

The International Committee on Taxonomy of Viruses

Taxonomy Proposal Form, 2025

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

|  |  |
| --- | --- |
| **Title:** | Reclassification of *Flaviviridae* (order *Amarillovirales*) and “flavi-like” viruses into three families, 12 genera and 3 subgenera |
| **Code assigned:** | 2025.006S.N.v1.Amarillovirales\_3reorgfam | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Author(s), affiliation and email address(es):** | | | | |
| **Given name (+middle initial(s))** | **Surname** | **Affiliation** | **Email address** | **Corr. author(s)** |
| Peter | Simmonds | Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom | Peter.Simmonds@ndm.ox.ac.uk | X |
| Anamarija | Butkovic | Viruses and RNA Interference Laboratory, Paris, France | anamarija.butkovic@pasteur.fr |  |
| Joe | Grove | MRC-University of Glasgow Centre for Virus Research, Glasgow, UK | Joe.Grove@glasgow.ac.uk |  |
| Richard | Mayne | Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom | richard.mayne@ndm.ox.ac.uk |  |
| Jon CO | Mifsud | Sydney Institute for Infectious Diseases, The University of Sydney, Sydney, Australia | jmif9945@uni.sydney.edu.au |  |
| Martin | Beer | Institute of Diagnostic Virology, Friedrich-Loeffler-Institut, Greifswald-Insel Riems, Germany | Martin.Beer@fli.de |  |
| Jens | Bukh | Copenhagen Hepatitis C Program(CO-HEP), Copenhagen University Hospital, Hvidovre, Denmark | jbukh@sund.ku.dk |  |
| Jan Felix | Drexler | Charité-Universitätsmedizin Berlin, Institute of Virology, Berlin, Germany | felix.drexler@charite.de |  |
| Amit | Kapoor | Center for Vaccines and Immunity, The Research Institute at Nationwide Children's Hospital, Columbus, Ohio, USA | amit.kapoor@nationwidechildrens.org |  |
| Volker | Lohmann | Department of Infectious Diseases, Molecular Virology, Heidelberg University, Heidelberg, Germany | Volker.Lohmann@med.uni-heidelberg.de |  |
| Donald B | Smith | Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom | donald.smith.ictv@gmail.com |  |
| Jack T | Stapleton | Departments of Internal Medicine, Microbiology and Immunology, University of Iowa and Iowa City VA Healthcare, Iowa City, Iowa, USA | jack-stapleton@uiowa.edu |  |
| Nikolaos | Vasilakis | Department of Pathology and Center for Vector-Borne and Zoonotic Diseases, University of Texas Medical Branch, Galveston, Texas, USA | nivasila@utmb.edu |  |
| Jens H | Kuhn | Frederick, Maryland, USA | jenshkuhn@comcast.net |  |

**Part 1b: Taxonomy Proposal Submission**

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

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

|  |  |  |  |
| --- | --- | --- | --- |
| **Optional – complete only if formally voted on by an ICTV Study Group:** | | | |
| **Study Group** | **Number of members** | | |
| **Votes in support** | **Votes against** | **No vote** |
| ICTV *Flaviviridae* Study Group | 9 | 0 |  |
|  |  |  |  |

|  |  |
| --- | --- |
| **Submission date:** | 26/06/2025 |

**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:** |
|  |

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

|  |
| --- |
| **Response of proposer:** |
|  |

|  |  |
| --- | --- |
| **Revision date:** |  |

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

<https://ictv.global/taxonomy/templates>

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

|  |  |
| --- | --- |
| **Etymology (origin) of proposed taxonomic names:** | |
| **Taxon name** | **Etymology of the term** |
| *Euflavivirus* | “True” flaviviruses, members of the original genus |
| *Crangovirus* | After a caridean shrimp species name, *Crango crango* (Linnaeus, 1758) |
| *Orthoflavivirus alphei* | Gen. sing. from Ancient Greek Ἀλφειός (Alpheiós), snapping shrimp |
| *Fusivirus* | After fusion |
| *Orthoflavivirus iunctionis* | Gen. sing. Latin of iunctio (Latin for junction) |
| *Tamanavirus* | After Tamana bat virus |
| *Tamanavirus parnelli* | Gen. Sing. Latin – from bat species name *Pteronotus* (*Phyllodia*) *parnellii* Gray, 1843 named after Richard Parnell |
| *Termitovirus* | After termite host |
| *Termitovirus isopterae* | Nom. Sing. Latin of *Isoptera*,an infraorder for termites |
| *Guaicovirus* | After Guaico culex virus |
| *Guaicovirus culicis* | Gen. Sing. Latin of *Culex* , a genus for mosquitoes |
| *Jingmenvirus* | After Jīngmén tick virus |
| *Jingmenvirus rhipicephali* | Gen. Sing. Latin of *Rhipicephalus*,a genus for ticks |
| *Arachnivirus* | derived from the Greek word ἀράχνη (arachni) |
| *Arachnivirus neosconae* | Gen. Sing. Latin of spider species name *Neoscona nautica* (L. Koch, 1875) |
| *Boletivirus* | From Bólè tick virus 4 |
| *Boletivirus hyalommae* | Gen. Sing. Latin of *Hyalomma*, agenus for ticks |
| *Chrysopivirus* | After Chrysopidae, a family for green lacewings |
| *Chrysopivirus vittae* | Gen. Sing. Latin of vitta (Latin for lace) |
| *Koshovirus* | After gentian Kobu-sho-associated virus |
| *Koshovirus sonchi* | Gen. Sing. Latin of *Sonchus, a genus for* sow thistles |

|  |  |  |
| --- | --- | --- |
| **Permission for use of names derived from a living person:** | | |
| **Taxon name** | **Full name of person from whom the name is derived** | **Attached** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

