

The International Committee on Taxonomy of Viruses

Taxonomy Proposal Form, 2024

**Part 1a: Details of taxonomy proposals**

|  |  |  |
| --- | --- | --- |
| **Title:** | Create a new order, Lineavirales, and a new family, the Oomyviridae, with 3 genera and 38 species in the class *Arfiviricetes* of the phylum *Cressdnaviricota* | |
| **Code assigned:** | 2024.004F.Uc.v2.Oomyviridae\_newfam |

|  |  |  |  |
| --- | --- | --- | --- |
| **Author(s), affiliation and email address(es):** | | | |
| **Name** | **Affiliation** | **Email address** | **Corresponding author(s)** X |
| Canuti M | Department of Veterinary and Animal Sciences, University of Copenhagen, Frederiksberg, Denmark | marta.canuti@gmail.com | X |
| Pénzes J | Institute for Quantitative Biomedicine,  Rutgers University, Piscataway, United States | Judycash08@gmail.com | X |

**Part 1b: Taxonomy Proposal Submission**

|  |  |  |  |
| --- | --- | --- | --- |
| **ICTV Subcommittee:** | | | |
| Animal DNA Viruses and Retroviruses |  | Bacterial viruses |  |
| Animal minus-strand and dsRNA viruses |  | Fungal and protist viruses | X |
| Animal positive-strand RNA viruses |  | Plant viruses |  |
| Archaeal viruses |  | General - |  |

|  |
| --- |
| **List the ICTV Study Group(s) that have seen or have been involved in creating this proposal:** |
|  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Optional – complete only if formally voted on by an ICTV Study Group:** | | | |
| **Study Group** | **Number of members** | | |
| **Votes in support** | **Votes against** | **No vote** |
|  |  |  |  |
|  |  |  |  |

|  |  |
| --- | --- |
| **Submission date:** | 09/06/2024 |

**Part 1c: Feedback from ICTV Executive Committee (EC) meeting**

|  |  |
| --- | --- |
| **Executive Committee Meeting Decision code:** | **X** |
| A – Accept |  |
| Ac – Accept subject to revision by relevant subcommittee chair. No further vote required |  |
| U – Accept without revision but with re-evaluation and email vote by the EC |  |
| Uc – Accept subject to revision and re-evaluation and email vote by the EC |  |
| Ud – Deferred to the next EC meeting, with an invitation to revise based on EC comments |  |
| J - Reject |  |
| W - Withdrawn |  |

|  |
| --- |
| **Comments from the Executive Committee:** |
| Please consider changing species epithets to fit current binomial format (no single numbers!). Please be aware that some of the proposed species are represented by partially sequenced isolates (please doublecheck if they are partial or coding-complete (highlighted yellow in the Excel file). Finally, the labelling of trees could be improved. |

**Part 1d: Revised Taxonomy Proposal Submission**

|  |
| --- |
| **Response of proposer:** |
| - Species epithets have been changed as requested and the documents have been updated accordingly  - We originally indicate a few strains as partial genome because they missed part of an hypothetical small ORF after the Cap gene; the genomes of all proposed species include the full coding region of the Rep and Cap genes. Since the expression of this hypothetical ORF has never been confirmed, we modified all PG in the Excel file to CCG. |

|  |  |
| --- | --- |
| **Revision date:** | 28/10/2024 |

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

|  |
| --- |
| **Name of accompanying Excel module:** |
| 2024.004F.Uc.v2.Oomyviridae\_newfam.xlsx |

|  |  |  |  |
| --- | --- | --- | --- |
| **Taxonomic changes proposed:** | | | |
| Establish new taxon | X | Split taxon |  |
| Abolish taxon |  | Merge taxon |  |
| Move taxon |  | Promote taxon |  |
| Rename taxon |  | Demote taxon |  |
| Move and rename |  |

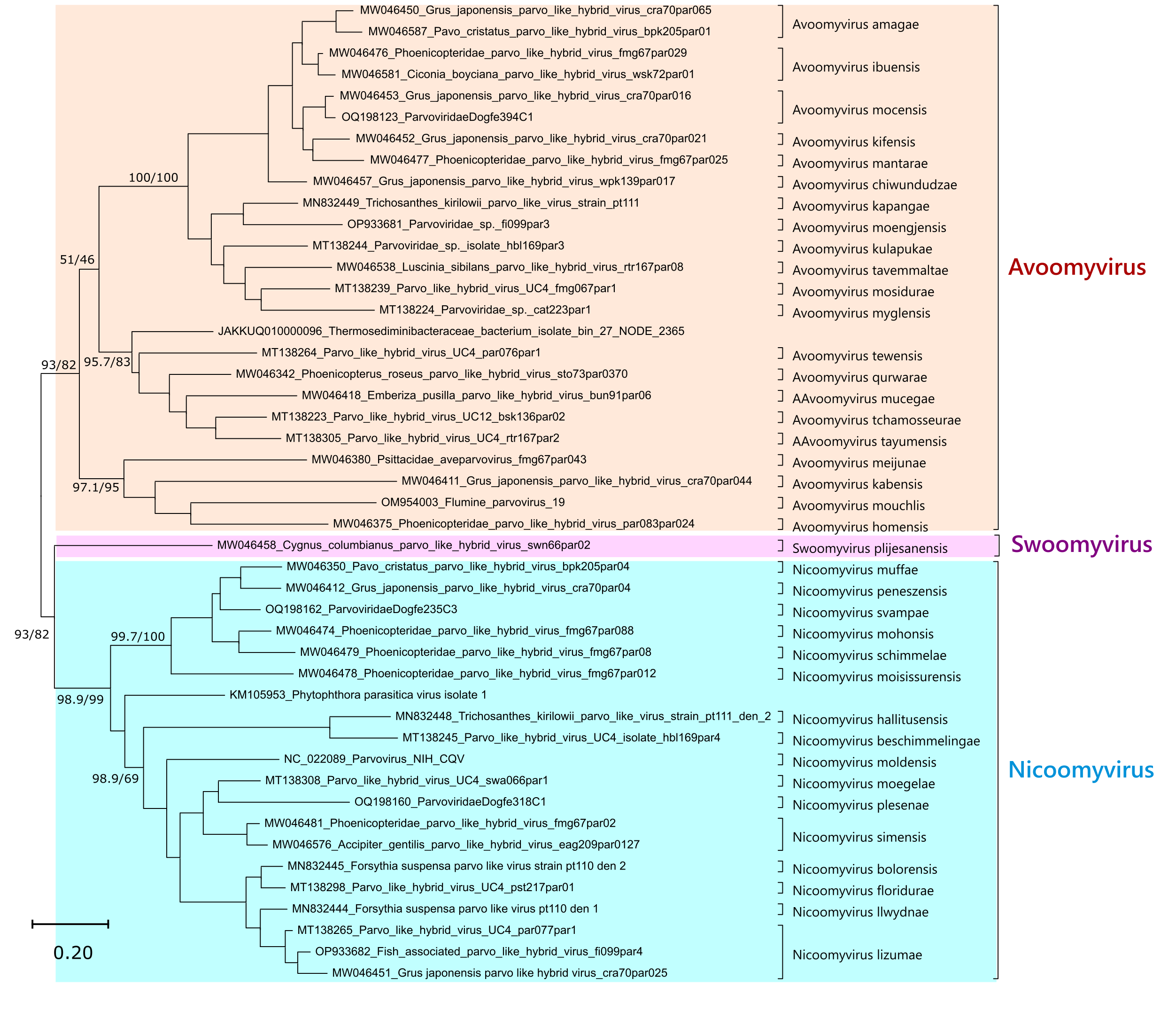
|  |  |  |
| --- | --- | --- |
| **Is any taxon name used here derived from that of a living person:** | | **Y/N** |
| **Taxon name** | **Person from whom the name is derived** | **Attached X** |
|  |  |  |
|  |  |  |
|  |  |  |

