

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

|  |  |  |
| --- | --- | --- |
| **Code assigned:** | ***2023.022M*** |  |
| **Short title:** Create thirty-two 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.hmberto@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 | 10 | 0 | 4 |

**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)** |
| N/A | N/A | N/A |

**Submission dates**

|  |  |
| --- | --- |
| Date first submitted to SC Chair | June 23, 2023 |
| Date of this revision (if different to above) |  |

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

|  |
| --- |
|  |

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

|  |
| --- |
| N/A |

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

**Name of accompanying Excel module**

|  |
| --- |
| 2023.022M.N.v1.Varicosavirus\_32nsp.xlxs |

**Abstract**

|  |
| --- |
| Viruses classified in the family *Rhabdoviridae* infect vertebrates, invertebrates, and plants. Thirty-two novel plant-infecting rhabdoviruses were discovered recently and their coding-complete genomes were determined. This proposal aims to classify taxonomically these viruses into thirty-two new species in the genus *Varicosavirus.* |

**Text of proposal**

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

|  |
| --- |
| We propose the taxonomic classification of thirty-two novel plant-infecting rhabdoviruses into thirty-two new species in the established genus *Varicosavirus* in the subfamily *Betarhabdovirinae,* family *Rhabdoviridae*: **1)** **Aconitum virus 1 (AcoV1)** was identified from an *in silico* analysis of transcriptome data of Bei Wu Tou (*Aconitum kusnezoffii*) tissues from Jilin, China. AcoV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,483 nucleotides and RNA2 of 5,561 nucleotides (BK061734 and BK061735) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains five ORFs (**Figure 1**). The nucleotide sequence of AcoV1 L ORF has the highest sequence identity with that of Zostera-associated varicosavirus 1 (ZaVV1; 61.16%) [1]. Based on ML trees generated from complete L protein sequences, AcoV1 is placed within a subclade of varicosaviruses with ZaVV1 and vitis varicosavirus (**Figure 2**).**2)** **Apera virus 1 (ApeV1)** was identified from an *in silico* analysis of transcriptome data of common windgrass (*Apera spica-venti*) tissues from Denmark. ApeV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,512 nucleotides and RNA2 of 6,552 nucleotides (BK061737 and BK061738) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains five ORFs (**Figure 1**). The nucleotide sequence of ApeV1 L ORF has the highest sequence identity with that of Melampyrum roseum virus 1 (MelRoV1; 52.47%) [1]. Based on ML trees generated from complete L protein sequences, ApeV1 is placed within a subclade of varicosaviruses with MelRoV1 and Guizotia virus 1 (**Figure 2**).**3)** **Aponogeton virus 1 (ApoV1)** was identified from an *in silico* analysis of transcriptome data of lace plant (*Aponogeton madagascariensis*) tissues from Canada. ApoV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,678 nucleotides and RNA2 of 5,628 nucleotides (BK061739 and BK061740) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of ApoV1 L ORF has the highest sequence identity with that of Brassica virus 1 (52.50%) [1]. Based on ML trees generated from complete L protein sequences, ApoV1 forms a well-supported clade with other varicosaviruses (**Figure 2**).**4)** **Arceuthobium virus 8 (ArcV8)** was identified from an *in silico* analysis of transcriptome data of dwarf mistletoe (*Arceuthobium sichuanence*) tissues from Qinghai, China. ArcV8 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,628 nucleotides and RNA2 of 4,149 nucleotides (BK061732 and BK061733) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of ArcV8 L ORF has the highest sequence identity with that of Brassica virus 1 (45.44%) [1]. Based on ML trees generated from complete L protein sequences, ArcV8 forms a well-supported clade with other varicosaviruses (**Figure 2**).**5) Artemisia virus 1 (ArtV1)** was identified from an *in silico* analysis of transcriptome data of wormwood (*Artemisia absinthium*) tissues from the United Kingdom. ArtV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 7,373 nucleotides and RNA2 of 4,497 nucleotides (BK061741 and BK061742) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of ArtV1 L ORF has the highest sequence identity with that of Tanacetum virus 1 (TanV1, 65.64%) [1]. Based on ML trees generated from complete L protein sequences, ArtV1 is placed within a subclade of varicosaviruses with TanV1 and Leucanthemum virus 1 (**Figure 2**).