|  |
| --- |
| **Abstract of Taxonomy Proposal:** |
| *Taxonomic rank(s) affected*: *Flaviviridae*, *Amarillovirales*  *Description of current taxonomy*: *Flaviviridae* is a family for non-segmented positive-sense enveloped RNA viruses many of which are significant pathogens, including hepatitis C virus and yellow fever virus. *Flaviviridae* isthe sole family included in order *Amarillovirales* and is subdivided into four genera: *Orthoflavivirus*, including 52 species into which arthropod-borne and insect-specific flavivirids are classified; *Pestivirus* (19 species), *Hepacivirus* (14 species); and *Pegivirus* (11 species).  *Proposed* *taxonomic change(s):* Recent large-scale metagenomic surveys have identified many diverse RNA viruses related to classical orthoflaviviruses and pestiviruses but possessing quite different genome lengths and configurations. They have a hugely expanded host range that spans multiple animal phyla (including mollusks, cnidarians and stramenopiles), and plants.  Phylogenetic analysis of RNA-directed RNA polymerase (RdRP) hallmark gene sequences splits flavivirid and ‘flavi-like’ viruses into four divergent clades and multiple lineages within them. The tree is congruent with helicase gene phylogeny, PPHMM profile comparisons, and RdRP protein structure predicted relationships predicted by AlphFold2. These results support their classification into the established order, *Amarillovirales* as three separate families (*Flaviviridae, Pestiviridae*,and *Hepaciviridae*), and a total of at least 12 genera.  *Justification*: Although the current classified members of the *Flaviviridae* and flavi-like viruses form a monophyletic group separate from other RNA viruses, they are far more divergent from each other than other RNA virus genera, supporting their assignment as three separate families and several genera within them. Taxonomic assignments based on RdRP hallmark gene evolutionary relationships provides a stable reference for assignment of further members of this order, and a framework from which major genome re-organisational events can be understood. |
| **Text of Taxonomy proposal:** |
| *Taxonomic rank(s) affected*: *Flaviviridae*, *Amarillovirales*  *Description of current taxonomy*: *Flaviviridae* is a family of positive-sense RNA viruses and the only family included in order *Amarillovirales.* The family includes four genera: *Orthoflavivirus*, including 52 species into which arthropod-borne and insect-specific flavivirids are classified, *Pestivirus* (19 species) for viruses of ruminants and a wider range of mammals, *Hepacivirus* (14 species), and *Pegivirus* (11 species). Flavivirid genome organization is shared across the four genera, with a single 8000–10,700 base open reading frame that is translated and cleaved to generate core and envelope proteins, and a shared complement of protease, helicase and RNA-directed RNA polymerase (RdRP) non-structural proteins. Viruses of different flavivirid genera differ, however, in their mechanisms of translation initiation and virion formation.  Collectively, the *Flaviviridae* incorporates several major human and veterinary pathogens, including hepatitis C virus (HCV) and a wide range of often highly virulent arthropod-borne viruses, including yellow fever virus (YFV) and dengue viruses [2]. The *Flavivirus* and *Pestivirus* genera were included in the family, *Flaviviridae* [3, 4], a group that was subsequently expanded to incorporate a third genus, *Hepacivirus*, for HCV [5] and relatives infecting non-human primates, bats, and rodents (reviewed in [6]). Finally, a fourth genus, *Pegivirus*,was added in 2012 for a range of apparently non-pathogenic RNA viruses infecting humans (human pegivirus [HPgV]) and a broad range of non-human primates, bats, and other mammals [7] and one avian (goose) host [8]. Since then, there have been only minor changes to flavivirid classification, essentially limited to the expansion of the number of species assigned to the *Pestivirus, Pegivirus* and *Hepacivirus* generaand the renaming of the genus *Flavivirus* to *Orthoflavivirus* [9].  Shared features of current members of the *Flaviviridae* include a common genome organization (positive-sense, non-segmented, linear, non-3'-polyadenylated RNA 9.0–13 kb in length), protein expression strategies (synthesis of a single polyprotein with a conserved organization that is cleaved into structural proteins located at the N terminus and nonstructural proteins at the C terminus), and a primarily mammalian host range (in case of orthoflaviviruses, also arthropods) [2]. In addition to the RdRP gene, flavivirids are homologous in their superfamily 2 helicase (NS3) and serine protease domain sequences [10].  However, among flavivirids, there are also some major differences, including possession of structurally distinct capsid proteins in viruses of different genera (and the apparent lack of a capsid in pegiviruses), packaging mechanisms, polyprotein translation strategy that may be 5'-cap-dependent (orthoflaviviruses) or driven through an internal ribosomal entry site (IRES; other flavivirids), and two disparate fusion glycoprotein systems. (Fig. 1). Indeed, other than the serine protease, helicase and RdRP, no other nonstructural protein genes are homologous across viruses of all four genera.  High throughput sequencing technologies have detected and genetically characterized viruses related to flaviviridsin a much wider range of potential hosts, particularly of arthropods and more recently fish and other aquatic life. These include insect-only (non-vectored) flavivirids [11-16] and the expansion of species assignments to the *Pestivirus*, *Pegivirus*, and *Hepacivirus* genera for viruses that infect primates, other mammals, and birds (reviewed in [6, 17-20]). The taxonomic assignment of a much wider range of ‘flavi-like’ viruses [21] is more problematic. Their RdRP gene sequences group phylogenetically with those of classified members of *Flaviviridae*, but they often possess quite different genome organisations, genome lengths, and host ranges. These include the ‘large-genome flaviviruses (LGFs); [22-24], with genomes substantially longer than currently classified flavivirids, the longest to date being maximus pesti-like virus’s genome of ≈40 kb [25]. Jīngmén tick virus (JMTV) [26] and related jingmenviruses have multisegmented genomes with four or more separate segments [27-30]. In addition, the majority of ‘flavi-like’ viruses have been discovered outside the primarily mammalian and vector host range of classified flavivirids, being distributed across the animal kingdom, from poriferans (sponges) [25], cnidarians (jellies) [31], mollusks (squid) [32], arthropods (insects [22, 29, 33]; diplurans [29]; scorpions [29]; crustaceans [31, 32, 34]), nematodes [35], platyhelminths [36, 37]) to echinoderms (sea cucumbers) [38], hemichordates (acorn worms) [31], cartilaginous and bony fish [23, 31, 39-42], amphibians (frogs) [31, 43], and reptiles [41]. Although only a few are known so far, ‘flavi-like’ viruses have also been discovered in stramenopiles (diatoms and oomycotes) [44, 45] and in angiosperm plants [1, 28, 46].  *Proposed* *taxonomic change(s):* Our proposed re-classification of flavivirids is based upon a submitted paper [47] that applies several analysis methods to reinvestigate their genetic relatedness to the wide range of currently described flavi-like viruses. For classification purposes, we have assigned primacy to the RdRP (hallmark) gene phylogeny using previously established principles for a genomics-based taxonomy of viruses [48] that state that assignments should be based on the most evolutionarily conserved gene within virus groups [49-51]. As members of the realm *Riboviria*, the RdRP is used as the hallmark gene for delineating RNA virus evolutionary histories. These define taxonomic assignments irrespective of relationships between other genome regions. Many flavi-like viruses differ substantially in genome organisation from classified flavivirids, including genome segmentation and instances of gene loss and acquisition. Structural genes encoding virion components are particularly variable and may be driven by major changes in host range or ecologies. Examples of the diversity of genome organisations of flavi-like viruses is shown in Fig. 1  To depict phylogeny relationships between classified flavivirids and flavi-like viruses, we accessed an alignment of their RdRP genes and a comprehensive sample of coding-complete sequences of currently unclassified ‘flavi-like’ viruses from a previous study [52] (listed in Table S1; Suppl. Data of [47], that includes a detailed methodological description of analysis methods) . This dataset included representatives from groups of segmented ‘flavi-like’ viruses, divergent ‘pesti-like’ viruses, many with extended genomes, many additional ‘hepaci-’ and ‘pegi-like’ viruses primarily found in marine vertebrates, plant-infecting ‘koshoviruses’ and a selection of ‘flavi-like’ viruses recovered from environmental samples such as diatom colony-associated virus (DCAV).  Using the distantly related tombusvirids as an outgroup, the RdRP domain phylogeny of the sequence set, calculated with IQ-TREE, resulted in four main bootstrap-supported clades (I–IV), each containing a number of bootstrap-supported lineages (labelled as Ia–l, IIm–s, IIIt–w anticlockwise around the tree) (Fig. 2). Viruses of the four currently established flavivirid genera were distributed in three of the four clades, with hepaciviruses and pegiviruses clustering together in and largely defining clade III. In this phylogeny, ‘jingmenviruses’, ‘tamanaviruses’, and numerous ‘insect-specific flaviviruses (ISFs)’ cluster with orthoflaviviruses in clade I. The viruses that were previously loosely defined as ‘LGFs’ join pestiviruses in clade II, and include diverse viruses distributed across multiple lineages, e.g., Bólè tick virus (a potentially tick-vectored pathogen of mammals) [24] in clade IIq and plant-infecting ‘koshoviruses’ [1, 28, 46] in clade IIs (Fig. 2; Table S1 in [47]).  To investigate the robustness of the clade and lineage groupings, the IQ-TREE-based maximum-likelihood phylogeny generated for Fig. 2 was compared with trees generated through a temporal reconstruction using the Bayesian evolutionary analysis by sampling trees cross-platform program (BEAST) and using the protein distance-based unweighted pair group method with arithmetic mean (UPGMA) phylogeny method (Fig. 3). BEAST and UPGMA results reproduced clades I–IV with bootstrap support comparable to the original maximum likelihood analysis (Figs. 2, 3A). However, in the time-rooted BEAST tree, *Tombusviridae* became an inlier (Fig. 3C), in marked contrast to its clearly defined outgroup positions in the IQ-TREE and UPGMA trees. The grouping of individual sequences within lineages a–w defined by IQ-TREE analysis were almost entirely reproduced in the BEAST and UPGMA trees, with the exception of eight sequences that did not group with their lineages in the UPGMA tree (Fig. 3B). Despite these, the clade and lineage assignments were relatively robust to different evolutionary reconstruction methods.  These RdRP phylogenies provide an excellent framework for the taxonomic reorganization of the *Flaviviridae*. Nonetheless, we recognized that phylogenetic inference over these very large genetic distances is challenging and therefore sought to corroborate (or, indeed, refute) these apparent taxonomic groupings with alternative and complementary approaches.  *Helicase domain phylogeny.* While there is evidence for modular exchange of structural or assessor proteins among flavivirids of the four established current genera, for instance acquisition of glycoprotein-encoding genes [52, 53], it is less clear whether sequences encoding the core replication module of these viruses (including serine protease, helicase and RdRP) evolve as a unit or are similarly subject to genome exchange and rearrangements. To investigate this, and to complement our RdRP analyses, we deduced flavivirid and ‘flavi-like’ helicase domain amino acid sequences, aligned them, performed IQ-TREE analysis, and compared the result to the RdRP IQ-TREE tree using a tanglegram (Fig. 4). Phylogenetic groupings of viruses of the four current genera were highly concordant, with only minor differences in branching order within genera. The positions of ‘flavi-like’ viruses grouping with orthoflaviviruses were generally concordant between regions, but with some exceptions. For instance, there was a change in the topology of the deeper branches underlying the ‘LGF’, pestivirus and hepaci/pegivirus groupings, creating paraphyletic groups not observed in the RdRP tree. This result is not necessarily surprising because the helicase domain is shorter and more divergent than the RdRP domain and hence likely reflects lower resolution of relationships rather than indicating genome reorganization between the two domains. Other minor exceptions included Wēnlǐng moray eel hepacivirus (WMEHV) moving from an outlier position in clade III (lineage IIIt) into *Hepacivirus* in the helicase tree, a finding not incompatible with potential sequence errors in the deposited RdRP region sequence. Bólè tick virus 4 (BoTV4) fell within the IIq lineage in the RdRP domain tree, but as an outlier to pestiviruses in the helicase domain tree. Finally, ‘flavi-like’ viruses from environmental samples (clade IV in the RdRP tree) were located within the pestivirus-‘LGF’ branches in the helicase domain tree.  *Protein structure relationships.* The function of any given protein is primarily a feature of its three-dimensional structure. Consequently, protein structure is fundamentally more conserved than the underlying protein sequence. The advent of accurate protein structure prediction through machine-learning (e.g., AlphaFold2) is enabling surveys of protein form and function at enormous scales [54, 55]; and has driven the development of new high-throughput structure comparisons tools (e.g., Foldseek) that enable structure-guided inference of deep evolutionary relationships [56, 57].  In a recent investigation of glycoproteins, protein structure prediction was systematically applied to flavivirids [52], generating thousands of structures spanning the complete polyproteins of all viruses  represented in the RdRP phylogeny (Fig. 2; Table S1, Suppl. Data in [47]). Drawing on this dataset, we generated complete NS5/NS5b RdRP domain structure predictions for each virus (see Methods). After filtering for prediction confidence and length, we analysed 400 RdRP structures using FoldTree [57] to produce a structure-guided tree based on the local distance difference test (lddt) structural similarity metric. This structure-based approach does not explicitly consider protein sequence and, therefore, represents an independent recapitulation of the sequence-based phylogeny (Fig. 2).  