|  |
| --- |
| **Abstract of Taxonomy Proposal:** |
| *Taxonomic rank(s) affected*: Phylum *Cressdnaviricota* and class *Arfiviricetes*.  *Description of current taxonomy*: Currently unclassified.  *Proposed* *taxonomic change(s):* Create a new order, Lineavirales, and a new family, the Oomyviridae, with 3 genera (Nicoomyvirus, Avoomyvirus, and Swoomyvirus) and 38 species, in the class *Arfiviricetes* of the phylum *Cressdnaviricota*.  *Justification*: In 2013 a novel virus that was considered to be a “hybrid” between a parvovirus and a circovirus (“parvovirus-like hybrid virus) was discovered. With the increased use of metagenomics, several recent publications described similar viruses, proposing their classification as parvoviruses and erroneously labeling them in GenBank as parvoviruses. This misclassification issue is continuously increasing and is in dire need to be rectified. Here, we show that these viruses comprise a distinct linear ssDNA virus family (Oomyviridae) within the *Cressdnaviricota* and that their unique features and phylogenetic relationships with other members of the class *Arfiviricetes*, are strong reasons to include these viruses in a distinct order, for which we propose the name Lineavirales, owning to the linear genome organization these viruses were found to possess thus far. We also show that, although most of these viruses were identified in samples collected from animals, their likely hosts are organisms of the eukaryotic clade Stramenopiles (SAR supergroup). |