**6)** **Asclepias syriaca virus 3 (AscSyV3)** was identified from an *in silico* analysis of transcriptome data of common milkweed (*Asclepias syriaca*) tissues from Illinois, USA. AscSyV3 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,506 nucleotides and RNA2 of 6,280 nucleotides (BK0617443 and BK061744) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains five ORFs (**Figure 1**). The nucleotide sequence of AscSyV3 L ORF has the highest sequence identity with that of Primula virus 1 (PriV1, 59.01%) [1]. Based on ML trees generated from complete L protein sequences, AscSyV3 is placed within a subclade of varicosaviruses with PriV1 and Brassica virus 2 (**Figure 2**). **7)** **Brassica virus 2 (BrV2)** was identified from an *in silico* analysis of transcriptome data of several Brassica species. Three strains of this virus, named as BrV2\_Inc; BrV2\_Jun and BrV2\_Ole, were identified, one from transcriptome data of shortpod mustard (*Hirschfeldia incana*), another one from Indian mustard (*Brassica juncea* var. *rugosa*), and the third one from Chinese kale (*Brassica oleracea* var. *alboglabra*) tissues from the USA, India and Hunan, China, respectively. BrV2\_Inc; BrV2\_Jun and BrV2\_Ole genomes are bi-segmented with a coding-complete genome (CCG) RNA1 of 6,316 nucleotides and RNA2 of 5,616, 5,537 and 5,647 nucleotides, respectively (BK061747, BK061748, BK061749, BK061750, BK061751 and BK061752) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequences of the L proteins of the tree BrV2 strains share identities that ranged from 98.57 to 99.75%, while the L ORFs of the three strains have the highest sequence identity with that of Asclepias syriaca virus 3 (AscSyV3, 56.04%, 55.84% and 56.14%, respectively) [1]. Based on ML trees generated from complete L protein sequences, BrV2 is placed within a subclade of varicosaviruses with AscSyV3 and Primula virus 1 (**Figure 2**).**8)** **Caladenia virus (CalV1)** was identified from an *in silico* analysis of transcriptome data of crab-lipped spider orchid (*Caladenia plicata*) tissues from Western Australia. CalV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,454 nucleotides and RNA2 of 5,011 nucleotides (BK061755 and BK061756) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of CalV1 L ORF has the highest sequence identity with that of Artemisia virus 1 (ArtV1, 55.15%) [1]. Based on ML trees generated from complete L protein sequences, CalV1 is placed within a subclade of varicosaviruses with ArtV1, Leucanthemum virus 1 and Tanacetum virus 1 (**Figure 2**).**9)** **Centaurea virus 1 (CenV1)** was identified from an *in silico* analysis of transcriptome data of cornflower (*Centaurea cyanus*) tissues from New South Wales, Australia. CenV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,789 nucleotides and RNA2 of 4,567 nucleotides (BK061757 and BK061758) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of CenV1 L ORF has the highest sequence identity with that of Brassica virus 1 (50.77%) [1]. Based on ML trees generated from complete L protein sequences, CenV1 forms a well-supported clade with other varicosaviruses (**Figure 2**).**10)** **Cucumis virus 1 (CucV1)** was identified from an *in silico* analysis of transcriptome data of melon (*Cucumis melo*) tissues from China. CucV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,919 nucleotides and RNA2 of 5,322 nucleotides (BK061761 and BK061762) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of CucV1 L ORF has the highest sequence identity with that of Luffa virus 1 (LufV1, 62.74%) [1]. Based on ML trees generated from complete L protein sequences, CucV1 is placed within a subclade of varicosaviruses with LufV1 and Streptoglossa virus 1 (**Figure 2**).**11)** **Didymochlaena virus 1 (DidV1)** was identified from an *in silico* analysis of transcriptome data of tree maidenhair fern (*Didymochlaena truncatula*) tissues from Indonesia. DidV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,319 nucleotides and RNA2 of 5,924 nucleotides (BK061764 and BK061765) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains five ORFs (**Figure 1**). The nucleotide sequence of DidV1 L ORF has the highest sequence identity with that of tree fern varicosa-like virus (TfVV, 74.16%) [1]. Based on ML trees generated from complete L protein sequences, DidV1 is placed within a subclade of varicosaviruses with TfVV and Treubia virus 1 (**Figure 2**).**12)** **Erysimum virus 1 (EryV1)** was identified from an *in silico* analysis of transcriptome data of wallflower (*Erysimum bastetanum*) tissues from Spain. EryV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,676 nucleotides and RNA2 of 3,980 nucleotides (BK061766 and BK061767) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of EryV1 L ORF has the highest sequence identity with that of Brassica virus 1 (BrV1, 62.22%) [1]. Based on ML trees generated from complete L protein sequences, EryV1 is placed within a subclade of varicosaviruses with BrV1 and Raphanus virus 1 (**Figure 2**).