The topology of the structure-based tree was remarkably similar to that of the sequence-based tree and supports the validity of the same four major clades (Fig. 5A). Moreover, most of the lineages represented in the sequence phylogeny were consistent in their position and composition. The only exceptions are lineages Ia, which in the structure-based tree formed a basal branch from lineage Ic, and lineages Ii and IIIv which moved subtly in relation to their neighboring clades. Clades Ij, IIn, and IIIt were lost from the structure-based tree due to filtering of lower confidence structural predictions. Major clade IV, containing ‘flavi-like’ viruses from environmental samples, also shifted position slightly, branching between clade II and the *Tombusviridae* outgroup in the structure-based tree. However, this divergent and basal taxon will likely remain difficult to place by any method without further discovery of similar viruses.  Example structural alignments of clade representatives corroborate their distribution on the tree (Fig. 5B). Dengue virus type 1 and HCV genotype 1 were at opposite ends of the tree (clade I and III) and their respective RdRPs align with a relatively low lddt score (0.47), whereas pairings within clade I, II and III give higher scores: 0.67, 0.61 and 0.74 respectively (note that an lddt score of 1.0 represents perfect alignment of identical structures). Thus, inferring evolutionary relatedness through structure-only analysis corroborates sequence-based approaches and is highly supportive of the organization of flavivirids and ‘flavi-like’ virusesinto four main clades.  *Analysis by GRAViTy version 2.* Genome Relationships Applied to Virus Taxonomy (GRAViTy) is a non-supervised, alignment free method to assess the relatedness of virus genome sequences though calculation of protein profile hidden Markov model (PPHMM) homologies and through metrics of genome organization such as the order and orientation of genes [58]. We performed GRAViTy analysis using our flavivirid and ‘flavi-like’ virus dataset (Table S1) for phylogeny and RdRP structure comparisons; genomes of segmented ‘jingmenviruses’were concatenated in order of segment length from short-to-long. Results were remarkably concordant with those determined by RdRP and helicase domain phylogenies (Figs. 2–4), with bootstrap-supported segregation of the same sequences into three main clades I–III (Fig. 6A). ‘Flavi-like’ viruses from environmental samples (formed an outlier position as clade IV.  To depict the wider inter-relationships of classified flavivirids and ‘flavi-like’ viruses, we compared their sequences with sequences representing viruses of all other established ribovirian families. Flavivirids and ‘flavi-like’ viruses were monophyletic in the GRAViTy dendrogram, supporting their assignment to a common higher taxonomic rank, the established order *Amarillovirales* (Fig. 6B). The Jaccard distances calculated by GRAViTy do not provide a precise quantitative estimate of evolutionary distance or thresholds for taxonomic assignments. However, ribovirian families differ from each other in the distance range of 0.7–0.85, whereas viruses within families typically are associated with Jaccard distances of approximately 0.9 (Fig. 6B). Although very general, the distances between clades I–IV were within the range of between-family distances elsewhere in the dendrogram, whereas their combined grouping occurs at the rank of order assignments for other virus families assigned to orders (e.g., *Nidovirales* and *Picornavirales*)*.*  *Association between taxonomic groupings and host range.* Within clade I, lineage Ib encompasses the currently classified members of the *Orthoflavivirus* genus, as well as a high number of currently unclassified primarily insect-specific flaviviruses (ISFs). Lineage Ic is similarly populated by ISFs, including the current unclassified cell fusing agent virus (CFAV), consistent with a previous proposed assignment to a new genus within *Flaviviridae* [59]. The similarly divergent Tamana bat virus (TABV) falls in the very diverse lineage If that also contains ‘flavi-like’ viruses infecting lumpfish (*Cyclopterus lumpus* Linnaeus, 1758) and southern pygmy squid (*Xipholeptos notoides* ([Berry](https://en.wikipedia.org/wiki/Samuel_Stillman_Berry), 1921)). The segmented JMTV and Guaico Culex virus (GCuV) and their relatives were assigned to lineages Ik and Il, adjacent to lineage Ij encompassing the non-segmented infectious precocity virus. Extending the host range of this clade were Cnidaria flavivirus and Harrimaniidae flavivirus infecting basal metazoa such as jellies and acorn worms [31] in lineage Ia. Despite the evident lineage (and host) diversity, clade I is clearly monophyletic with a bootstrap-supported long branch separating members from other clades.  Clade II similarly sub-divides into a number of bootstrap-supported lineages IIm–s, with currently classified pestiviruses clustering exclusively in lineage IIo along with currently unassigned more divergent viruses exclusively infecting vertebrates, as does lineage IIn (fish). Viruses assigned to other lineages within clade II are arthropod-hosted with the striking exception of the plant-infecting Apis flavivirus, carrot flavi-like virus 1, Gentian Kobu-sho-associated virus, Coptis virus 1, and Sonchus virus 1 that form a sub-group, termed ‘koshoviruses’, within lineage IIs [1, 46, 60]. Arthropod and plant-infecting members of clade II frequently have substantially longer genomes than the vertebrate-infecting members of lineages IIn and IIo (ranges 11,038–11,450 and 11,555–15,154 respectively) compared with lineage IIm (20,432–22,622), IIq (13,599–18,696), IIr (14,731–26,314), and IIs (18,749–27,708).  Members of clade III are all non-segmented, possessing similar genome lengths (8294–12,290) and form a well-defined separate grouping from other flavivirids. In contrast to the wide host range of clade I and II, all clade III members have presumed or demonstrated vertebrate hosts, spanning a wide range of mammals, birds, reptiles, and bony and cartilaginous fish. Two lineages, IIIu and IIIw, encompass currently classified members of the *Pegivirus* and *Hepacivirus* genera along with a range of more divergent viruses with a greater host range beyond mammals and birds.  *Proposals for assignments.* The reproducible phylogenetic relationships of flavivirids and ‘flavi-like’ viruses using different tree construction methods in the RdRP domain sequences (Fig. 2 & 3) were recapitulated by protein structure-guided analysis (Fig. 5). RdRP is evidently co-evolving with the helicase gene (Fig. 4), suggesting that the essential replicase of flavivirids and ‘flavi-like’ viruses traces a single coherent evolutionary history. This is also reflected in the relationships identified by an alignment-free method for analyzing complete genome sequences (Fig. 6). Therefore, we are confident that the foundational RdRP phylogeny in Figure 2 provides a robust framework for a genomics-based re-classification of flavivirids.  These analyses based on clade and lineage relationships provide a framework to map clades and lineages onto families and genera. However, there are quite variable branch lengths within lineages (Fig. 2) and differences in thresholds that split lineages in BEAST and UPGMA tree across clades I, II and III (Fig. 3). These observations indicate that a purely cladistic classification may not conform to specific sequence divergence thresholds that are often used elsewhere in virus taxonomy. Indeed, formal comparison of mean amino acid sequence identities between and within lineages in the three clades vary considerably (Fig. 7), particularly in clade I where sequence identities between lineages Ia, Ib, Ic, and Id were all much greater than within-group values of other lineages (notably If, Ih, Ii and Ij, and lineage n in clade II).  With this caveat, we propose the following re-classification and expansion of the *Flaviviridae* through incorporation of the large number of additional ‘flavi-like’ viruses (Table 1). The following proposals for a radically reorganised and expanded flavivirid taxonomy are presented below.  **Proposed taxonomic assignments and demarcation criteria (Table 2):**  **Order** Flavivirid and ‘flavi-like’ viruses form a monophyletic group in comparison with all other members of the *Riboviria*, enabling their assignment to a single taxonomic rank. Although not precise, divergence among flaviviruses at this rank was comparable to that of members of virus orders in GRAViTy analysis (Fig. 6B). We therefore propose that all flavivirid and ‘flavi-like’ viruses can be assigned to the established order *Amarillovirales*.  **Family** The level of sequence divergence among members of clades I–IV was comparable to inter-family distances of other RNA viruses on GRAViTy analysis. There is no formal divergence threshold by which families may be defined in the *Riboviria* or in other virus realms. However, a maximum intra-family Jaccard distance score of between 0.7-0.75 was observed in the majority of RNA virus families on analysis of (coding-) complete genome sequences by GRAViTy v2, including the three proposed families in the current proposal (Fig. 6B). This threshold was consistently below the level between different orders of RNA viruses. We therefore propose assignment of members of lineages I, II, and III to individual families. We maintain the name *Flaviviridae* for one of them, and propose ‘*Pestiviridae*’ and ‘*Hepaciviridae*’, respectively, for the other two, reflecting the historical key virus members of these families*.*  While a separate, bootstrap-supported lineage IV was consistently observed, its members are highly divergent genetically and derive from environmental samples. Where hosts are suspected, these are extremely diverse and require further verification. We therefore do not propose to create a family for lineage IV at this time.  **Genus** All three suggested families contain a number of bootstrap-supported lineages of sequences, with members often possessing distinct genome organisations, lengths and host tropisms. We propose that the lineage assignments form the basis for genus demarcation.  *Current genera.* Clades Ib, IIo, IIIu, and IIIw contain currently classified flavivirids and we propose that the assignments of members of these clades to *Orthoflavivirus*, *Pegivirus*, *Pestivirus* and *Hepacivirus* genera are retained. However, we propose that the genera *Pestivirus* and *Hepacivirus* genera are renamed as ‘*Orthopestivirus*’and ‘*Orthohepacivirus*’ (to avoid confusion of vernacular names with those of the newly assigned *Pestiviridae* and *Hepaciviridae* familynames). The proposed genus names maintain a degree of continuity with established nomenclature.  *New genera.* We propose assignment of further genera within each of the three proposed families in the basis of forming bootstrap-supported groupings consistent between multiple analysis methods, and a minimum of three members so that the extent of grouping can be assessed. In proposing new genera, we give priority to clades corresponding to established genera and to those containing previously described and well-characterised viruses, such as cell fusing agent virus (CFAV) and Tamana bat virus (TABV).  To guide assignments, however, assignments, there is however no specific level of genetic divergence of flavivirids that would define genus membership. Indeed, amino acid sequence identities of RdRP gene amino acid sequences between and within lineages in the three clades showed considerable variation (Fig. 7). A threshold of 0.3 amino acid sequence identity would create several genera within each proposed family but identities between lineages Ia, Ib, Ic, and Id were greater and inconsistent with within-group values of other lineages, as were lineages IIIu, IIIv and IIIw. Contrastingly If, Ih and Ii were a little lower than the threshold. A demarcation line in the UPGMA and BEAST trees (blue dotted line in Figs. 3B and 3C) would, however, support provisional genus assignments across all three clades / families, and in separate GRAViTy 2 analysis (Fig. 6A).  On this basis, we propose the establishment and naming of a total of eight additional genera and three subgenera in addition to the four genera currently assigned (Table 2). These would include four further genera within the *Flaviviridae* family, *Tamanavirus, Termitovirus, Guaicovirus* and *Jingmenvirus*, and four additional genera in the *Pestiviridae* family, *Arachnivirus, Boletivirus, Chrysopivirus* and *Koshovirus* in addition to the established, renamed genus *Orthopestivirus*.  However, based on divergence and grouping below the blue threshold line, we propose that lineages Ia- Id should be combined into single genus with a proposed name *Orthoflavivirus*, with individual groupings within these of Ib, Ic and Id assigned as subgenera *Euflavivirus, Fusivirus*, and *Crangovirus.* These subgenus assignments possess biological relevance in terms of host range and vector transmission. As a result, all established *Orthoflavivirus* species will remain but the genus will be expanded by viruses not considered “classical” orthoflaviviruses before.  **Species** Assignments of species in established and newly proposed genera await the future development of genome-based assignment thresholds that can be more widely applied across the three families in the *Amarillovirales.* However, exemplar species for all newly proposed genera and subgenera have been proposed, typically corresponding the first described example of each group possessing a coding complete genome sequence. Assignments of established species have been retained.  A diagrammatic summary of the proposed taxonomic changes is provided in Fig. 8.  *Justification*: This taxonomy proposal makes a robust case for an evolutionarily-based reclassification of flavivirids using an approach that puts primacy on genetic relationships of the hallmark RdRP gene. The radical differences in genome organization, such as segmentation, changes to translation mechanisms and exchanges of structural gene modules of most ‘flavi-like’ viruses had historically prevented their incorporation into virus taxonomy when there was uncertainty over which criteria were most taxonomically informative. The use of protein structure relationships of RdRP and potentially other replication-associated enzymes provides a valuable new method to determine deeper evolutionary histories of RNA viruses and has substantiated the family and order relationships proposed. |