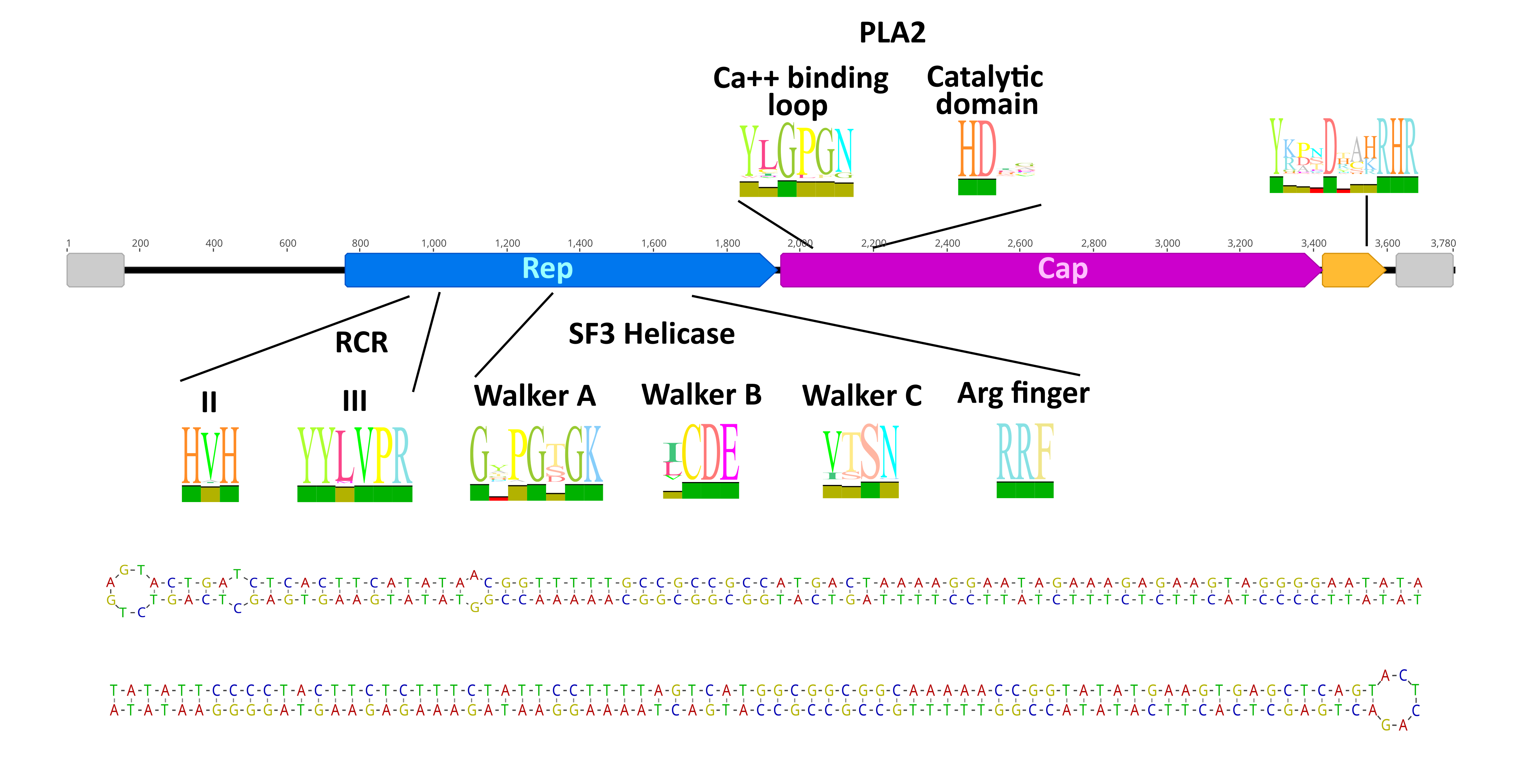
|  |
| --- |
| * **Text of Taxonomy proposal:** |
| *Taxonomic rank(s) affected*: Phylum *Cressdnaviricota* and class *Arfiviricetes*.  *Description of current taxonomy*: Currently unclassified.  *Proposed* *taxonomic change(s)*: Introducing a new ssDNA virus family, with 3 genera and 38 species (**Figure 1**), within the class *Arfiviricetes*, of the *Cressdnaviricota* phylum under the name **Oomyviridae**. The name derives from the most likely host for these viruses, which are members of the class Oomycetes (water molds). To accommodate this family, we would like to introduce a new order, **Lineavirales**, since the viruses within this family are, so far, the only linear members of the Cressdnaviricota.  Introducing the genus **Nicoomyvirus** with 16 species. The name derives from the first discovered virus in the family, strain NIH-CQV. Each species contains viruses fulfilling the demarcation criteria described below. All species names are derived from the word for “mold” in different languages.  - Nicoomyvirus moldensis, including Parvovirus NIH\_CQV, the prototypical and first described virus for the family, and related strains. These viruses were identified as contaminants in spin columns used for DNA isolations (1,2). The name derives from the word “mold”.  - Nicoomyvirus muffae, including Pavo cristatus parvo-like hybrid virus bpk205par04, found in a Chinese peafowl (3).The name derives from the Italian word “muffa”.  - Nicoomyvirus peneszensis, including Grus japonensis parvo-like hybrid virus cra70par04, found in a Chinese red-crowned crane (3) and ParvoviridaeDogfe368C2, found in a Chinese dog (4). The name derives from the Hungarian word “penesz”.  - Nicoomyvirus svampae, including Parvoviridae Dogfe235C3, found in a Chinese dog (4). The name derives from the Danish word “skimmelsvamp”.  - Nicoomyvirus mohonsis, including Phoenicopteridae parvo-like hybrid virus fmg67par088, found in Chinese flamingos (3). The name derives from the Spanish word “moho”.  - Nicoomyvirus schimmelae, including Phoenicopteridae parvo-like hybrid virus fmg67par08, found in Chinese flamingos (3). The name derives from the German word “Schimmel”.  - Nicoomyvirus moisissurensis, including Phoenicopteridae parvo-like hybrid virus isolate fmg67par012, found in Chinese flamingos (3). The name derives from the French word “moisissures”.  - Nicoomyvirus hallitusensis, including Trichosanthes kirilowii parvo-like virus strain pt111-den-2, found on Chinese cucumber plants (5). The name derives from the Estonian word “Hallitus”.  - Nicoomyvirus beschimmelingae, including Parvo-like hybrid virus UC4 hbl169par4 and cra070par1, identified in fecal samples from Chinese redstarts and cranes (6). The name derives from the Dutch word “beschimmeling”.  - Nicoomyvirus moegelae, including Parvo-like hybrid virus UC4 swa066par1, identified in fecal samples of a Chinese swan (6). The name derives from the Swedish word “mögel”.  - Nicoomyvirus plesenae, including ParvoviridaeDogfe318C1, found in a Chinese dog (4). The name derives from the Russian word “plesen”.  - Nicoomyvirus simensis, including Phoenicopteridae parvo-like hybrid virus fmg67par02, found in Chinese flamingos (3), and Accipiter gentilis parvo-like hybrid virus eag209par0127, identified in a Chinese goshawk (3). The name derives from the Esperanto word “simo”.  - Nicoomyvirus bolorensis, including Forsythia suspensa parvo-like virus strain pt110-den-2, found on Chinese forsythia plants (5). The name derives from the Portuguese word “bolor”.  - Nicoomyvirus floridurae, including Parvo-like hybrid virus UC4 isolate pst217par01, identified in a Chinese pheasant (6). The name derives from the Catalan word “floridures”.  - Nicoomyvirus llwydnae, including Forsythia suspensa parvo-like virus strain pt110-den-1, found on Chinese forsythia plants (5). The name derives from the Welsh word “llwydni”.  - Nicoomyvirus lizumae, including Parvo-like hybrid virus UC4 isolate par077par1, identified in a Chinese parrot (6), Fish-associated parvo-like hybrid virus fi099par4, found in intestinal content of Chinese fish (7), and the Grus japonensis parvo-like hybrid virus cra70par025, identified in a Chinese crane (3). The name derives from the Basc word “lizum”.  Introducing the genus **Avoomyvirus** with 21 species. The name derives from “avian”, since most members identified so far were detected in samples from birds. Each species contains viruses fulfilling the demarcation criteria described below. All species names are derived from the word for “mold” in different languages.  - Avoomyvirus amagae, including the Grus japonensis parvo-like hybrid virus cra70par065, identified in Chinese cranes (3), and the Pavo cristatus parvo-like hybrid virus bpk205par01, identified in a Chinese peafowl (3). The name derives from the Tagalog word “amag”.  - Avoomyvirus ibuensis, including the Phoenicopteridae parvo-like hybrid virus fmg67par029, identified in Chinese flamingos (3), and the Ciconia boyciana parvo-like hybrid virus wsk72par01, identified in a Chinese stork (3). The name derives from the Yoruba word “ibu”.  - Avoomyvirus mocensis, including the Grus japonensis parvo-like hybrid virus cra70par0169, identified in Chinese cranes (3), and the ParvoviridaeDogfe394C1, found in a Chinese dog (4). The name derives from the Vietnamese word “moc”.  - Avoomyvirus kifensis, including the Grus japonensis parvo-like hybrid virus cra70par02, identified in Chinese cranes (3). The name derives from the Azerbaijani word “kif”.  - Avoomyvirus mantarae, including the Phoenicopteridae parvo-like hybrid virus fmg67par025, identified in Chinese flamingos (3). The name derives from the Turkish word “kuf mantari”.  - Avoomyvirus chiwundudzae, including the Grus japonensis parvo-like hybrid virus wpk139par017, identified in Chinese cranes (3). The name derives from the Shona word “chiwundudzi”.  - Avoomyvirus kapangae, including the Trichosanthes kirilowii parvo-like virus pt111-phy-9-plant, found on Chinese cucumber plants (5). The name derives from the Indonesian word “kapang”.  - Avoomyvirus moengjensis, including the Parvoviridae sp. isolate fi099par3found in the intestinal content of Chinese fish (7). The name derives from the Vahcuengh word “moengj”.  - Avoomyvirus kulapukae, including Parvoviridae sp. hbl169par3, identified in the intestinal content of Chinese birds (6). The name derives from the Malay word “kulapuk”.  - Avoomyvirus tavemmaltae, including the Luscinia sibilans parvo-like hybrid virus rtr167par08, identified in a Chinese robin (3). The name derives from the Kabyle word “taɣemmalt”.  - Avoomyvirus mosidurae, including the Parvo-like hybrid virus UC4 fmg067par1, identified in Chinese flamingos (6). The name derives from the Occitan word “mosidura”.  - Avoomyvirus myglensis, including the Parvoviridae sp. cat223par1, identified in the intestinal content of Chinese birds (6). The name derives from the Icelandic word “mygla”.  - Avoomyvirus tewensis, including the Parvo-like hybrid virus UC4 par076par1, identified in a Chinese parakeet (6). The name derives from the Haitian creole word “tewo”.  - Avoomyvirus qurwarae, including the Phoenicopterus roseus parvo-like hybrid virus sto73par0370, identified in a Chinese flamingo (3). The name derives from the Quechuan word “qurwara”.  - Avoomyvirus mucegae, including the Emberiza pusilla parvo-like hybrid virus bun91par06, identified in a Chinese bunting (3). The name derives from the Romanian word “mucegai”.  - Avoomyvirus tchamosseurae, including the Parvo-like hybrid virus UC12 bsk136par02, identified in a Chinese shrike (6). The name derives from the Walloon word “tchamosseure”.  - Avoomyvirus tayumensis, including the Parvo-like hybrid virus UC4 rtr167par2, identified in a Chinese robin (6). The name derives from the Javanese word “tayum”.  - Avoomyvirus meijunae, including the Psittacidae aveparvovirus fmg67par043, identified in Chinese flamingos (3). The name derives from the Chinese word “méi jūn”.  - Avoomyvirus kabensis, including the Grus japonensis parvo-like hybrid virus cra70par044, identified in Chinese cranes (3). The name derives from the Japanese word “kabi”.  - Avoomyvirus mouchlis, including the Flumine parvovirus 19, identified in environmental samples in New Zealand (8). The name derives from the Greek word “mouchla”.  - Avoomyvirus homensis, including the Phoenicopteridae parvo-like hybrid virus par083par024, identified in Chinese flamingos (3). The name derives from the Finnish word “home”.  Introducing the genus **Swoomyvirus** with 1 species. The name derives from “swan”, since the only member in this genus was detected in a swan. This strain is equidistant from members of the other two genera and could not be definitively assigned to either of the 2. Therefore, a separate third genus was created to accommodate this virus.  - Swoomyvirus plijesanensis, including the Cygnus columbianus parvo-like hybrid virus swn66par02, identified in Chinese swans (3). The name derives from the Croatian word for “mold”: “plijesan”.  *Demarcation criteria:*  **Virus definition**:  For a sequence to be classified in the proposed family Oomyviridae, it must be in one piece and contain the intact complete coding regions of the nonstructural protein (Rep), which must possess RCR endonuclease and SF3 helicase domains in its protein sequence, and of the structural protein (Cap). Considering difficulties in obtaining reliable terminal sequences, the 5’ and 3’ non-coding regions are not strictly necessary for classification. All sequences must be reported in a credible peer-reviewed publication. This definition is designed to allow the inclusion of viruses identified by virus discovery approaches while possibly avoiding endogenous viral elements.  **Demarcation criteria and nomenclature**:  Species: two oomyviruses can be potentially classified in one species if their Rep share at least 90% protein sequence identity. Species must be designated under a binomial name, consisting of the genus name, within which the given virus is classified, and a Latinized name.  Genus: two oomyviruses can be potentially classified in one genus if they cluster as a robust monophyletic lineage based on their complete Rep protein sequence in a family-wide phylogeny and their Rep proteins share 50% identity. Flexibility in these numbers may apply.  *Justification*:  In 2013 two independent research groups in the USA and China described the discovery through metagenomic methods of a novel virus that was considered to be a “hybrid” between a parvovirus and a circovirus, and which was named “parvovirus-like hybrid virus” (1,2). While originally linked to cases of human hepatitis (1), the virus was later found to be a laboratory contaminant originating from silica columns used for DNA isolations (2). The “hybrid” nature of the virus was defined because phylogenetic analyses performed at that time, when knowledge about viral diversity in the *Cressdnaviricota* was extremely limited, showed how this virus was located in an intermediate position between viruses from the *Parvoviridae* and viruses from the *Circoviridae.* However, the genome structure was similar to parvoviruses, possessing a ssDNA genome with two main ORFs coding for a non-structural protein with HUH and SF3 helicase domains and a structural protein with a PLA2 domain, flanked by almost identical inverted terminal repeats at both 5’ and 3’ sides, predicted to be capable of folding into hairpin structures(1,2) (**Figure 2**).  Throughout the years, with the widespread utilization of metagenomics, several related viruses have been discovered (3,4,6–8). Likely because of the originally given name, many of the studies describing sequences similar to the original “parvovirus-like hybrid virus” tended to propose a classification of these viruses as parvoviruses or include them within the *Parvoviridae* when describing the diversity of viruses within their investigations. These viruses, however, have never been officially classified within the family *Parvoviridae*. As many of these sequences are labelled in GenBank as belonging to the *Parvoviridae,* with some of them even erroneously labeled as specific existing parvovirus species, the problem of misclassification is growing, owning to a vicious cycle of misclassification that gets amplified by wrongly labeled viruses uploaded in public repositories. To rectify this and clarify the continuous misclassification events, we propose to give an official name to these viruses and create a new viral family to accommodate them.  Here, we will show that i), these viruses comprise a distinct ssDNA virus family that is united by strong monophyletic support, genome organization, as well as potential host spectrum, ii) they form a lineage that is not closely related to the *Parvoviridae*, iii) the most suitable megataxonomic position to classify the novel family is within a distinct and new order within the *Cressdnaviricota*, demonstrating that the linear nature of these genomes evolved independently from parvoviruses, and iv) the likely host of these viruses are the eukaryotic clade Stramenopiles (SAR supergroup). Further on in this proposal, we will refer to this putative family as oomyviruses instead of the previously dubbed “parvo-like hybrid virus”.     1. **A distinct ssDNA virus family**   Oomyviruses comprise three lineages according to their Rep protein sequences, which are all homologous and highly conserved (aa pairwise identity of 65-95% with a >95% overlap) (**Figure 1**). Although most viruses in this proposed family lack a completely determined genome sequence (the terminal non-coding regions are mostly lacking), fully sequenced viruses are approximately 4 Kb in size and are characterized by the presence of three open reading frames (ORFs) coding for:  - A double-domain Rep protein with a typical endonuclease, including the HUH domain (RCR motif II),  a tyrosine-rich RCR III motif (YYLVPR), and the 4 CRESS DNA virus-like helicase subdomains, i.e., Walker A (GxxGxGK), Walker B (xCDE), Walker C ((V/I)(T/S)SN), and Arginine finger (AxxRRF).These Reps are short, at an approximate length of 400 residues or smaller.  - A Cap protein, which often contains a phospholipase A2 (PLA2) domain including a Ca++ binding loop (YxGPGN) and a catalytic subdomain (HDxx). The presence of the PLA2 is not ubiquitous throughout the proposed family.  - A small hypothetical ORF coding for a protein with an unknown function with a conserved YxxxDxxxRHR domain.  Fully sequenced viruses include almost identical inverted terminal repeats at the two extremities that can fold into hairpins (**Figure 2**).   1. **Host spectrum**   Although members of the proposed family have been mostly derived from metagenomes of multiple eukaryotic realms, i.e. plants (5), animals (3,4,6,7), and water molds (GenBank accession number: KM105953), there is no evidence to date confirming the replication of these viruses in the first two. However, multiple endogenous viral elements (EVEs), derived from both the Rep and structural proteins of these viruses are present in the genomes of water molds (Oomycetes) of the SAR division. For example, EVEs are found in chromosome 5 of *Phytophthora pluivora* (accession number: CP125260), several genome projects derived from *Phytophthora parasitica* (KI669648) and P. *fragariae* (QXFX00000000), *Aphanomyces astaci* (QUTE01010000), and from *Pythium insidiosum* (BBXB02000000),or in a transcript of *Plasmopara halstedii* (XM\_0247260951), *Phytophthora parasitica (XM\_008916763), Aphanomyces astaci* (XM\_009840055), and *Aphanomyces invadans* (XM\_008879548) (**Figure 3**). No oomyvirus EVEs can be found in the genome of any other eukaryotic realm. However, we did find CRESS-like transcripts and EVEs in alveolates, which harbored Reps more similar at aa level to the oomyviruses than to other CRESS DNA viruses, despite of being derived from exogenous relatives with a circular genome. Examples include *Gregarina niphandrodes* (XM\_011133028), and the same genome projects containing also the oomyvirus-like EVE (e.g., QXFX00000000). The genome of *Schmidingerella arcuata* also includes one of these closely related EVEs (JAMFLK010002335).   1. **The position of the proposed Oomyviridae from a megataxonomy perspective**   The proposed Oomyviridae family clusters within the *Cressdnaviricota* in an SF3 helicase-based phylogenetic tree, distinct from the similarly linear *Parvoviridae* (*Cossaviricota*) (**Figure 4**). Although the lack of conservation in the helicases across the large number of distinct ssDNA virus families results in a limited region with suitability for phylogenetic reconstructions, the relationships within the *Cressdnaviricota* can be resolved by the incorporation of the complete Rep protein sequence throughout the phylum (**Figure 5**). In this phylogenetic inference, it becomes evident that the Oomyviridae clusters within the class *Arfiviricetes* as a separate lineage, prompting us to classify this proposed family into its own proposed order, Lineavirales. Interestingly, the closest relatives to the Lineavirales are members of the *Baphyvirales*, an order of viruses that infect diatoms, close relatives of the Oomycetes in the Stramenopiles.  Besides their phylogenetic position, structural predictions and overall Rep homology also support this placement of the newly proposed family (**Figure 6**). If a BlastP search is executed, searching the Rep of any Oomyviridae member against the *Cressdnaviricota*, reliable coverage of 60% to 90% at an identity of 30% to 50% can be achieved with any member of the phylum, encompassing the highly conserved endonuclease and helicase domains. Parvovirus NS proteins (the parvoviral homologue of the CRESS DNA viral Rep) only result in hits against the *Cressdnaviricota* if the search is executed by PSIblast and, in this case, coverage no higher than 30% with an identity lower than 20% can be identified. Furthermore, parvovirus NS1 proteins are large, complex, multidomain proteins, which typically exceed 600 residues in length. This is radically different from the typically short Rep of CRESS DNA viruses that possess a double-domain architecture and length similar to those of the oomyvirus Reps.  Structural homology modeling is also in concordance with these findings. Fold recognition and template search identifies the Rep protein of porcine circovirus 2 (9) (PDB ID: 7LAS, p-value of 4e-04 by pGenThreader) to be structurally most similar to the oomyviral Rep, yet the homologous region has been structurally characterized in parvoviruses as well. A homology model, executed by Alphafold2 (pLDDT=87.1), when superimposed on the helicase domain of circoviral and parvoviral Reps, respectively, indicates that the oomyvirus Rep possesses the short linker region between the two N-terminal α-helices and the five-stranded β-sheet, which form the core of the domain, just like in the PCV2 Rep. In contrast, the parvoviral NS helicase domain displays an insertion here, consisting of three α-helices (**Figure 6**). It is notable, however, that the homology modeling of an additional C-terminal domain of the oomyvirus helicase suggests a similar structure for this subdomain as the parvoviral N-terminally located triple helix insertion. This may be in concordance with the complex helicase task required to package a linear ssDNA virus genome into pre-assembled capsids. This further implies that the linear organization of parvovirus and oomyvirus genomes evolved independently, albeit also possibly requiring the same replicase protein functions, besides the characteristic terminal hairpins.  Homology modeling of the capsid proteins could not be executed with sufficiently reliable results. Despite the lack of detectable sequence-based homology, the VPs of parvoviruses thus far all possess a characteristic eight-stranded jelly roll fold, which is distinguishable from similarly folded RNA- and DNA virus structural proteins by their elongated surface loops, linking the structurally conserved β-strands together (10). This fold could be modeled even in case of distant parvoviral lineages (11). The lack of success in homology modeling suggests that oomyviruses do not harbor parvovirus-like capsids, despite the majority possessing a PLA2 domain. This further supports our hypothesis that the family Oomyviridae is not closely related to the *Parvoviridae*, despite harboring a similarly linear genome. Penzes et al., has recently shown (12) that such domains were possibly acquired independently by various ancestral parvoviral lineages, which may be true for the oomyviruses as well. This is supported by the non-ubiquitous presence of such domains throughout the proposed family. Consequently, we base our proposed classification on the Rep.  Linearization of viral genomes from circular ancestors has also occurred multiple times in various viral lineages, such as the *Adenoviridae* (13)and the *Myoviridae* (14) families in the *Varidnaviria* and *Caudoviricetes*, respectively. Consequently, we do not consider this attribute to be a synapomorphy between oomyviruses and the *Parvoviridae*, but rather the result of convergent evolution. |

