

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

|  |  |  |
| --- | --- | --- |
| **Code assigned:** | ***2023.019D*** |  |
| **Short title:** Establishing one new family, associated genera and species in the order *Cirlivirales* | | |
|  | | |

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

|  |  |
| --- | --- |
| Varsani A, Hopkins A, Lund MC, Kraberger S, Krupovic M | Arvind.varsani@asu.edu;  adhopki1@asu.edu; mclund2@asu.edu;  Simona.kraberger@asu.edu; mart.krupovic@pasteur.fr |

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

|  |
| --- |
| The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life sciences, Arizona State University, Tempe, AZ 85287-5001, USA [AV, AH, MCL, SK]  Institut Pasteur, Université Paris Cité, Archaeal Virology Unit, 25 rue du Dr Roux, 75015 Paris, France [MK] |

**Corresponding author**

|  |
| --- |
| Arvind Varsani |

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

|  |
| --- |
|  |

**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** |
|  |  |  |  |
|  |  |  |  |

**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 | 14 July 2023 |
| Date of this revision (if different to above) | 5 Oct 2023 |

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

|  |
| --- |
| The proposal was deemed acceptable in the form presented at the EC meeting. |

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

**Text of proposal**

|  |
| --- |
|  |

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

**Name of accompanying Excel module**

|  |
| --- |
| 2023.019D.N.v2.Cirlivirales\_1nf\_4ng\_10nsp.xlsx |

**Abstract**

|  |
| --- |
| We aim to establish a new family, *Endolinaviridae,* in the order *Cirlivirales* to classify 39 new virus genomes that have been identified from mammals. The new family will include 4 genera and 10 species. |

**Text of proposal**

|  |  |
| --- | --- |
| |  | | --- | | *Circoviridae* and *Vilyaviridae* are the two families in the order *Cirlivirales*. The Reps of 39 viruses that are informally referred to as hudisaviruses [1-3] cluster with those of members of the families *Circoviridae* and *Vilyaviridae*.  Analyses by Kinsella et al. [1] predicted that hudiaviruses infect the protist *Endolimax nana*. Thus, we propose the name *Endolinaviridae,* derived from the host name ***Endoli****max* ***na****na.*  **Genus and species demarcation**  We performed comparative genomics and phylogenetic analyses of the Rep proteins to determine the relationships between viruses in the proposed family *Endolinaviridae*. Genera were delineated based on phylogenetic analyses coupled with pairwise identities and also the genome organization relative to the *rep* open reading frame (Figures 2 - 3).  For species demarcation, we used a 78% pairwise nucleotide genome-wide sequence identity which is similar to that used other cressdnaviruses [4-6].  Thus, all viral genomes showing identities higher than 78% should be considered as variants of the existing species. There may be situations where it is difficult to assign species because a particular new sequence is  1. >78% similar to sequences from a particular species but is <78% similar to other variants of that same species;  2. >78% similar to sequences from two or more different species.  To resolve the above conflicts, we suggest adopting a similar approach proposed for geminiviruses [5] and genomoviruses [5]. To resolve conflict 1, we suggest that the new sequence be classified within any species in which it shares >78% identity to any one variant formerly classified as belonging to that species, even if it is <78% identical to other viruses within that species. To resolve conflict 2, we suggest that the new sequence be considered as belonging to the species with sequences with which it shares the highest degree of similarity.  **Genera**  Etymology of genus names  We used the name ‘disa’ from the name hudisavirus that was given by Altan et al. [1-2] and add a prefix to it from the most common host / source for the viruses in the genus.   1. *Bovdisavirus:* ***Bov****ine* 2. *Camdisavirus:* ***Cam****el* 3. *Chidisavirus:* ***Chi****ken* 4. *Hudisavirus:* ***Hu****man*   Etymology of species epithets   1. *runais: human in Quechwa - runa* 2. *sewis: human in amharic - sewi* 3. *jaqis: human in Aymata - jaqi* 4. *niois: bird in Chiinese - niao* 5. *ryguais: chicken in Guarani - ryguasu* 6. *miluis: elk in chinese - milu* 7. *kauis: cow in Māori - kau* 8. *jamalis: camal inb arabic - jamal* 9. *lingis: monken in Thai - ling* 10. *nainis: cow in Chinese - nainiu* | |