|  |
| --- |
| **References:** |
| 1. **Debat H, Bejerman N**. Two novel flavi-like viruses shed light on the plant-infecting koshoviruses. *Arch Virol* 2023;168(7):184.  2. **Simmonds P, Becher P, Bukh J, Gould EA, Meyers G et al.** ICTV virus taxonomy profile: *Flaviviridae*. *J Gen Virol* 2017;98(1):2-3.  3. **Francki RIB, Fauquet CM, Knudson DL, Brown F** **(editors)**. *Classification and nomenclature of viruses. Fifth report of the International Committee on Taxonomy of Viruses. Archives of Virology Supplementum 2*. Vienna, Austria: Springer-Verlag; 1991.  4. **Westaway EG, Brinton MA, Gaidamovich SY, Horzinek MC, Igarashi A et al.** *Flaviviridae*. Report of the *Togaviridae* Study Group, Vertebrate Virus Subcommittee, International Committee on Taxonomy of Viruses. *Intervirology* 1985;24(4):183-192.  5. **Choo QL, Richman KH, Han JH, Berger K, Lee C et al.** Genetic organization and diversity of the hepatitis C virus. *ProcNatlAcadSciUSA* 1991;88:2451-2455.  6. **Scheel TK, Simmonds P, Kapoor A**. Surveying the global virome: identification and characterization of HCV-related animal hepaciviruses. *Antiviral Res* 2015;115:83-93.  7. **Stapleton JT, Foung S, Muerhoff AS, Bukh J, Simmonds P**. The GB viruses: a review and proposed classification of GBV-A, GBV-C (HGV), and GBV-D in genus Pegivirus within the family Flaviviridae. *J Gen Virol* 2011;92(Pt 2):233-246.  8. **Wu Z, Wu Y, Zhang W, Merits A, Simmonds P et al.** The First Nonmammalian Pegivirus Demonstrates Efficient In Vitro Replication and High Lymphotropism. *J Virol* 2020;94(20).  9. **Postler TS, Beer M, Blitvich BJ, Bukh J, de Lamballerie X et al.** Renaming of the genus Flavivirus to Orthoflavivirus and extension of binomial species names within the family Flaviviridae. *Arch Virol* 2023;168(9):224.  10. **Choo Q-L, Richman KH, Han JH, Berger K, Lee C et al.** Genetic organization and diversity of the hepatitis C virus. *Proc Natl Acad Sci U S A* 1991;88(6):2451-2455.  11. **Orba Y, Matsuno K, Nakao R, Kryukov K, Saito Y et al.** Diverse mosquito-specific flaviviruses in the Bolivian Amazon basin. *J Gen Virol* 2021;102(3).  12. **Blasdell KR, Wynne JW, Perera D, Firth C**. First detection of a novel 'unknown host' flavivirus in a Malaysian rodent. *Access Microbiol* 2021;3(4):000223.  13. **Gravina HD, Suzukawa AA, Zanluca C, Cardozo Segovia FM, Tschá MK et al.** Identification of insect-specific flaviviruses in areas of Brazil and Paraguay experiencing endemic arbovirus transmission and the description of a novel flavivirus infecting Sabethes belisarioi. *Virology* 2019;527:98-106.  14. **Kobayashi D, Watanabe M, Faizah AN, Amoa-Bosompem M, Higa Y et al.** Discovery of a Novel Flavivirus (Flaviviridae) From the Horse Fly, Tabanus rufidens (Diptera: Tabanidae): The Possible Coevolutionary Relationships Between the Classical Insect-Specific Flaviviruses and Host Dipteran Insects. *J Med Entomol* 2021;58(2):880-890.  15. **Korkusol A, Takhampunya R, Hang J, Jarman RG, Tippayachai B et al.** A novel flavivirus detected in two Aedes spp. collected near the demilitarized zone of the Republic of Korea. *J Gen Virol* 2017;98(5):1122-1131.  16. **McLean BJ, Hobson-Peters J, Webb CE, Watterson D, Prow NA et al.** A novel insect-specific flavivirus replicates only in Aedes-derived cells and persists at high prevalence in wild Aedes vigilax populations in Sydney, Australia. *Virology* 2015;486:272-283.  17. **Smith DB, Meyers G, Bukh J, Gould EA, Monath T et al.** Proposed revision to the taxonomy of the genus Pestivirus, family Flaviviridae. *J Gen Virol* 2017;98(8):2106-2112.  18. **Smith DB, Becher P, Bukh J, Gould EA, Meyers G et al.** Proposed update to the taxonomy of the genera Hepacivirus and Pegivirus within the Flaviviridae family. *J Gen Virol* 2016;97(11):2894-2907.  19. **Hartlage AS, Cullen JM, Kapoor A**. The Strange, Expanding World of Animal Hepaciviruses. *Annual review of virology* 2016;3(1):53-75.  20. **Ridpath JF**. A need to define characteristics to be used in the taxonomy of the expanding pestivirus genus. *Berl Munch Tierarztl Wochenschr* 2013;126(11-12):462-467.  21. **Shi M, Lin XD, Vasilakis N, Tian JH, Li CX et al.** Divergent Viruses Discovered in Arthropods and Vertebrates Revise the Evolutionary History of the Flaviviridae and Related Viruses. *J Virol* 2015;90(2):659-669.  22. **Matsumura EE, Nerva L, Nigg JC, Falk BW, Nouri S**. Complete Genome Sequence of the Largest Known Flavi-Like Virus, Diaphorina citri flavi-like virus, a Novel Virus of the Asian Citrus Psyllid, Diaphorina citri. *Genome Announc* 2016;4(5).  23. **Shi M, Lin X-D, Vasilakis N, Tian J-H, Li C-X et al.** Divergent viruses discovered in arthropods and vertebrates revise the evolutionary history of the *Flaviviridae* and related viruses. *J Virol* 2016;90(2):659-669.  24. **Kartashov MY, Gladysheva AV, Shvalov AN, Tupota NL, Chernikova AA et al.** Novel Flavi-like virus in ixodid ticks and patients in Russia. *Ticks Tick Borne Dis* 2023;14(2):102101.  25. **Petrone ME, Grove J, Mélade J, Mifsud JCO, Parry RH et al.** A ~40-kb flavi-like virus does not encode a known error-correcting mechanism. *Proc Natl Acad Sci U S A* 2024;121(30):e2403805121.  26. **Qin X-C, Shi M, Tian J-H, Lin X-D, Gao D-Y et al.** A tick-borne segmented RNA virus contains genome segments derived from unsegmented viral ancestors. *Proc Natl Acad Sci U S A* 2014;111(18):6744-6749.  27. **Ladner JT, Wiley MR, Beitzel B, Auguste AJ, Dupuis AP, 2nd et al.** A multicomponent animal virus isolated from mosquitoes. *Cell host & microbe* 2016;20(3):357-367.  28. **Zhang S, Yang C, Qiu Y, Liao R, Xuan Z et al.** Conserved untranslated regions of multipartite viruses: Natural markers of novel viral genomic components and tags of viral evolution. *Virus evolution* 2024;10(1):veae004.  29. **Paraskevopoulou S, Käfer S, Zirkel F, Donath A, Petersen M et al.** Viromics of extant insect orders unveil the evolution of the flavi-like superfamily. *Virus evolution* 2021;7(1):veab030.  30. **Colmant AMG, Charrel RN, Coutard B**. Jingmenviruses: Ubiquitous, understudied, segmented flavi-like viruses. *Frontiers in microbiology* 2022;13:997058.  31. **Mifsud JCO, Costa VA, Petrone ME, Marzinelli EM, Holmes EC et al.** Transcriptome mining extends the host range of the Flaviviridae to non-bilaterians. *Virus evolution* 2023;9(1):veac124.  32. **Parry R, Asgari S**. Discovery of novel crustacean and cephalopod flaviviruses: insights into the evolution and circulation of flaviviruses between marine invertebrate and vertebrate hosts. *J Virol* 2019;93(14):e00432-00419.  33. **Lay CL, Shi M, Buček A, Bourguignon T, Lo N et al.** Unmapped RNA Virus Diversity in Termites and their Symbionts. *Viruses* 2020;12(10).  34. **Dong X, Wang G, Hu T, Li J, Li C et al.** A Novel Virus of Flaviviridae Associated with Sexual Precocity in Macrobrachium rosenbergii. *mSystems* 2021;6(3):e0000321.  35. **Bekal S, Domier LL, Gonfa B, McCoppin NK, Lambert KN et al.** A novel flavivirus in the soybean cyst nematode. *J Gen Virol* 2014;95(Pt 6):1272-1280.  36. **Dheilly NM, Lucas P, Blanchard Y, Rosario K**. A World of Viruses Nested within Parasites: Unraveling Viral Diversity within Parasitic Flatworms (Platyhelminthes). *Microbiol Spectr* 2022;10(3):e0013822.  37. **Chen Y-M, Sadiq S, Tian J-H, Chen X, Lin X-D et al.** RNA viromes from terrestrial sites across China expand environmental viral diversity. *Nat Microbiol* 2022;7(8):1312-1323.  38. **Hewson I, Johnson MR, Tibbetts IR**. An Unconventional Flavivirus and Other RNA Viruses in the Sea Cucumber (Holothuroidea; Echinodermata) Virome. *Viruses* 2020;12(9).  39. **Geoghegan JL, Di Giallonardo F, Cousins K, Shi M, Williamson JE et al.** Hidden diversity and evolution of viruses in market fish. *Virus evolution* 2018;4(2):vey031.  40. **Skoge RH, Brattespe J, Økland AL, Plarre H, Nylund A**. New virus of the family Flaviviridae detected in lumpfish (Cyclopterus lumpus). *Arch Virol* 2018;163(3):679-685.  41. **Soto E, Camus A, Yun S, Kurobe T, Leary JH et al.** First Isolation of a Novel Aquatic Flavivirus from Chinook Salmon (Oncorhynchus tshawytscha) and Its In Vivo Replication in a Piscine Animal Model. *J Virol* 2020;94(15).  42. **Costa VA, Mifsud JCO, Gilligan D, Williamson JE, Holmes EC et al.** Metagenomic sequencing reveals a lack of virus exchange between native and invasive freshwater fish across the Murray-Darling Basin, Australia. *Virus evolution* 2021;7(1):veab034.  43. **Parry RH, Slonchak A, Campbell LJ, Newton ND, Debat HJ et al.** A novel tamanavirus (Flaviviridae) of the European common frog (Rana temporaria) from the UK. *J Gen Virol* 2023;104(12).  44. **Urayama S-I, Takaki Y, Nunoura T**. FLDS: A Comprehensive dsRNA Sequencing Method for Intracellular RNA Virus Surveillance. *Microbes Environ* 2016;31(1):33-40.  45. **Forgia M, Daghino S, Chiapello M, Ciuffo M, Turina M**. New clades of viruses infecting the obligatory biotroph *Bremia lactucae* representing distinct evolutionary trajectory for viruses infecting oomycetes. *Virus evolution* 2024;10(1):veae003.  46. **Schönegger D, Marais A, Faure C, Candresse T**. A new flavi-like virus identified in populations of wild carrots. *Arch Virol* 2022;167(11):2407-2409.  47. **Simmonds P, Butković A, Grove J, Mayne R, Mifsud JCO et al.** Integrated analysis of protein sequence and structure redefines viral diversity and the taxonomy of the Flaviviridae. *bioRxiv* 2025.  48. **Simmonds P, Adriaenssens EM, Zerbini FM, Abrescia NGA, Aiewsakun P et al.** Four principles to establish a universal virus taxonomy. *PLoS biology* 2023;21(2):e3001922.  49. **Nasir A, Romero-Severson E, Claverie JM**. Investigating the Concept and Origin of Viruses. *Trends Microbiol* 2020;28(12):959-967.  50. **Brussow H**. The not so universal tree of life or the place of viruses in the living world. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 2009;364(1527):2263-2274.  51. **Krupovic M, Dolja VV, Koonin EV**. Origin of viruses: primordial replicators recruiting capsids from hosts. *Nat Rev Microbiol* 2019;17(7):449-458.  52. **Mifsud JCO, Lytras S, Oliver MR, Toon K, Costa VA et al.** Mapping glycoprotein structure reveals Flaviviridae evolutionary history. *Nature* 2024.  53. **Vaney MC, Dellarole M, Duquerroy S, Medits I, Tsouchnikas G et al.** Evolution and activation mechanism of the flavivirus class II membrane-fusion machinery. *Nat Commun* 2022;13(1):3718.  54. **Jumper J, Evans R, Pritzel A, Green T, Figurnov M et al.** Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;596(7873):583-589.  55. **Varadi M, Bertoni D, Magana P, Paramval U, Pidruchna I et al.** AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res* 2024;52(D1):D368-d375.  56. **van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J et al.** Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* 2024;42(2):243-246.  57. **Moi D, Bernard C, Steinegger M, Nevers Y, Langleib M et al.** Structural phylogenetics unravels the evolutionary diversification of communication systems in gram-positive bacteria and their viruses. *bioRxiv* 2023.  58. **Mayne R, Aiewsakun P, Turner D, Adriaenssens EM, Simmonds P**. GRAViTy-V2: a grounded viral taxonomy application. *bioRxiv* 2024:2024.2007.2026.605250.  59. **Marin MS, Zanotto PM, Gritsun TS, Gould EA**. Phylogeny of TYU, SRE, and CFA virus: different evolutionary rates in the genus Flavivirus. *Virology* 1995;206(2):1133-1139.  60. **Kobayashi K, Atsumi G, Iwadate Y, Tomita R, Chiba K et al.** Gentian Kobu-sho-associated virus: a tentative, novel double-stranded RNA virus that is relevant to gentian Kobu-sho syndrome. *J Gen Plant Pathol* 2013;79:56–63.  61. **Blum M, Andreeva A, Florentino LC, Chuguransky SR, Grego T et al.** InterPro: the protein sequence classification resource in 2025. *Nucleic Acids Res* 2024. |