|  |
| --- |
| **References:** |
| 1. Xu B, Zhi N, Hu G, Wan Z, Zheng X, Liu X, et al. Hybrid DNA virus in Chinese patients with seronegative hepatitis discovered by deep sequencing. Proc Natl Acad Sci U S A. 2013 Jun 18;110(25):10264–9.  2. Naccache SN, Greninger AL, Lee D, Coffey LL, Phan T, Rein-Weston A, et al. The perils of pathogen discovery: origin of a novel parvovirus-like hybrid genome traced to nucleic acid extraction spin columns. J Virol. 2013 Nov;87(22):11966–77.  3. Dai Z, Wang H, Wu H, Zhang Q, Ji L, Wang X, et al. Parvovirus dark matter in the cloaca of wild birds. GigaScience. 2023 Jan 1;12:giad001.  4. Jiang X, Liu J, Xi Y, Zhang Q, Wang Y, Zhao M, et al. Virome of high-altitude canine digestive tract and genetic characterization of novel viruses potentially threatening human health. mSphere. 2023 Sep 19;8(5):e00345-23.  5. Yang S, Mao Q, Wang Y, He J, Yang J, Chen X, et al. Expanding known viral diversity in plants: virome of 161 species alongside an ancient canal. Environ Microbiome. 2022 Nov 27;17(1):58.  6. Shan T, Yang S, Wang H, Wang H, Zhang J, Gong G, et al. Virome in the cloaca of wild and breeding birds revealed a diversity of significant viruses. Microbiome. 2022 Apr 12;10(1):60.  7. Xi Y, Jiang X, Xie X, Zhao M, Zhang H, Qin K, et al. Viromics Reveals the High Diversity of Viruses from Fishes of the Tibet Highland. Microbiol Spectr. 11(3):e00946-23.  8. French R, Charon J, Lay CL, Muller C, Holmes EC. Human land use impacts viral diversity and abundance in a New Zealand river. Virus Evol. 2022;8(1):veac032.  9. Tarasova E, Dhindwal S, Popp M, Hussain S, Khayat R. Mechanism of DNA Interaction and Translocation by the Replicase of a Circular Rep-Encoding Single-Stranded DNA Virus. mBio. 2021 Aug 31;12(4):e0076321.  10. Mietzsch M, Pénzes JJ, Agbandje-McKenna M. Twenty-Five Years of Structural Parvovirology. Viruses. 2019 Apr 20;11(4):362.  11. Pénzes JJ, de Souza WM, Agbandje-McKenna M, Gifford RJ. An Ancient Lineage of Highly Divergent Parvoviruses Infects both Vertebrate and Invertebrate Hosts. Viruses. 2019 Jun 6;11(6):525.  12. Pénzes JJ, Pham HT, Chipman P, Smith EW, McKenna R, Tijssen P. Bipartite genome and structural organization of the parvovirus Acheta domesticus segmented densovirus. Nat Commun. 2023 Jun 14;14(1):3515.  13. Davison AJ, Benkő M, Harrach B. Genetic content and evolution of adenoviruses. J Gen Virol. 2003 Nov;84(Pt 11):2895–908.  14. Hao Y, Wang S, Zhang M, Tang Q, Meng C, Wang L, et al. Isolation and characterization of a novel linear-plasmid phage from the sediment of the Mariana Trench. Virol Sin. 2022 Apr;37(2):311–3.  15. Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020. Nucleic Acids Res. 2021 Jan 8;49(D1):D458–60. |