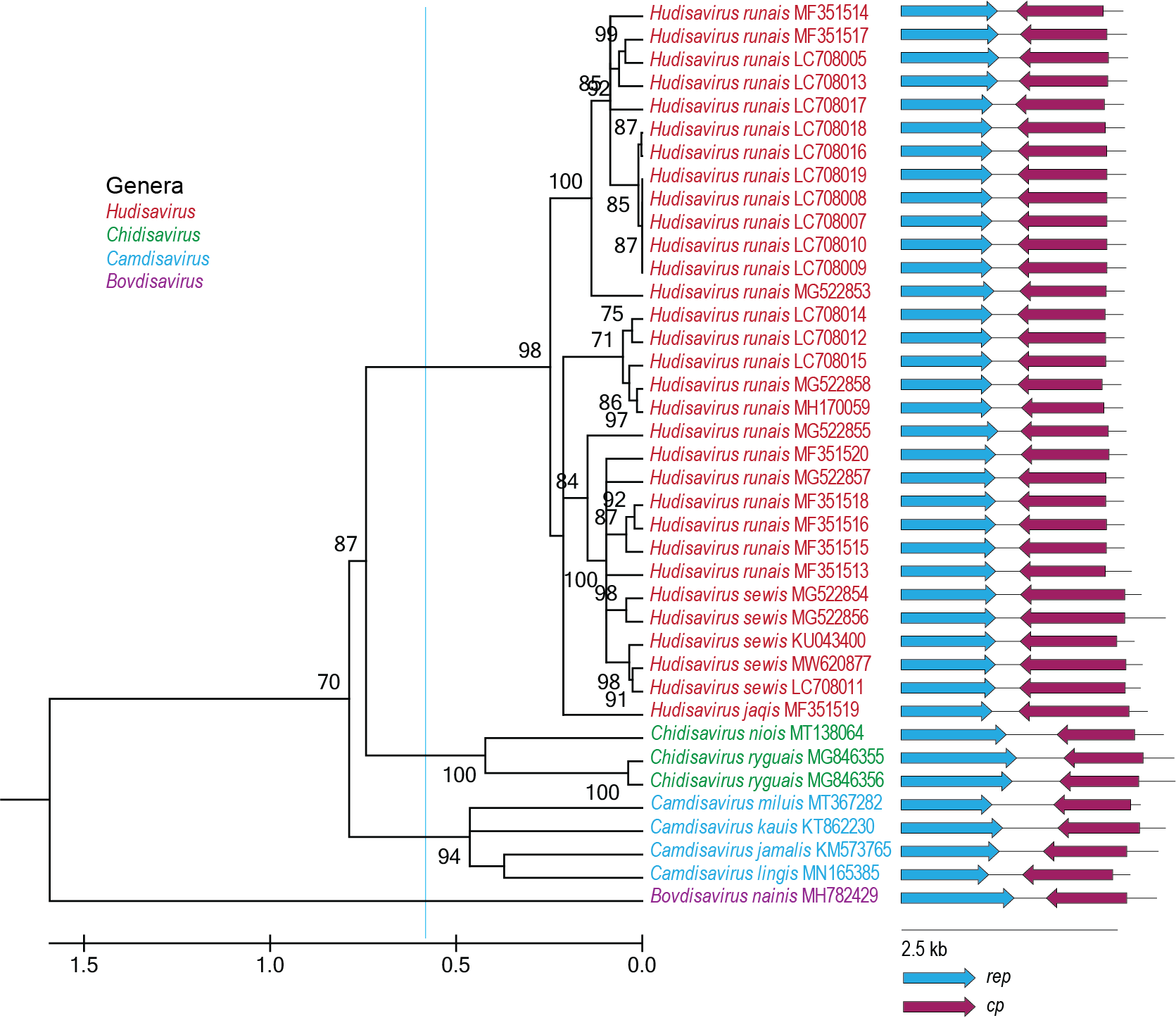
**Supporting evidence**

**Table 1:** Summary of the viruses classified in the new family *Endolinaviridae* in the order *Cirlivirales*.

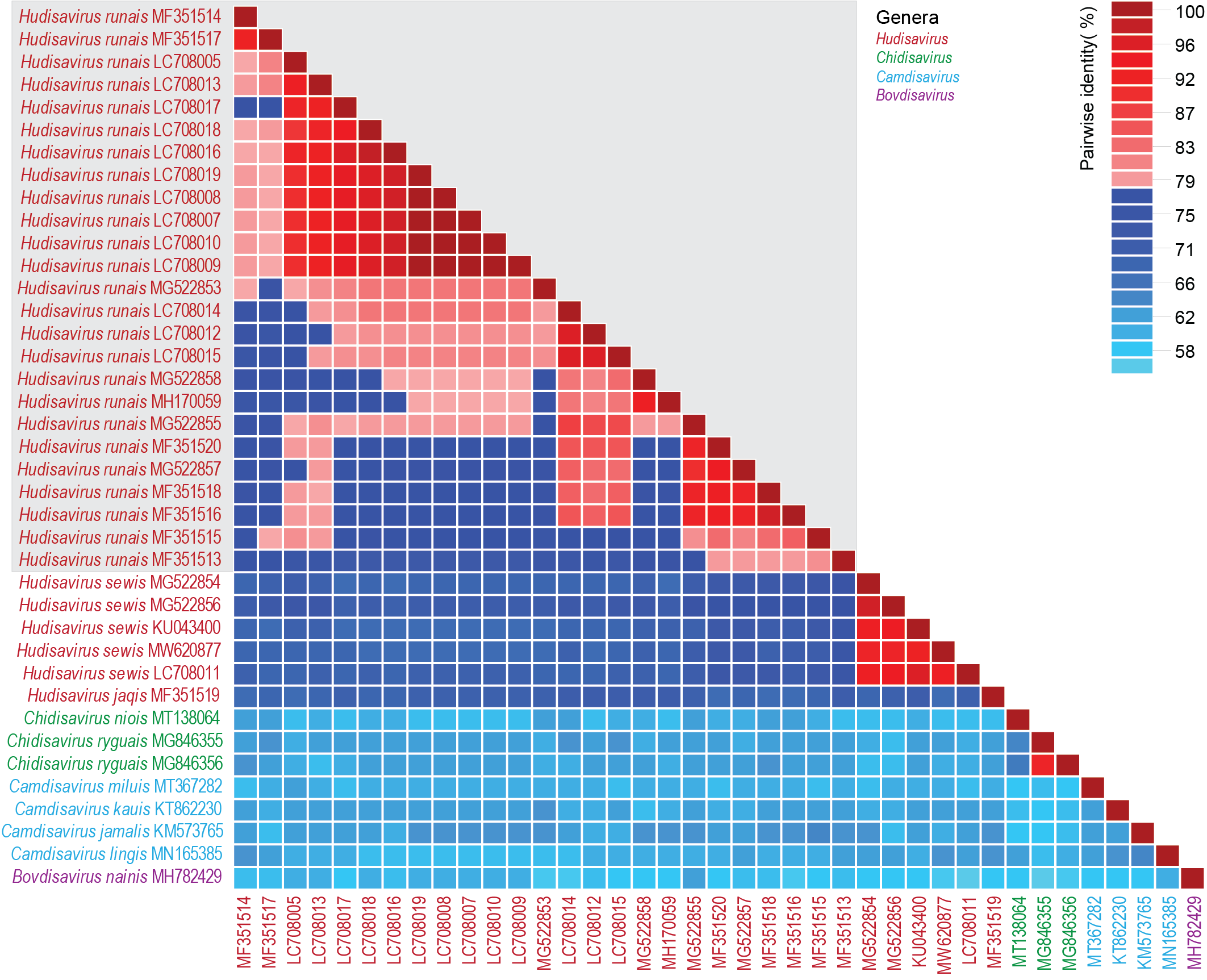
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Genus** | **Species** | **Accession #** | **Virus** | **Year** | **Country** | **Source** | **Isolate** |
| *Bovdisavirus* | *Bovdisavirus nainis* | MH782429 | Circoviridae sp. C042793 | 2016 | China | *Bos taurus* | C042793 |
| *Camdisavirus* | *Camdisavirus jamalis* | KM573765 | Dromedary stool-associated circular ssDNA virus DcSCV\_c941 | 2013 | United Arab Emirates | *Camelus dromedarius* | DcSCV\_c941 |
| *Camdisavirus* | *Camdisavirus kauis* | KT862230 | Bovine faeces associated circular DNA virus 1 GP3-46075\_cow2 | 2013 | New Zealand | *Bos taurus* | GP3-46075\_cow2 |
| *Camdisavirus* | *Camdisavirus lingis* | MN165385 | Macaca fascicularis stool-associated virus PPT018/MFS01 | 2017 | Thailand | *Macaca fascicularis* | PPT018/MFS01 |
| *Camdisavirus* | *Camdisavirus miluis* | MT367282 | CRESS virus sp. Elk10cre2 | 2015 | China | *Cervus canadensis* | Elk10cre2 |
| *Chidisavirus* | *Chidisavirus niois* | MT138064 | CRESS virus sp. zftwig07cir1 genomic sequence | 2016 | China | avian | zftwig07cir1 |
| *Chidisavirus* | *Chidisavirus ryguais* | MG846355 | Chicken stool associated circular virus 2 RS/BR/2015/1R | 2015 | Brazil | *Gallus gallus* | RS/BR/2015/1R |