|  |  |
| --- | --- |
| **Accompanying files:** | |
| **Filename** | **Description of contents** |
|  |  |
|  |  |

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

**Table 1**. Listing of clades and lineages of flavivirids and ‘flavi-like’ viruses and suggested revised and expanded taxonomy

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Clade** | **Lineage** | **n1** | **Genus2** | **Example3** | **Example virus name** | **Comments** | |
| I | ⓐ | 2 |  | OX394137 | Cnidaria flavivirus (CnFV) | Marine invertebrates, including acorn worms (Enteropneusta) and Cnidaria | |
| ⓑ | 95 | *Orthoflavivirus* | U87411 | dengue virus type 2 (DENV-2) | Expansion of established genus, several insect-specific flaviviruses | |
| ⓒ | 39 |  | KJ741267 | cell-fusing agent virus (CFAV) | Large group of ‘ISFs’ | |
| ⓓ | 5 |  | MK473878 | Crangon crangon flavivirus (CCFV) | Crustacean and other marine hosts | |
| ⓔ | 2 |  | MT075326 | salmon flavivirus (SaFV) |  | |
| ⓕ | 3 |  | AF285080 | Tamana bat virus (TABV) | Infects diverse lineage with marine hosts | |
| ⓖ | 3 |  | MW052131 | Waxsystermes virus | Diverse host range, arthropods, shark | |
| ⓗ | 2 |  | OX394159 | Chowder Bay tunicate associated flavi-like virus | Marine invertebrates | |
| ⓘ | 3 |  | BK059737, BK059738 | fluviflavili virus | Platyhelminth and environmental sources | |
| ⓙ | 2 |  | MT084113 | infectious precocity virus | Infect crustaceans | |
| ⓚ | 9 |  | KM521566 - KM521570 | Guaico Culex virus (GCuV) | Arthropod-specific viruses, 5 segments | |
| ⓛ | 19 |  | KJ001579 - KJ001582 | Jīngmén tick virus (JMTV) | Arthropod host (ticks), vector-borne transmission to mammals, 4 segments, | |
| II | ⓜ | 3 |  | KR902730 | Xīnzhōu spider virus (XSV) | Infect several arachnid species | |
| ⓝ | 2 |  | MG599985 | Xiàmén fanray pesti-like virus (XFPV) | Cartilaginous fish (Chondrichthyes) | |
| ⓞ | 34 | *Pestivirus* | M96751 | bovine viral diarrhea virus 1 | Expansion of established genus, primarily mammalian | |
| ⓟ | 1 |  | *SRR7976360* |  | Soil metagenome sequence | |
| ⓠ | 39 |  | KR902736 | Bólè tick virus 4 | Tick and diverse invertebrate hosts | |
| ⓡ | 15 |  | KR902734 | Shuāngào lacewing virus 2 | Wide range of insect hosts | |
| ⓢ | 20 |  | BK062903 | Sonchus virus | Diverse, hosts include arthropods, nematodes and angiosperm plants | |
| III | ⓣ | 1 |  | MG599990 | Wēnlǐng moray eel hepacivirus |  | |
| ⓤ | 54 | *Pegivirus* | U44402 | human pegivirus genotype 2 | Expansion of established genus, mammalian and avian hosts | |
| ⓥ | 1 |  | MG334001 | (hepacivirus sp.) | Detected in a terrapin | |
| ⓦ | 101 | *Hepacivirus* | AF009606 | hepatitis C virus genotype 1a | Expansion of established genus, mammalian, reptile, avian, cartilaginous and bony fish hosts | |
| IV | (unclassified) | | | AP014912 | diatom colony-associated ssRNA virus | | Environmental ‘flavi-like’ viruses |

1Available distinct coding-complete genome sequences; examples representing species range within the larger groups.

2Established genus assignments

3INDSC accession number of example virus.

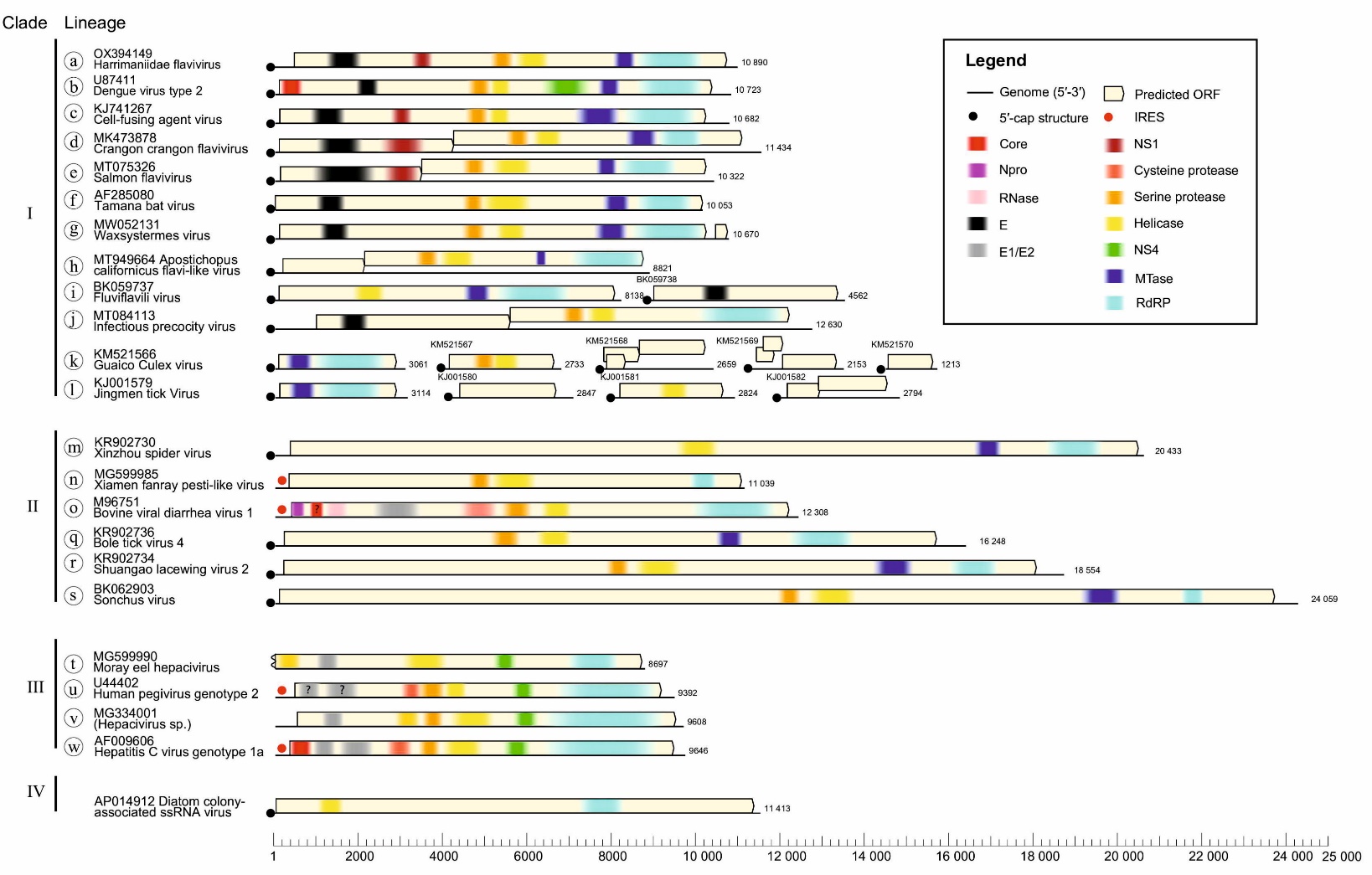
**Table 2**. Proposed taxonomic assignments of “flavi-like” viruses