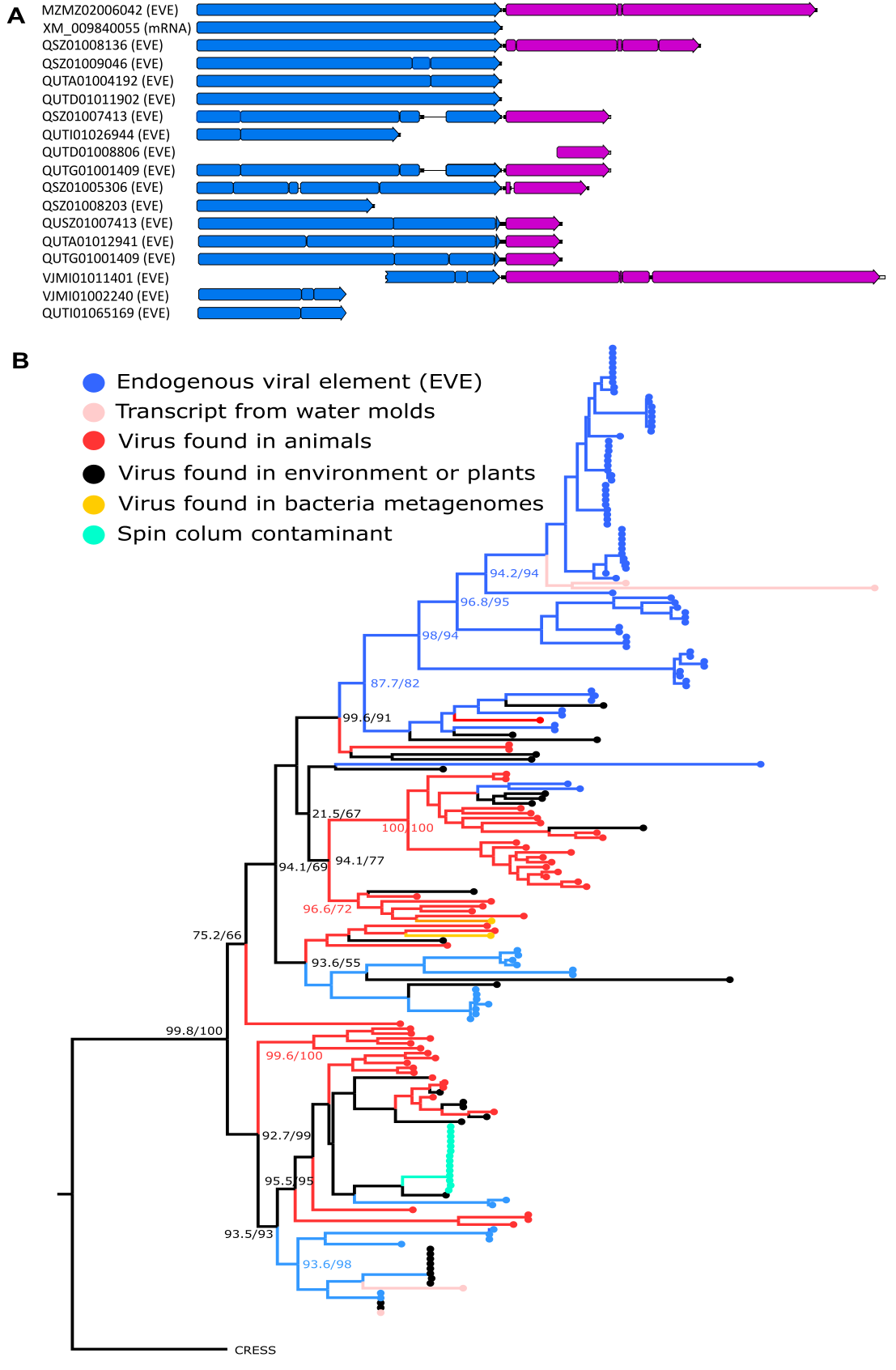
|  |
| --- |
| **Tables, Figures:** |



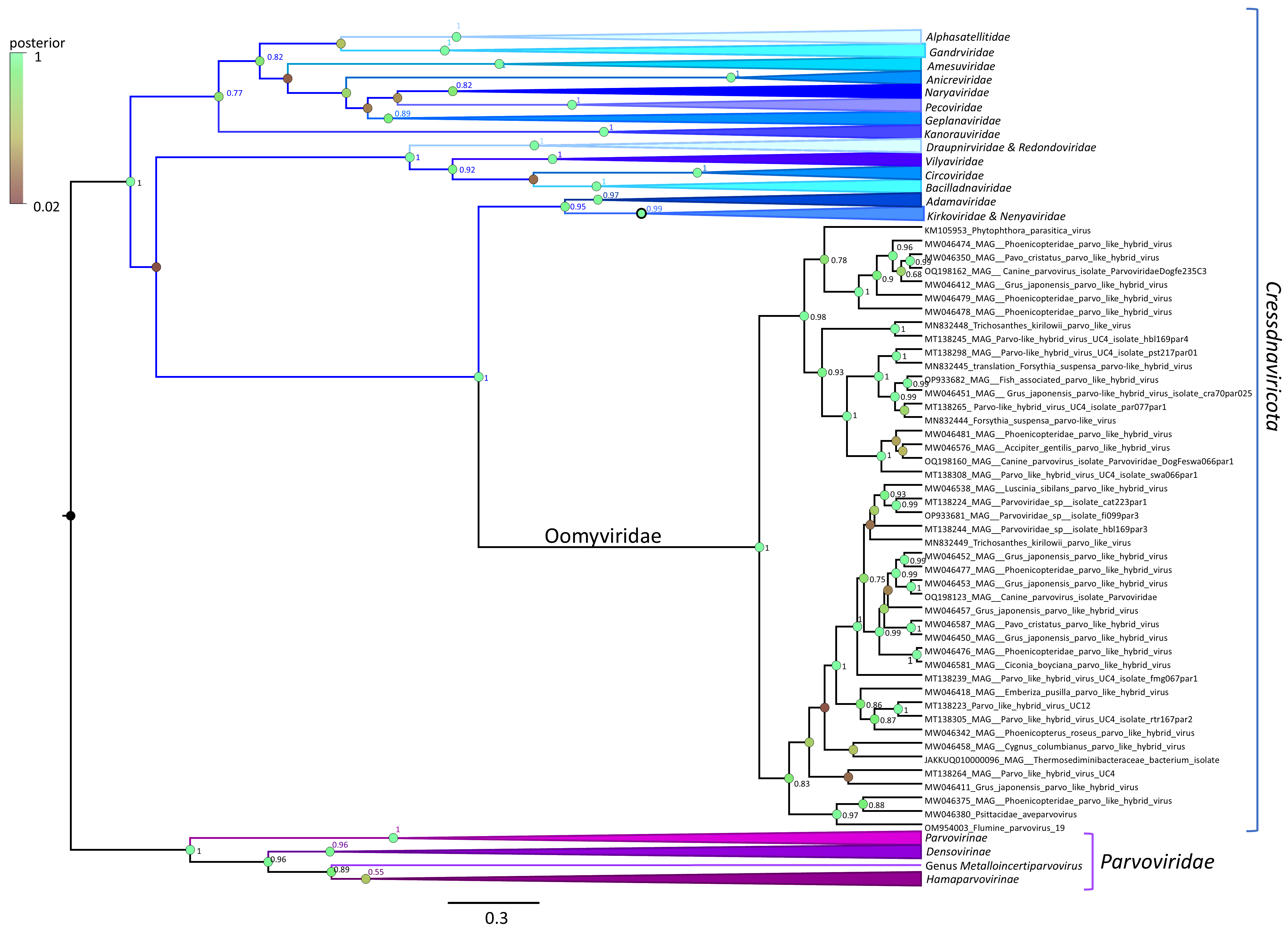
**Figure 1**.Maximum likelihood phylogenetic inference of the proposed family Oomyviridae, based on the full Rep protein-derived amino acid sequences and rooted at midpoint. The reliability of the tree topology is indicated by ultrafast bootstrap and SH-aLRT values, shown as node labels. The calculations were carried out by IQTREE2 under the LG+I+G4 substitution model. Each virus name is indicated after the GenBank accession number and proposed species names are indicated on the right. Proposed genera are labelled by colored shading and genera names are shown on the right.



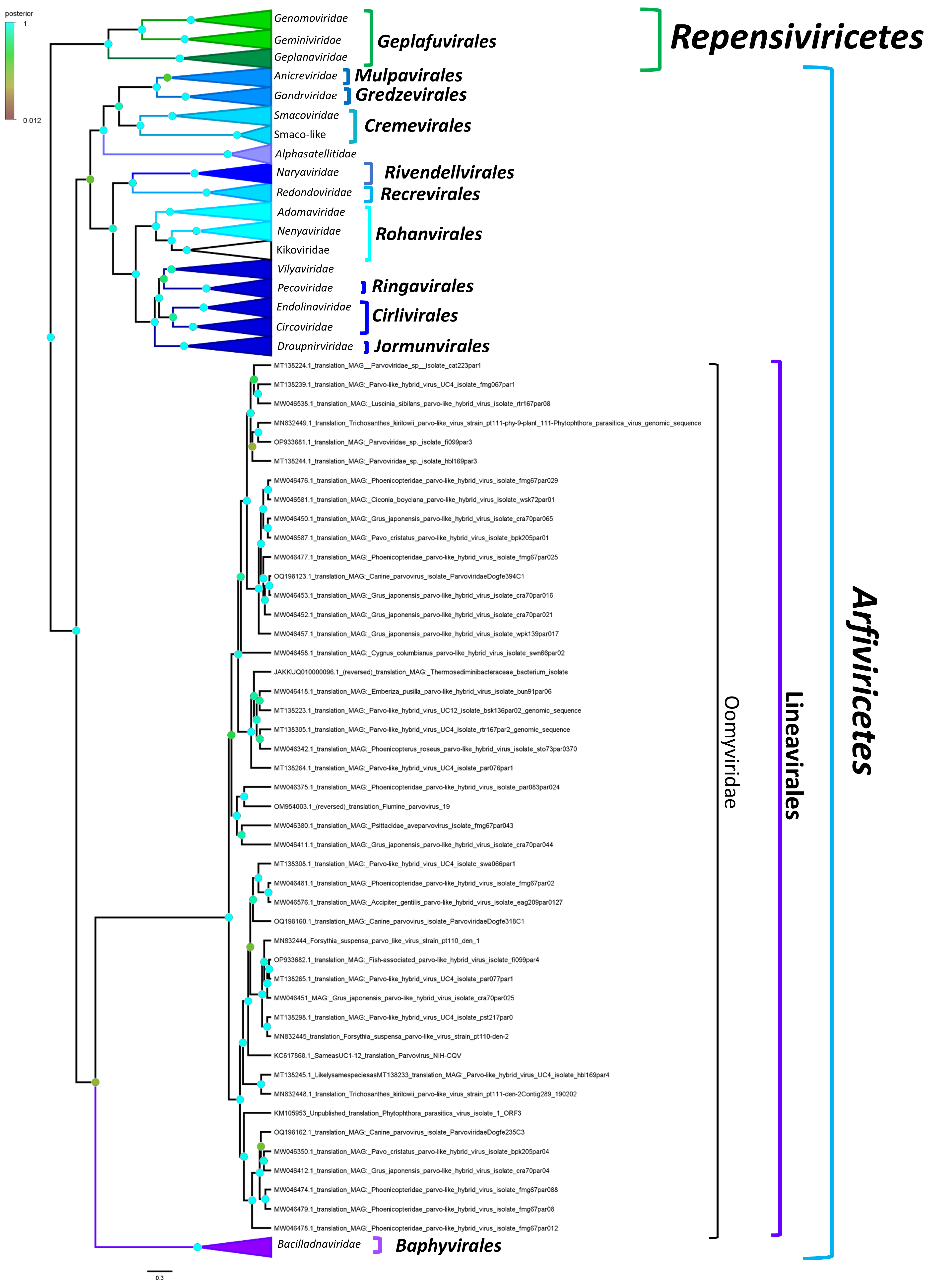
**Figure 2**.Genome organization of oomyviruses, built using the sequence of parvovirus NIH\_CQV (accession number: NC\_022089), the prototypical strain of the family. The genome is represented by a black line on which different domains are indicated by colored boxes. The two grey boxes indicate the region including the inverted terminal repeats (ITR) which can fold into hairpis (depicted at the bottom of the figure). The blue arrow represents the Rep protein, that includes rolling circle replication (RCR) domains II and III and a superfamily 3 (SF3) helicase domains (Walker A, B, and C, and the arginine finger), typical of CRESS viruses. The pink arrow depicts the structural protein Cap, that can include a phospholipase A2 (PLA2) domain, including a calcium binding loop (Ca++) and the catalytic domain. The yellow box indicates a protein of unknown function with an additional conserved domain. Domain conservation is illustrated by sequence logos.



