, morning glory varicosavirus, and ZaVV1 (**Figure 2**).**8)** **Xinjiang varicosavirus (XVV)** was identified from soil samples collected in Xinjiang, China. The XVV genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6729 nucleotides and RNA2 of 5849 nucleotides (GenBank #MW897032 and MW897033) [4]. RNA1 contains one large ORF encoding L, whereas RNA2 contains five ORFs (**Figure 1**). The nucleotide sequence of XVV L has the highest sequence identity with that of lettuce big vein-associated virus (LBVaV; 66.45%) (**Table 1**). Based on ML trees generated from complete L sequences, XVV forms a well-supported clade with the varicosavirus LBVaV (**Figure 2**).**9)** **Zostera-associated varicosavirus 1 (ZaVV1)** was identified during an *in silico* analysis of transcriptome data of common eelgrass (*Zostera marina* L.) tissues from Shandong, China. The ZaVV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6632 nucleotides and RNA2 of 4304 nucleotides (GenBank #BK014484 and BK014485) [5]. RNA1 contains one large ORF encoding L, whereas RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of ZaVV1 L has the highest sequence identity with that of Vitis varicosavirus (VVV; 62.57%) (**Table 1**). Based on ML trees generated from complete L sequences, ZaVV1 is placed within a subclade of varicosaviruses with Melampyrum roseum virus 1, morning glory varicosavirus and VVV (**Figure 2**).Current species demarcation criteria proposed for the genus *Varicosavirus* by the ICTV Rhabdoviridae Study Group are:Viruses assigned to different species within the genus Varicosavirus have several of the following characteristics: A) the nucleotide sequence identity of complete genomes is lower than 75%; B) they occupy different ecological niches as evidenced by differences in hosts and/or chytrid vectors; and C) they can be clearly distinguished in serological tests or by nucleic acid hybridisation.AAnV1, BrRV1, LoPV1, MelRoV1, MGVV, PiFleV1, VVV, XVV, and ZaVV1 meet criteria A and B. Consequently, we propose to classify AAnV1, BrRV1, LoPV1, MelRoV1, MGVV, PiFleV1, VVV, XVV, and ZaVV1 into nine new species |

 |

**Supporting evidence**

Table 1. Varicosavirus *L* ORF Percent Identity Matrix - created by SDT v1.2

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | AAnV1 | AMVV1 | BrRV1 | LBVaV | LoPV1 | MelRoV1 | MGVV | PiFleV1 | RCaVV | VVV | XVV | ZaVV1 |
| AAnV1 | 100 |  |  |  |  |  |  |  |  |  |  |  |
| AMVV1 | 59.54 | 100 |  |  |  |  |  |  |  |  |  |  |
| BrRV1 | 61.61 | 59.09 | 100 |  |  |  |  |  |  |  |  |  |
| LBVaV | 57.51 | 59.05 | 57.94 | 100 |  |  |  |  |  |  |  |  |
| LoPV1 | 60.57 | 70.07 | 60.38 | 60.31 | 100 |  |  |  |  |  |  |  |
| MelRoV1 | 60.53 | 59.80 | 60.96 | 57.58 | 60 | 100 |  |  |  |  |  |  |
| MGVV | 59.75 | 60.27 | 60.53 | 59.71 | 59.17 | 61.84 | 100 |  |  |  |  |  |
| PiFleV1 | 58.94 | 59.04 | 58.37 | 58.93 | 59.86 | 57.86 | 57.77 | 100 |  |  |  |  |
| RCaVV | 60.61 | 60.67 | 63 | 58.81 | 60.64 | 59.57 | 59.88 | 59.84 | 100 |  |  |  |
| VVV | 60.6 | 58.38 | 59.22 | 58.73 | 58.00 | 61.52 | 61.26 | 58.00 | 59.68 | 100 |  |  |
| XVV | 58.35 | 58.54 | 58.64 | 66.45 | 59.38 | 59.48 | 60.14 | 59.19 | 58.49 | 59.02 | 100 |  |
| ZaVV1 | 59.41 | 59.65 | 59.7 | 57.97 | 60.36 | 62.11 | 62.05 | 58.57 | 60.20 | 62.57 | 59.67 | 100 |

Virus name abbreviations: Allium angulosum virus 1 (AAnV1), Allopecurus myesuroides varicosavirus 1 (AMVV1), Brassica rapa virus 1 (BrRV1), lettuce big vein-associated virus (LBVaV), Lolium perenne virus 1 (LoPV1), Melampyrum roseum virus 1 (MelRoV1), morning glory varicosavirus (MGVV), Pinus flexilis virus 1 (PiFleV1), red clover-associated varicosavirus (RCaVV), Vitis varicosavirus (VVV), Xinjiang varicosavirus (XVV) and Zostera-associated varicosavirus (ZaVV1).

****

**Figure 1**. Genome graphs depicting architecture and gene products of viruses proposed to be included in species within genus *Varicosavirus.* Abbreviations: N: nucleoprotein; P2: protein 2; P3: protein 3; P4: protein 4; P5: protein 5; L: large protein. Virus name abbreviations: Allium angulosum virus 1 (AAnV1), Brassica rapa virus 1 (BrRV1), Lolium perenne virus 1 (LoPV1), Melampyrum roseum virus 1 (MelRoV1), morning glory varicosavirus (MGVV), Pinus flexilis virus 1 (PiFleV1), Vitis varicosavirus (VVV), Xinjiang varicosavirus (XVV) and Zostera-associated varicosavirus (ZaVV1).

****

**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. Nine viruses potentially belonging to the new species are indicated with green squares. Numbers at the nodes indicate bootstrap support (1000 replicates).

**References**

[1] Bejerman N, Dietzgen RG, Debat H (2021). Illuminating the Plant Rhabdovirus Landscape through Metatranscriptomics Data. Viruses 13:1303. PMID: 34372509, doi: 10.3390/v13071304.

[2] Zhao F, Liu H, Qiao Q, Wang Y, Zhang D, Wang S, Tian Y, Zhang Z (2021) Complete genome sequence of a novel varicosavirus infecting tall morning glory (*Ipomoea purpurea*). Arch Virol 166:3225-3228. PMID: 34519925, doi: 10.1007/s00705-021-05240-6.

[3] Nabeshima T, Abe J (2021). High-Throughput Sequencing Indicates Novel Varicosavirus, Emaravirus, and Deltapartitivirus Infections in *Vitis coignetiae*. Viruses 13:827. PMID: 34063722. doi: 10.3390/v13050827.

[4] Chen Y, Sadiq S, Tian J, Chen X, Lin X, et al (2021) RNA Virome Composition Is Shaped by Sampling Ecotype. SSRN Electronic Journal. <http://dx.doi.org/10.2139/ssrn.3934022>.

[5] Shin C, Choi D, Hahn Y (2021) Identification of the genome sequence of Zostera associated varicosavirus 1, a novel negative-sense RNA virus, in the common eelgrass (Zostera marina) transcriptome. Acta Virol 65:373-380. PMID: 34796713. doi: 10.4149/av\_2021\_404.