| *Chidisavirus* | *Chidisavirus ryguais* | MG846356 | Chicken stool associated circular virus 1 RS/BR/2015/1R | 2015 | Brazil | *Gallus gallus* | RS/BR/2015/1R |
| *Hudisavirus* | *Hudisavirus jaqis* | MF351519 | Hudisavirus sp. P13\_1 | - | Peru | *Homo sapiens* | P13\_1 |
| *Hudisavirus* | *Hudisavirus runais* | MF351514 | Hudisavirus sp. P19 | - | Peru | *Homo sapiens* | P19 |
| *Hudisavirus* | *Hudisavirus runais* | MF351517 | Hudisavirus sp. P29 | - | Peru | *Homo sapiens* | P29 |
| *Hudisavirus* | *Hudisavirus runais* | LC708005 | Hudisavirus sp. CRESSV19-84-AMS-01 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-01 |
| *Hudisavirus* | *Hudisavirus runais* | LC708013 | Hudisavirus sp. CRESSV19-84-AMS-09 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-09 |
| *Hudisavirus* | *Hudisavirus runais* | LC708017 | Hudisavirus sp. CRESSV19-84-AMS-13 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-13 |
| *Hudisavirus* | *Hudisavirus runais* | LC708018 | Hudisavirus sp. CRESSV19-84-AMS-14 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-14 |
| *Hudisavirus* | *Hudisavirus runais* | LC708016 | Hudisavirus sp. CRESSV19-84-AMS-12 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-12 |
| *Hudisavirus* | *Hudisavirus runais* | LC708019 | Hudisavirus sp. CRESSV19-84-AMS-15 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-15 |
| *Hudisavirus* | *Hudisavirus runais* | LC708008 | Hudisavirus sp. CRESSV19-84-AMS-04 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-04 |