**13)** **Frullania virus 1 (FruV1)** was identified from an *in silico* analysis of transcriptome data of the liverwort *Frullania orientalis* from Sichuan, China. FruV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,458 nucleotides and RNA2 of 4363 nucleotides (BK061768 and BK061769) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of FruV1 L ORF has the highest sequence identity with that of Monoclea gottschei varicosa-like virus (MgVV, 54.83%) [1]. Based on ML trees generated from complete L protein sequences, FruV1 forms a well-supported clade with the proposed varicosavirus MgVV (**Figure 2**).**14)** **Guizotia virus 1 (GuiV1)** was identified from an *in silico* analysis of transcriptome data of noug (*Guizotia abyssinica*) tissues from Australia. GuiV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,457 nucleotides and RNA2 of 4,722 nucleotides (BK061770 and BK061771) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of GuiV1 L ORF has the highest sequence identity with that of Melampyrum roseum virus 1 (MelRoV1; 60.08%) [1]. Based on ML trees generated from complete L protein sequences, GuiV1 is placed within a subclade of varicosaviruses with MelRoV1 and Apera virus 1 (**Figure 2**).**15)** **Holcus virus 1 (HolV1)** was identified from an *in silico* analysis of transcriptome data of common velvet grass (*Holcus lanatus*) tissues from Northern Ireland, United Kingdom. HolV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,571 nucleotides and RNA2 of 4,397 nucleotides (BK061772 and BK061773) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of HolV1 L ORF has the highest sequence identity with that of Alopecurus myosuroides varicosavirus 1 (AMVV1, 64.76%) [1]. Based on ML trees generated from complete L protein sequences, HolV1 is placed within a subclade of varicosaviruses with AMVV1 and Lolium virus 1 (**Figure 2**).**16)** **Leucanthemum virus 1 (LeuV1)** was identified from an *in silico* analysis of transcriptome data of oxeye daisy (*Leucanthemum vulgare*) tissues from New South Wales, Australia. LeuV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,763 nucleotides and RNA2 of 4,775 nucleotides (BK061774 and BK061775) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of LeuV1 L ORF has the highest sequence identity with that of Tanacetum virus 1 (TanV1, 70.49%) [1]. Based on ML trees generated from complete L protein sequences, LeuV1 is placed within a subclade of varicosaviruses with TanV1, Artemisia virus 1 and Caladenia virus 1 (**Figure 2**).**17)** **Luffa virus 1 (LufV1)** was identified from an *in silico* analysis of transcriptome data of sponge gourd (*Luffa aegyptiaca*) tissues from Australia. LufV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,693 nucleotides and RNA2 of 4,961 nucleotides (BK061780 and BK061781) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of LufV1 L ORF has the highest sequence identity with that of Cucumis virus 1 (CucV1, 62.74%) [1]. Based on ML trees generated from complete L protein sequences, LufV1 is placed within a subclade of varicosaviruses with CucV1 and Streptoglossa virus 1 (**Figure 2**).**18)** **Melilotus virus 1 (MelV1)** was identified from an *in silico* analysis of transcriptome data of two Melilotus species. Two strains of this virus, named as MelV1\_Alb and MelV1\_Off, were identified from transcriptome data of honey clover (*Melilotus albus*), and yellow sweet clover (*Melilotus officinalis*) tissues from China and Canada, respectively. MelV1\_Alb and MelV1\_Off genomes are bi-segmented with a coding-complete genome (CCG) RNA1 of 6,657 and 6,433 nucleotides, respectively, and RNA2 of 3,985 and 3,781 nucleotides, respectively (BK061784, BK061785, BK061786 and BK061787) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of the L ORF of both MelV1 strains share an identity of 91.67%, while the L ORF of both MelV1\_Alb and MelV1\_Off have the highest sequence identity with that of red clover-associated varicosavirus (RCaVV, 64.71% and 65.11%, respectively) [1]. Based on ML trees generated from complete L protein sequences, MelV1 is placed within a subclade of varicosaviruses with RCaVV and Triticum virus 1 (**Figure 2**).**19)** **Monoclea gottschei varicosa-like virus (MgVV)** was identified from an *in silico* analysis of transcriptome data of the liverwort *Monoclea gottschei*. MgVV genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,297 nucleotides and RNA2 of 4,979 nucleotides (OX380363 and OX380364) [2]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of MgVV L ORF has the highest sequence identity with that of Frullania virus 1 (FruV1, 54.83%). Based on ML trees generated from complete L protein sequences, MgVV forms a well-supported clade with the proposed varicosavirus FruV1 (**Figure 2**).