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Phylogeny** | | **Proposed new taxonomic assignments1** | | | | **Exemplar** | |
| **Clade** | **Lineage** | **Family** | **Genus** | **Subgenus** | **Species name2** | **Virus name** | **Accession no.2** |
| *Families* | | | | | | | |
| I |  | ***Flaviviridae*** |  |  |  |  |  |
| II | *Pestiviridae* |  |  |  |  |  |
| III | *Hepaciviridae* |  |  |  |  |  |
| *Genera, subgenera and species* | | | | | | | |
| I | Ia, Ib, Ic, Id | ***Flaviviridae*** | ***Orthoflavivirus*** |  | | | |
| Ib | ***Euflavivirus*** | ***Orthoflavivirus dengue*** | dengue virus type 2 (DENV-2) | U87411 |
| Ic | *Fusivirus* | *Orthoflavivirus iunctionis* | cell-fusing agent virus (CFAV) | MK473878 |
| Id | *Crangovirus* | *Orthoflavivirus alphei* | Crangon crangon flavivirus (CCFV) | MK473878 |
| If | *Tamanavirus* |  | *Tamanavirus parnellis* | Tamana bat virus (TABV) | AF285080 |
| Ig | *Termitovirus* |  | *Termitovirus isoptera* | Waxsystermes virus | MW052131 |
| Ik | *Guaicovirus* |  | *Guaicovirus culicis* | Guaico Culex virus (GCuV) | KM521566 - KM521570 |
| Il | *Jingmenvirus* |  | *Jingmenvirus rhipicephali* | Jīngmén tick virus (JMTV) | KJ001579 - KJ001582 |
|  | | | | | | | |
| II | IIm | *Pestiviridae* | *Arachnivirus* |  | *Arachnivirus neosconae* | Xīnzhōu spider virus 3 (XSV) | KR902730 |
| IIo | ***Orthopestivirus*** |  | ***Orthopestivirus bovis*** | bovine viral diarrhea virus 1 | M96751 |
| IIq | *Boletivirus* |  | *Boletivirus hyalommae* | Bólè tick virus 4 | KR902736 |
| IIr | *Chrysopivirus* |  | *Chrysopivirus vittae* | Shuāngào lacewing virus 2 | KR902734 |
| IIs | *Koshovirus* |  | *Koshovirus sonchi* | Sonchus virus | BK062903 |
|  | | | | | | | |
| III | IIIu | *Hepaciviridae* | ***Pegivirus*** |  | ***Pegivirus hominis*** | human pegivirus genotype 2 | U44402 |
| IIIw | ***Orthohepacivirus*** |  | ***Orthohepacivirus hominis*** | hepatitis C virus genotype 1a | AF009606 |

1Established taxa are shown in ***bold***; ***red*** if renamed; genera may contain additional members and potential species

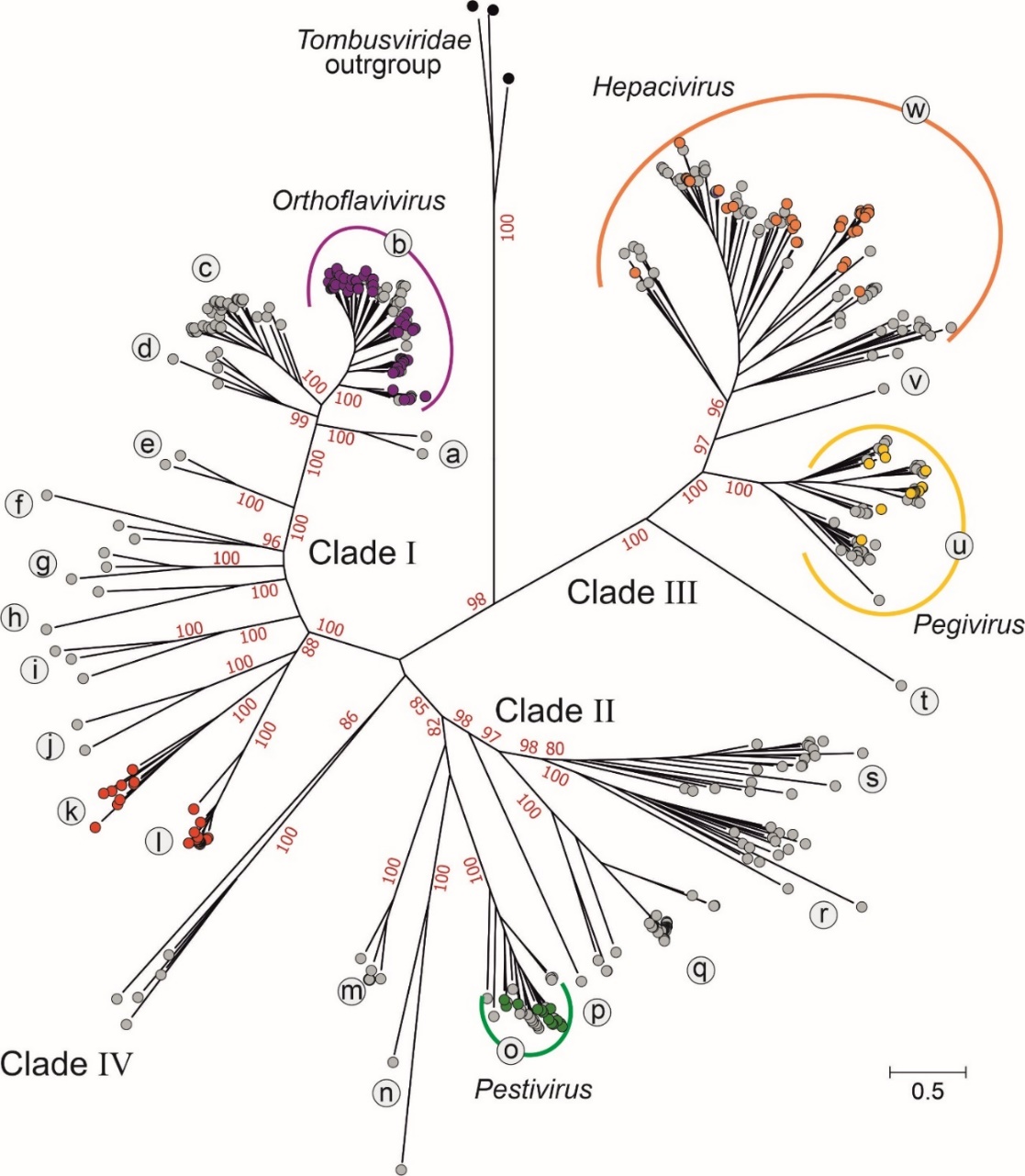
2Exemplar species shown

**Fig. 1**. Organization of example genomes in each lineage of flavivirids and ‘flavi-like’ viruses.

****

Genome diagrams for the example viruses listed in Table 1 drawn to scale (lower scale bar) and main functional domains identified by InterProScan browser v. 103 (<https://www.ebi.ac.uk/interpro/search/sequence/>) [61]

**Fig. 2.**  RNA-directed RNA polymerase domain (RdRP) amino-acid sequence phylogenies differentiate flavivirids and ‘flavi-like’ viruses into four highly supported main clades.



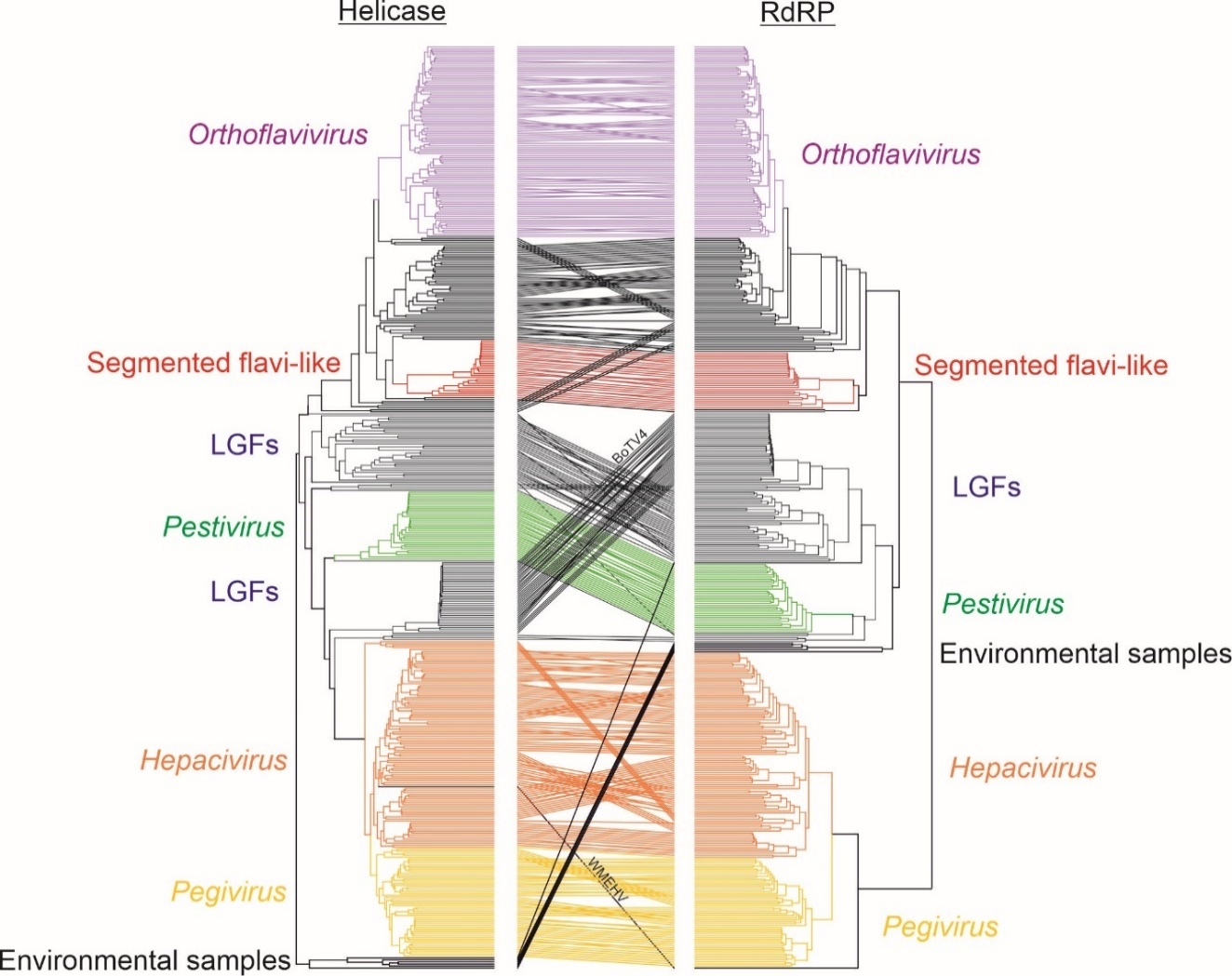
Maximum likelihood tree of aligned flavivirid and ‘flavi-like’ RNA-directed RNA polymerase (RdRP) domain amino-acid sequences, estimated using the LG+F+R10 model using IQ-TREE with tombusvirid sequences as an outgroup. Bootstrap support values for the main branches of the tree are shown in red if ≥70%. Already classified flavivirids are shown in color-filled circles. The main clades are numbered I–IV and lineages are labeled with lower-case letters. The component sequences within each clade are provided in a fully annotated tree (RdRP\_tree.NWK (File Sq; Suppl. Data in [47]) and listed in Table S1 (Suppl. Data; [47]).