| *Hudisavirus* | *Hudisavirus runais* | LC708007 | Hudisavirus sp. CRESSV19-84-AMS-03 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-03 |
| *Hudisavirus* | *Hudisavirus runais* | LC708010 | Hudisavirus sp. CRESSV19-84-AMS-06 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-06 |
| *Hudisavirus* | *Hudisavirus runais* | LC708009 | Hudisavirus sp. CRESSV19-84-AMS-05 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-05 |
| *Hudisavirus* | *Hudisavirus runais* | MG522853 | Hudisavirus sp. ETH\_P4\_2016 | 2016 | Ethiopia | *Homo sapiens* | ETH\_P4\_2016 |
| *Hudisavirus* | *Hudisavirus runais* | LC708014 | Hudisavirus sp. CRESSV19-84-AMS-10 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-10 |
| *Hudisavirus* | *Hudisavirus runais* | LC708012 | Hudisavirus sp. CRESSV19-84-AMS-08 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-08 |
| *Hudisavirus* | *Hudisavirus runais* | LC708015 | Hudisavirus sp. CRESSV19-84-AMS-11 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-11 |
| *Hudisavirus* | *Hudisavirus runais* | MG522858 | Hudisavirus sp. ETH\_P25\_2016 | 2016 | Ethiopia | *Homo sapiens* | ETH\_P25\_2016 |
| *Hudisavirus* | *Hudisavirus runais* | MH170059 | Porcine serum-associated circular virus PoSCV 7B RS/BR | 2016 | Brazil | *Sus scrofa domesticus* | PoSCV 7B RS/BR |
| *Hudisavirus* | *Hudisavirus runais* | MG522855 | Hudisavirus sp. ETH\_P8\_2016 | 2016 | Ethiopia | *Homo sapiens* | ETH\_P8\_2016 |