**20)** **Pennisetum virus 1 (PenV1)** was identified from an *in silico* analysis of transcriptome data of purple grass (*Pennisetum violaceum*) tissues from India. PenV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,284 nucleotides and RNA2 of 3,407 nucleotides (BK061790 and BK061791) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of PenV1 L ORF has the highest sequence identity with that of Lolium virus 1 (LolV1, 51.13%) [1]. Based on ML trees generated from complete L protein sequences, PenV1 is placed within a subclade of varicosaviruses with LolV1, Alopecurus myosuroides varicosavirus 1 and Holcus virus 1 (**Figure 2**).**21)** **Primula virus 1 (PriV1)** was identified from an *in silico* analysis of transcriptome data of splendor primrose (*Primula oreodoxa*) tissues from Sichuan, China. PriV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,352 nucleotides and RNA2 of 6,283 nucleotides (BK061795 and BK061796) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains five ORFs (**Figure 1**). The nucleotide sequence of PriV1 L ORF has the highest sequence identity with that of Asclepias syriaca virus 3 (AscSyV3, 59.01%) [1]. Based on ML trees generated from complete L protein sequences, PriV1 is placed within a subclade of varicosaviruses with AscSyV3 and Brassica virus 2 (**Figure 2**).**22)** **Ranunculus virus 1 (RanV1)** was identified from an *in silico* analysis of transcriptome data of goldilocks buttercup (*Ranunculus auricomus*) tissues from Slovakia. RanV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,481 nucleotides and RNA2 of 6,269 nucleotides (BK061797 and BK061798) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains five ORFs (**Figure 1**). The nucleotide sequence of RanV1 L ORF has the highest sequence identity with that of Vincetoxicum virus 1 (VinV1, 52.24%) [1]. Based on ML trees generated from complete L protein sequences, RanV1 forms a well-supported clade with the proposed varicosavirus VinV1 (**Figure 2**).**23)** **Raphanus virus 1 (RapV1)** was identified from an *in silico* analysis of transcriptome data of radish (*Raphanus sativus*) tissues from China. RapV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,410 nucleotides and RNA2 of 4,144 nucleotides (BK061799 and BK061800) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of RapV1 L ORF has the highest sequence identity with that of Brassica virus 1 (BrV1, 68.70%) [1]. Based on ML trees generated from complete L protein sequences, RapV1 is placed within a subclade of varicosaviruses with BrV1 and Erysimum virus 1 (**Figure 2**).**24)** **Ribes virus 1 (RibV1)** was identified from an *in silico* analysis of transcriptome data of Siberian currant (*Ribes diacanthum*) tissues from China. RibV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,323 nucleotides and RNA2 of 5,201 nucleotides (BK061801 and BK061802) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of RibV1 L ORF has the highest sequence identity with that of Aponogeton virus 1 (48.99%) [1]. Based on ML trees generated from complete L protein sequences, RibV1 forms a well-supported clade with other varicosaviruses (**Figure 2**).**25)** **Silene virus 1 (SilV1)** was identified from an *in silico* analysis of transcriptome data of bladder campion (*Silene vulgaris*) tissues from Slovakia. SilV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,391 nucleotides and RNA2 of 4,363 nucleotides (BK061807 and BK061808) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of SilV1 L ORF has the highest sequence identity with that of Spinach virus 1 (SpV1, 59.73%) [1]. Based on ML trees generated from complete L protein sequences, SilV1 forms a well-supported clade with the varicosavirus SpV1 (**Figure 2**).**26)** **Streptoglossa virus 1 (StrV1)** was identified from an *in silico* analysis of transcriptome data of broadhead daisy (*Streptoglossa macrocephala*) tissues from Western Australia. StrV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,776 nucleotides and RNA2 of 5,130 nucleotides (BK061813 and BK061814) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of StrV1 L ORF has the highest sequence identity with that of Luffa virus 1 (LufV1, 52.71%) [1]. Based on ML trees generated from complete L protein sequences, StrV1 is placed within a subclade of varicosaviruses with LufV1 and Cucumis virus 1 (**Figure 2**).**27)** **Tanacetum virus 1 (TanV1)** was identified from an *in silico* analysis of transcriptome data of tansy (*Tanacetum vulgare*) tissues from Germany. TanV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,888 nucleotides and RNA2 of 4,608 nucleotides (BK061815 and BK061816) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of TanV1 L ORF has the highest sequence identity with that of Leucanthemum virus 1 (LeuV1, 70.49%) [1]. Based on ML trees generated from complete L protein sequences, TanV1 is placed within a subclade of varicosaviruses with LeuV1, Artemisia virus 1 and Caladenia virus 1 (**Figure 2**).