Fig. 3. RNA-directed RNA polymerase domain (NS5/NS5B) amino-acid sequence phylogenies differentiate flavivirids and ‘flavi-like’ viruses into four highly supported main clades

## 

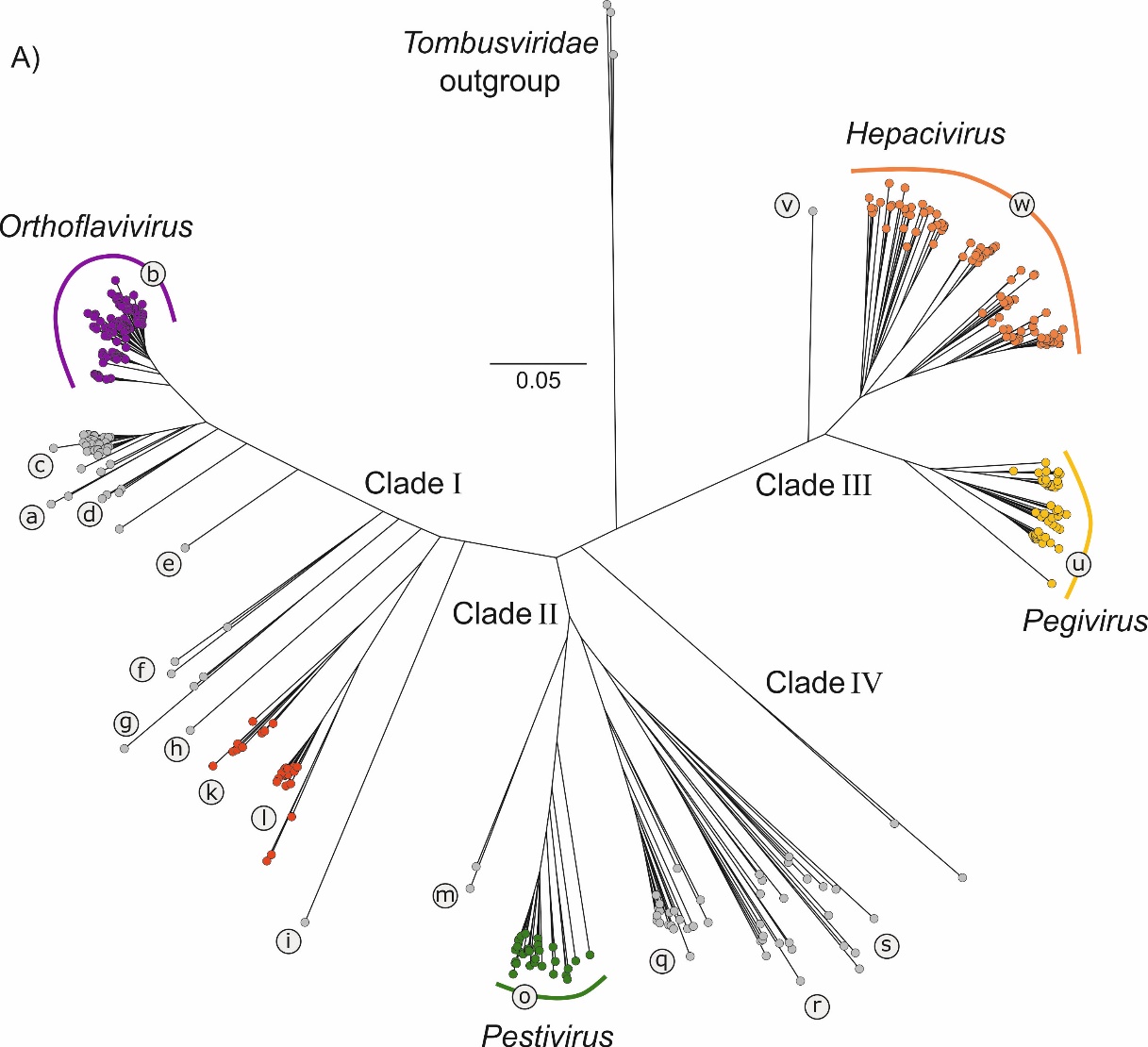
Phylogenetic trees constructed by likelihood (A, B) and distance-based (C) methods using flavivirid and ‘flavi-like’ RNA-directed RNA polymerase (RdRP) domain amino-acid sequences. Clades (I–IV) and lineages (a–w) labelled in each tree are based on those in Fig. 2. Tentative threshold levels of divergence separating clades and lineages are shown as red dotted lines in BEAST and UPGMA trees. An alternative threshold corresponding to the assignment of lineages Ia-Id to a common lineage is shown in a blue dotted line. Abbreviations: BP, before present; BEAST, Bayesian evolutionary analysis by sampling trees cross-platform program; JTT, Jones-Taylor-Thornton matrix; ML, maximum likelihood; UPGMA, unweighted pair group method with arithmetic mean.

## Fig. 4. Helicase domain (NS3) amino-acid sequence phylogeny supports partition of flavivirids and ‘flavi-like’ viruses into four main clades



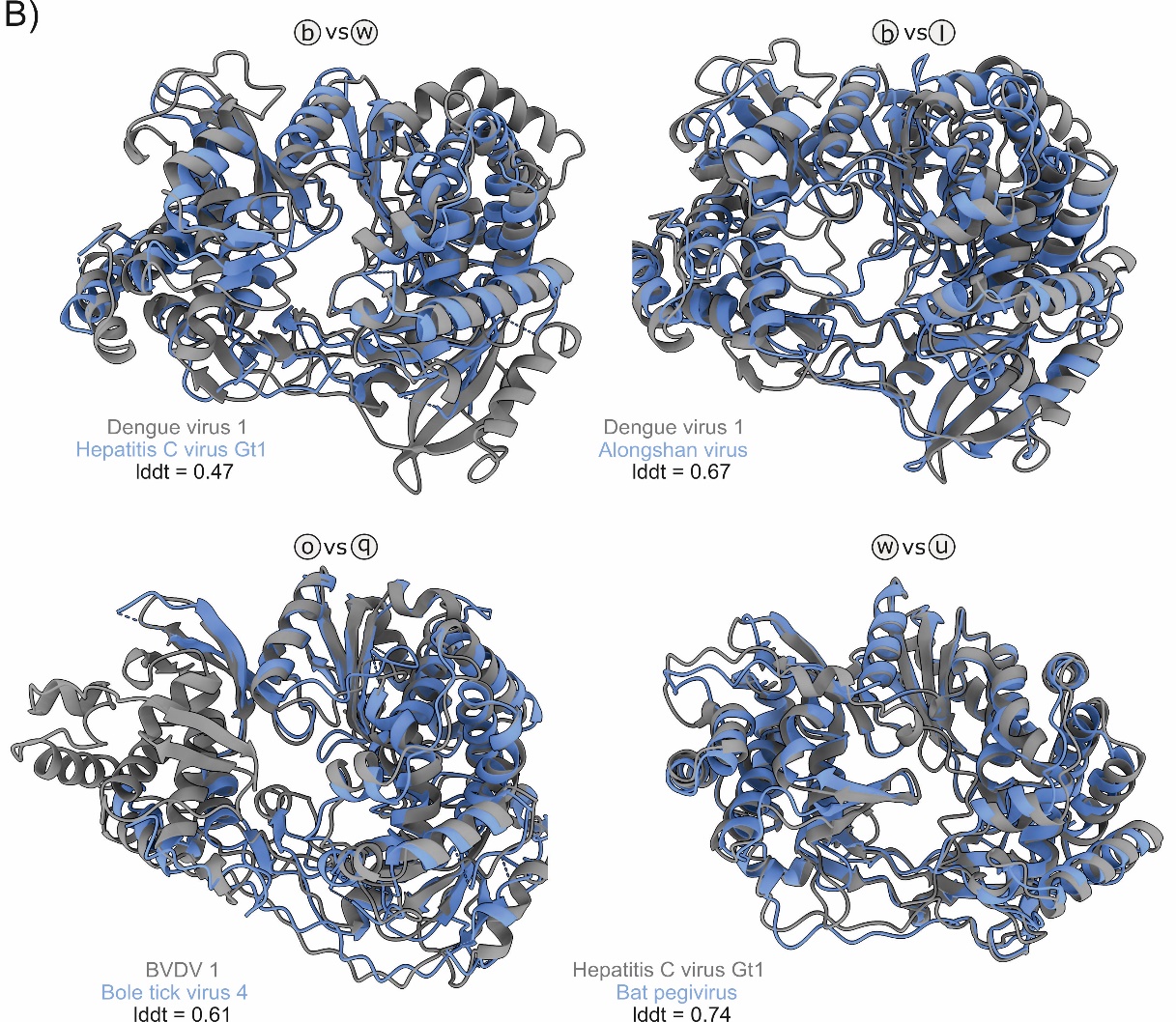
Tanglegram of flavivirid and ‘flavi-like’ virus helicase and RNA-directed RNA polymerase (RdRP) domain sequences constructed from ML phylogenetic trees generated by IQ-TREE trees, with established flavivirid genera colored as in Figs. 2 and 3, and segmented ‘flavi-like’ viruses (lineages k and l in red). BoTV4, Bólè tick virus 4; LGFs, large genome flaviviruses, WMEHV, Wēnlǐng moray eel hepacivirus. A copy of the figure with the branches individually labelled is provided as Fig. S1 [47].