| *Hudisavirus* | *Hudisavirus runais* | MF351520 | Hudisavirus sp. P13\_2 | - | Peru | *Homo sapiens* | P13\_2 |
| *Hudisavirus* | *Hudisavirus runais* | MG522857 | Hudisavirus sp. ETH\_P22\_2016 | 2016 | Ethiopia | *Homo sapiens* | ETH\_P22\_2016 |
| *Hudisavirus* | *Hudisavirus runais* | MF351518 | Hudisavirus sp. P3\_2 | - | Peru | *Homo sapiens* | P3\_2 |
| *Hudisavirus* | *Hudisavirus runais* | MF351516 | Hudisavirus sp. P23 | - | Peru | *Homo sapiens* | P23 |
| *Hudisavirus* | *Hudisavirus runais* | MF351515 | Hudisavirus sp. P22 | - | Peru | *Homo sapiens* | P22 |
| *Hudisavirus* | *Hudisavirus runais* | MF351513 | Hudisavirus sp. P18 | - | Peru | *Homo sapiens* | P18 |
| *Hudisavirus* | *Hudisavirus sewis* | MG522854 | Hudisavirus sp. ETH\_P7\_2016 | 2016 | Ethiopia | *Homo sapiens* | ETH\_P7\_2016 |
| *Hudisavirus* | *Hudisavirus sewis* | MG522856 | Hudisavirus sp. ETH\_P20\_2016 | 2016 | Ethiopia | *Homo sapiens* | ETH\_P20\_2016 |
| *Hudisavirus* | *Hudisavirus sewis* | KU043400 | Unidentified circular ssDNA virus cg\_8815 | 2014 | USA | *Macaca* | cg\_8815 |
| *Hudisavirus* | *Hudisavirus sewis* | LC708011 | Hudisavirus sp. CRESSV19-84-AMS-07 DNA | 1984 | Netherlands | *Homo sapiens* | CRESSV19-84-AMS-07 |
| *Hudisavirus* | *Hudisavirus sewis* | MW620877 | Porcine associated circular DNA virus 5 | 2014 | Brazil | *Sus scrofa domesticus* | P20\_41\_BR |