**28)** **Tree fern varicosa-like virus** **(TfVV)** was identified from an *in silico* analysis of transcriptome data of tree fern (*Thyrsopteris elegans*) tissues. TfVV genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,395 nucleotides and RNA2 of 5,794 nucleotides (OW528630 and OW528632) [2]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains five ORFs (**Figure 1**). The nucleotide sequence of TfVV L ORF has the highest sequence identity with that of Didymochlaena virus 1 (DidV1, 74.16%) [1]. Based on ML trees generated from complete L protein sequences, TfVV is placed within a subclade of varicosaviruses with DidV1 and Treubia virus 1 (**Figure 2**).**29)** **Treubia virus 1 (TreV1)** was identified from an *in silico* analysis of transcriptome data of the liverwort *Treubia lacunosa* from China. TreV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,684 nucleotides and RNA2 of 4,940 nucleotides (BK061819 and BK061820) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of TreV1 L ORF has the highest sequence identity with that of tree fern varicosa-like virus (TfVV, 53.91%) [1]. Based on ML trees generated from complete L protein sequences, TreV1 is placed within a subclade of varicosaviruses with TfVV and Didymochlaena virus 1 (**Figure 2**).**30)** **Triticum virus 1 (TriV1)** was identified from an *in silico* analysis of transcriptome data of wheat (*Triticum aestivum*) tissues from Ohio, USA. TriV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,290 nucleotides and RNA2 of 4,103 nucleotides (BK061821 and BK061822) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains three ORFs (**Figure 1**). The nucleotide sequence of TriV1 L ORF has the highest sequence identity with that of red clover-associated varicosavirus (RCaVV, 72.54%) [1]. Based on ML trees generated from complete L protein sequences, TriV1 is placed within a subclade of varicosaviruses with RCaVV and Melilotus virus 1 (**Figure 2**).**31)** **Vincetoxicum virus 1 (VinV1)** was identified from an *in silico* analysis of transcriptome data of variegated swallow-wort (*Vincetoxicum versicolor*) tissues from China. RapV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,598 nucleotides and RNA2 of 4,655 nucleotides (BK061823 and BK061824) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of VinV1 L ORF has the highest sequence identity with that of Ranunculus virus 1 (RanV1, 52.24%) [1]. Based on ML trees generated from complete L protein sequences, VinV1 forms a well-supported clade with the varicosavirus RanV1 (**Figure 2**).**32)** **Zea virus 1 (ZeaV1)** was identified from an *in silico* analysis of transcriptome data of corn (*Zea mays*) tissues from Germany. ZeaV1 genome is bi-segmented with a coding-complete genome (CCG) RNA1 of 6,345 nucleotides and RNA2 of 4,607 nucleotides (BK061825 and BK061826) [1]. RNA1 contains one large ORF for the L polymerase protein while RNA2 contains four ORFs (**Figure 1**). The nucleotide sequence of ZeaV1 L ORF has the highest sequence identity with that of Holcus virus 1 (HolV1, 51.40%) [1]. Based on ML trees generated from complete L protein sequences, ZeaV1 is placed within a subclade of varicosaviruses with HolV1, AMVV1, LolV1 and PenV1 (**Figure 2**).We propose to classify AcoV1, ApeV1, ApoV1, ArcV8, ArtV1, AscSyV3, BrV2, CalV1, CenV1, CucV1, DidV1, EryV1, FruV1, GuiV1, HolV1, LeuV1, LufV1, MelV1, MgVV, PenV1, PriV1, RanV1, RapV1, RibV1, SilV1, StrV1, TanV1, TfVV, TreV1, TriV1, VinV1, and ZeaV1 in the new species *Varicosavirus aconiti, Varicosavirus aperae, Varicosavirus aponogeti, Varicosavirus arceuthobii, Varicosavirus artemisiae, Varicosavirus asclepiadis, Varicosavirus betabrassicae, Varicosavirus caladeniae, Varicosavirus centaureae, Varicosavirus cucumis, Varicosavirus didymochlaenae, Varicosavirus erysimi, Varicosavirus frullaniae, Varicosavirus guizotiae, Varicosavirus holci, Varicosavirus leucanthemi, Varicosavirus luffae, Varicosavirus meliloti, Varicosavirus monocleae, Varicosavirus penniseti, Varicosavirus primulae, Varicosavirus ranunculi, Varicosavirus raphani, Varicosavirus ribes, Varicosavirus silenis, Varicosavirus streptoglossae, Varicosavirus tanaceti, Varicosavirus thrysopteris, Varicosavirus treubiae, Varicosavirus tritici, Varicosavirus vincetoxici,* and *Varicosavirus zeae,* respectively,in the genus *Varicosavirus,* subfamily *Betarhabdovirinae,* family *Rhabdoviridae*. Viruses assigned to different species within the genus *Varicosavirus* have several of the following characteristics: 1. nucleotide sequence identity lower than 75% for the L ORF
2. occupy different ecological niches as evidenced by differences in hosts
3. can be clearly distinguished in serological tests or by nucleic acid hybridization