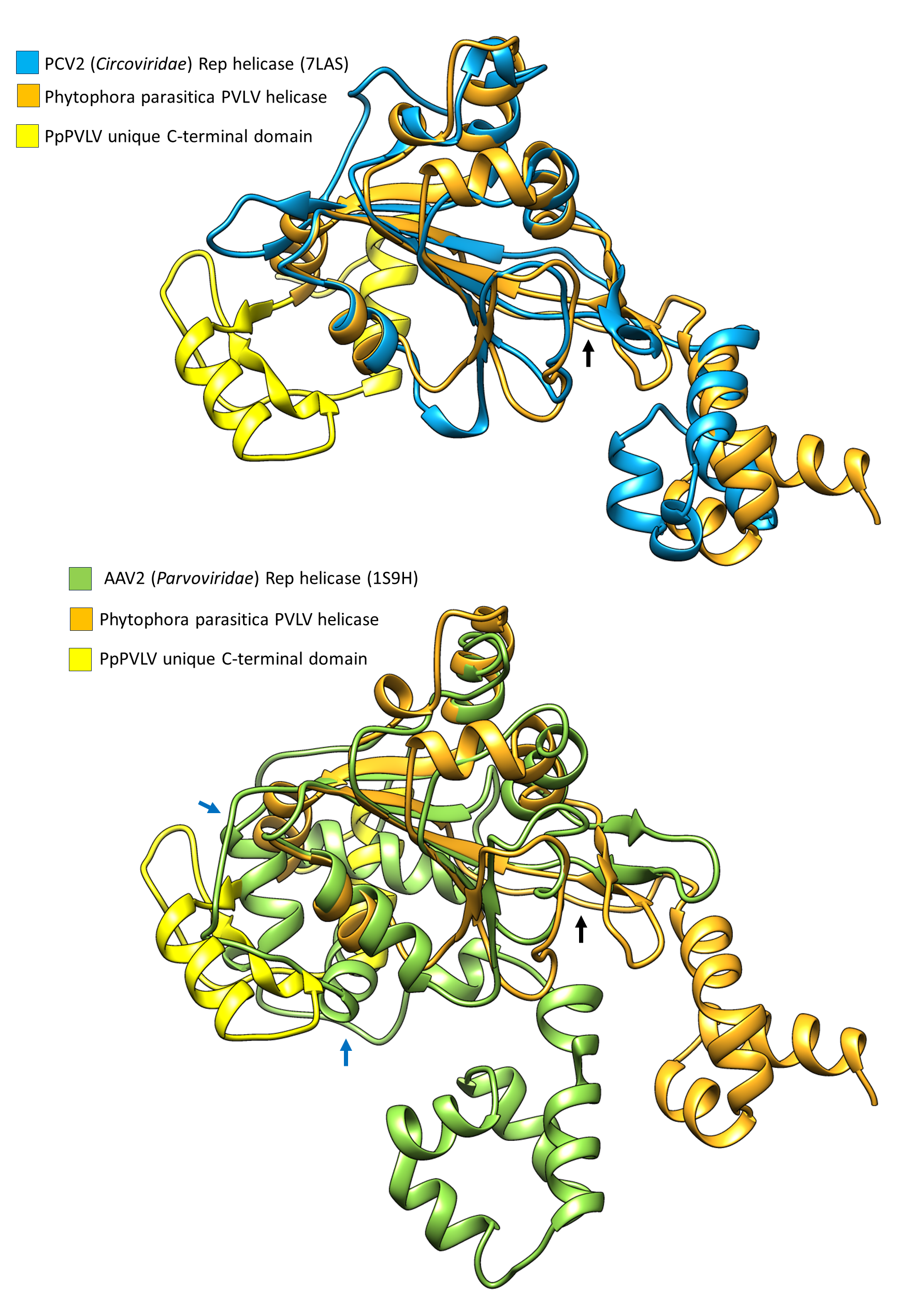
**Figure 3**. Oomyvirus-like endogenous viral elements (EVE). The figure in panel A depicts a few examples of EVEs identified in sequenced genomes of *Aphanomyces astaci*. On the left, the accession number of the genomic sequence that contains the element is noted together with the indication of whether it derived from a full genome sequencing (EVE) or a transcriptomic (mRNA) project. The blue and pink boxes represent Rep- and Cap-derived coding sequence showing homology to oomyvirus proteins. Panel B shows the phylogenetic relationships among partial Rep sequences from viruses identified in different environments and oomyvis-like EVEs identified in Oomycetes as indicated by the legend on top-left. The maximum likelihood phylogenetic inference was inferred by IQTREE2 under the LG+R5 substitution model and the reliability of its topology is indicated by ultrafast bootstrap and SH-aLRT values, shown at main nodes.



**Figure 4.**Bayesian inference of the isolated SF3 helicase domain (160 aa) protein sequence throughout the entire *Shotokuvirae* kingdom. Representatives, which lacked a detectable SF3 pham domain by the SMART (Simple Modular Architecture Research Tool) search (15) [were omitted from the analysis. Although the SF3 helicase is not sufficient to resolve inter family relationships within the](https://doi.org/10.1093/nar/gkaa937) phylum *Cressdnaviricota*, the proposed Oomyviridae still clusters in this phylum as opposed to the similarly linear *Parvoviridae* (phylum *Cossaviricota*). The alignment was obtained by an initial structure-based algorithm of t-coffee expresso, which was then aligned with more representatives of all families with the regressive algorithm of t-coffee, using a batch sub-alignement size of 1000 sequences. The phylogenetic inference was carried out by BEAST v1.4.10, using an uncorrelated lognormal relaxed clock, a Yule speciation model and an LG+G4 substitution model. The simulation was run for 50 million generations.



**Figure 5.** Bayesian inference of the complete Rep protein sequences of the phylum *Cressdnaviricota*, including the proposed family Oomyviridae. Oomyviridae clusters as a distinct lineage within the class *Arfiviricetes*, in its own proposed order, Lineavirales. The inference was carried out by BEAST v1.4.10, using an uncorrelated lognormal relaxed clock, a Yule speciation model and an LG+I+G4 substitution model. The simulation was run for 50 million generations.



**Figure 6**. Homology model of the Phytophora parasitica parvo-like virus (PpPVLV), superimposed on the porcine circovirus 2 (PCV2) (*Circoviridae*, *Cressdnaviricota*) Rep helicase domain (top) as well as the corresponding domain of the adeno-associated virus 2 (AAV2) (*Parvoviridae*, *Cossaviricota*) (bottom). The black arrow indicates the short linker, missing the helical insertion present N-terminally of the Walker A motif in parvoviral SF3 helicases (blue arrows). Note the similar secondary- and tertiary structural morphology of this region to that of the unique C-terminal helical domain of the PpPVLV homology model, shown in bright yellow.