****

**Figure 1:** Maximum likelihood phylogenetic tree inferred from Rep proteins of members of the phylum *Cressdnaviricota*. Related sequence groups are collapsed into triangles, the side lengths of which are proportional to the distances between the closest and farthest leaf nodes. The alignment was trimmed with TrimAL [8] with gap threshold of 0.2. The maximum likelihood phylogenetic tree was constructed using IQtree [9] with automatic selection of the best-fit substitution model for a given alignment, which was Q.pfam+F+R10. Numbers at the nodes represent aLRT branch supports. The scale bar represents the number of substitutions per site.

****

**Figure 2:** Maximum likelihood phylogenetic tree of the Rep sequences of the members of the*Endolinaviridae* family inferred with PhyML 3.0 [10] with LG+I+G model determined as the best substitution model using ProtTest 3 [11] and rooted with representative sequences of members of the family *Circoviridae*. The species belonging to the same genus are indicated with the same color. Numbers at the nodes represent aLRT branch supports. The cyan line shows a proposed demarcation of genera. The genome organization relative to the *rep* ORF is shown to the right of the phylogeny.



**Figure 3:** A ‘two color’ pairwise identity matrix of members of the family *Endolinaviridae* with 78% species threshold *s* inferred using SDT v1.2 [12]. The grey box shows members of the species *Hudisavirus runais* with conflictsof 78% species demarcation which we resolved using the criteria outlined in the proposal coupled with phylogenetic support of the genome sequence-based maximum likelihood phylogeny (Figure 4).



**Figure 4:** Maximum likelihood phylogenetic tree of the genome sequences of the members of the*Endolinaviridae* family inferred with PhyML 3.0 [10] with GTR model determined as the best substitution model using jModelTest 3 [13]. The tree is midpoint rooted.

**References**

1. Kinsella CM, Deijs M, Becker C, Broekhuizen P, van Gool T, Bart A, Schaefer AS, van der Hoek L. (2022) Host prediction for disease-associated gastrointestinal cressdnaviruses. Virus Evol. 8(2):veac087. doi: 10.1093/ve/veac087. PMID: 36325032; PMCID: PMC9615429.
2. Altan E, Del Valle Mendoza J, Deng X, Phan TG, Sadeghi M, Delwart EL. (2017). Small circular Rep-encoding single-stranded DNA genomes in Peruvian diarrhea virome. Genome Announc. 5(38):e00822-17. doi: 10.1128/genomeA.00822-17. PMID: 28935725; PMCID: PMC5609404.
3. Altan E, Aiemjoy K, Phan TG, Deng X, Aragie S, Tadesse Z, Callahan KE, Keenan J, Delwart E. (2018) Enteric virome of Ethiopian children participating in a clean water intervention trial. PLoS One. 13(8):e0202054. doi: 10.1371/journal.pone.0202054. PMID: 30114205; PMCID: PMC6095524.
4. Krupovic M, Varsani A, 2022. Naryaviridae, Nenyaviridae, and Vilyaviridae: three new families of single-stranded DNA viruses in the phylum Cressdnaviricota. Arch. Virol. 167, 2907–2921.PMID: 36098801 DOI: 10.1007/s00705-022-05557-w
5. Varsani A, Krupovic M (2017) Sequence-based taxonomic framework for the classification of uncultured single-stranded DNA viruses of the family Genomoviridae. Virus Evol 3:vew037. doi:10.1093/ve/vew037. PMID:28458911
6. Varsani A, Krupovic M (2018) Smacoviridae: a new family of animal-associated single-stranded DNA viruses. Arch Virol 163:2005-2015. doi:10.1007/s00705-018-3820-z. PMID:29572596
7. 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(6):1411-24. doi: 10.1007/s00705-012-1601-7. Epub 2013 Jan 23. PMID: 23340592.
8. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972-3. doi:10.1093/bioinformatics/btp348. PMID:19505945
9. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol 37:1530-1534. doi:10.1093/molbev/msaa015. PMID:32011700
10. 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. PMID: 20525638; doi: 10.1093/sysbio/syq010
11. Darriba D, Taboada GL, Doallo R, Posada D. (2011).ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics. 27(8):1164-5. doi: 10.1093/bioinformatics/btr088. Epub 2011 Feb 17. PMID: 21335321; PMCID: PMC5215816.
12. Muhire BM, Varsani A, Martin DP (2014) SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. Plos One 9:e108277. PMID: 25259891; PMCID: PMC4178126.
13. Darriba D, Taboada GL, Doallo R, Posada D. (2012). jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 9(8):772. doi: 10.1038/nmeth.2109. PMID: 22847109; PMCID: PMC4594756.