AcoV1, ApeV1, ApoV1, ArcV8, ArtV1, AscSyV3, BrV2, CalV1, CenV1, CucV1, DidV1, EryV1, FruV1, GuiV1, HolV1, LeuV1, LufV1, MelV1, MgVV, PenV1, PriV1, RanV1, RapV1, RibV1, SilV1, StrV1, TanV1, TfVV, TreV1, TriV1, VinV1, and ZeaV1 meet criteria A and B. |

 |

**Supporting evidence**

****

**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: RNA-dependent RNA polymerase. Virus name abbreviations: Aconitum virus 1 (AcoV1), Apera virus 1 (ApeV1), Aponogeton virus 1 (ApoV1), Arceuthobium virus 8 (ArcV8), Artemisia virus 1 (ArtV1), Asclepias syriaca virus 3 (AscSyV3), Brassica virus 2 (BrV2), Caladenia virus 1 (CalV1), Centaurea (CenV1), Cucumis virus 1 (CucV1), Didymochlaena virus 1 (DidV1), Erysimum virus 1 (EryV1), Frullania virus 1 (FruV1), Guizotia virus 1 (GuiV1), Holcus virus 1 (HolV1), Leucanthemum virus 1 (LeuV1), Luffa virus 1 (LufV1), Melilotus virus 1 (MelV1), Monoclea gottschei varicosa-like virus (MgVV), Pennisetum virus 1 (PenV1), Primula virus 1 (PriV1), Ranunculus virus 1 (RanV1), Raphanus virus 1 (RapV1), Ribes virus 1 (RibV1), Silene virus 1 (SilV1), Streptoglossa virus 1 (StrV1), Tanacetum virus 1 (TanV1), tree fern varicosa-like virus (TfVV), Treubia virus 1 (TreV1), Triticum virus 1 (TriV1), Vincetoxicum virus 1 (VinV1), and Zea virus 1 (ZeaV1).

****

**Figure 2.** AMaximum Likelihood (ML) phylogenetic tree of plant-infecting rhabdovirus L polymerase 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. Thirty-two 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] Mifsud JCO, Gallagher RV, Holmes EC, Geoghegan JL (2022). Transcriptome Mining Expands Knowledge of RNA Viruses across the Plant Kingdom. Journal of Virology 96:e00260-22. PMID:35638822, doi 10.1128/jvi.00260-22.