**Fig. 5.** RNA-directed RNA polymerase domain (NS5/NS5B) structural comparison supports partition of flavivirids and ‘flavi-like’ viruses into four main clades



A)

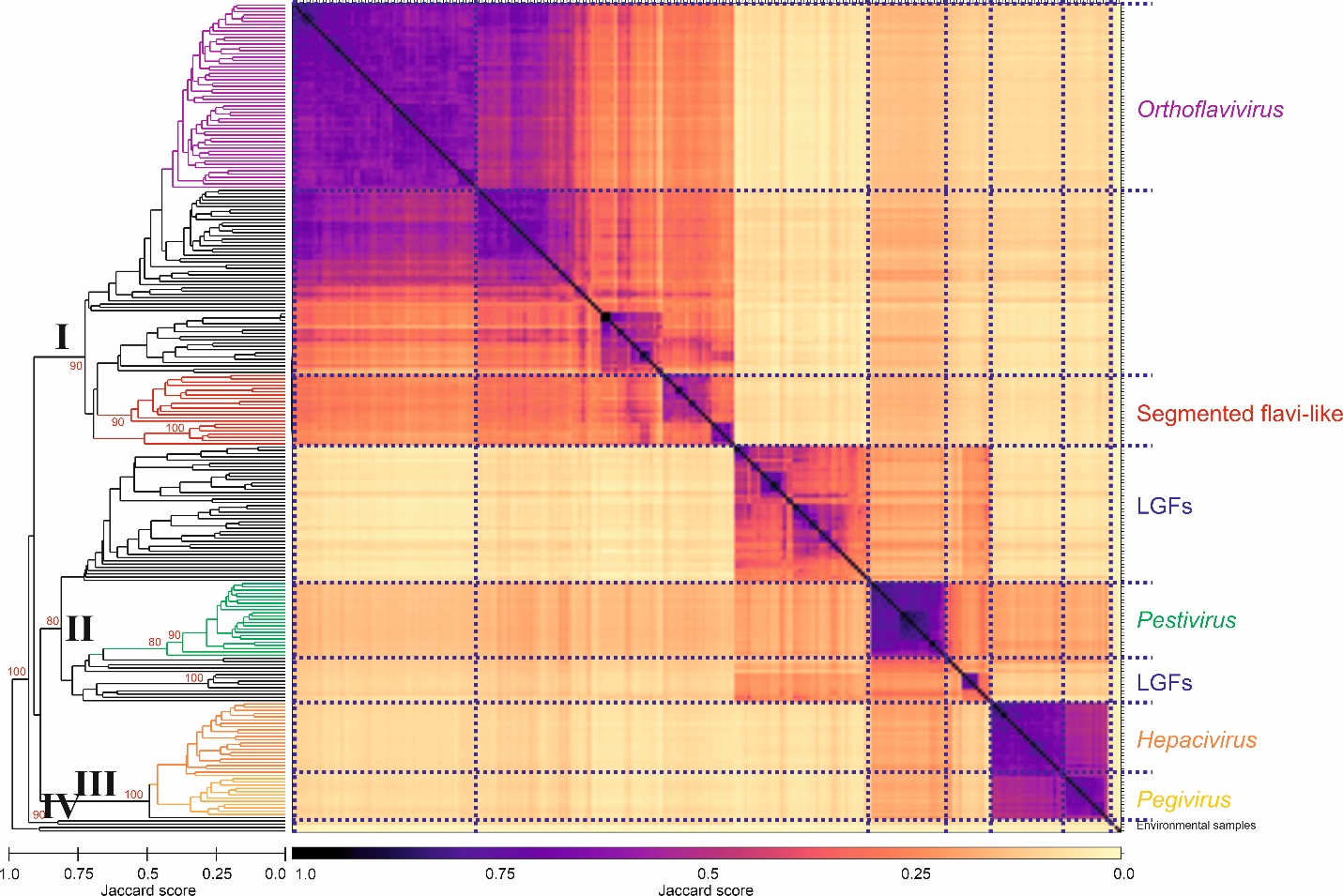
Structure-based tree of 400 flavivirids and ‘flavi-like’ viruses’ RdRP domains, derived from a lddt distance matrix (calculated by FoldTree [57], powered by Foldseek [56]), scale bar indicates lddt distance (which is approximate to the inverse of the pairwise lddt score). Main clades and lineages are labelled as in Figure 2.

****

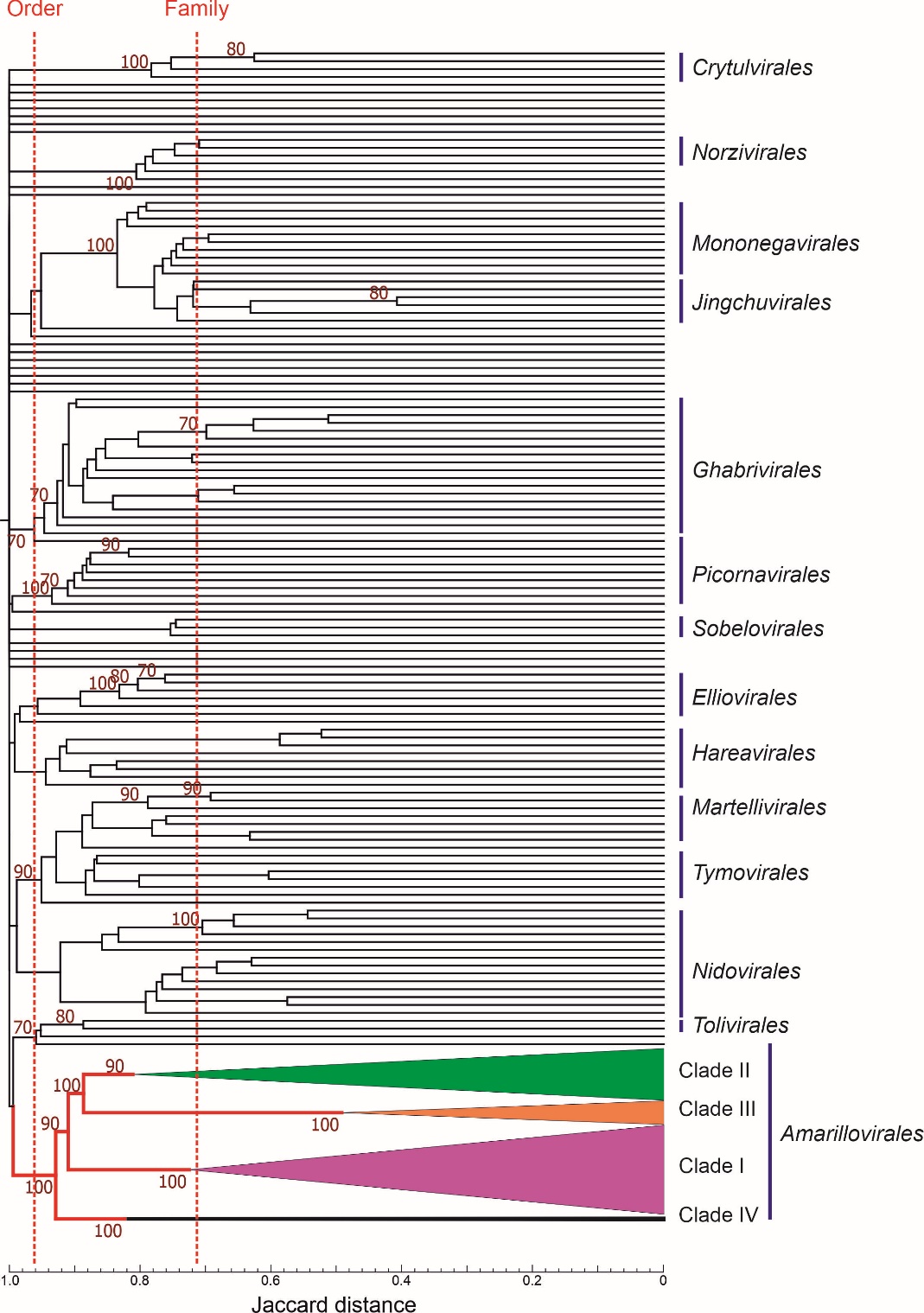
B)

Examples of aligned RdRP domain structures, color-coded as stated in the label, lineage identifiers (e.g. b vs w) indicate the position of the compared structures on the tree. lddt values represent structural similarity, with values of 1 being perfectly aligned identical structures. BVDV, bovine viral diarrhea virus.

## Fig. 6. Alignment-free hidden Markov model homology analysis supports partition of flavivirids and ‘flavi-like’ viruses into one order and four family rank clades

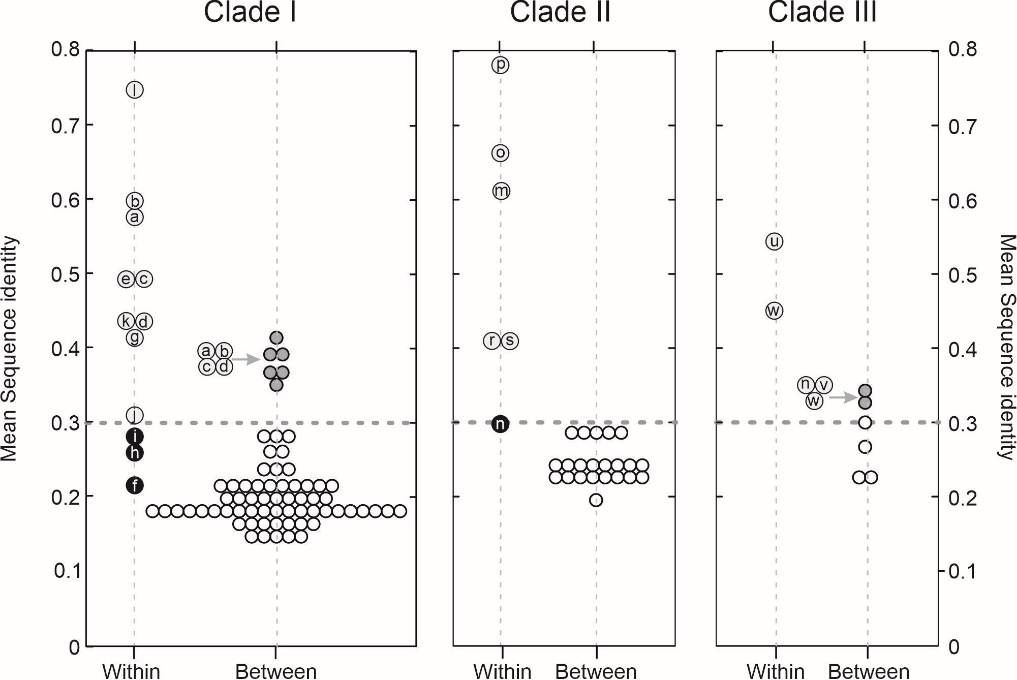


A) GRAViTy Jaccard distances calculated for classified flavivirids and ’flavi-like viruses and a representative member of each established RNA virus family in ribovirian kingdom *Orthornavirae* (n=135), showing approximate demarcation thresholds for orders and families (dashed vertical red lines). Bootstrap support values (10 iterations) are shown in red if ≥70%. A copy of the figure with the branches individually labelled is provided as Fig. S2 [47].



B) Heatmap and dendrogram depicting relationships among classified flaviviridsand ‘flavi-like’ viruses. Clades I–IV identified in the RdRP phylogeny (Fig. 2) were added to equivalent branches in dendrogram. Bootstrap support values (10 iterations) for deeper branches are shown in red if ≥70%.

**Fig. 7**. Within- and between-sequence diversity of lineages identified on phylogenetic analysis of RdRP sequences

****

Mean pairwise sequence identities of RdRP domain amino acid sequences between and within lineages of clades I-III. Dotted line indicates an approximate threshold dividing within- and between-lineage distances; between-lineage comparisons above threshold shaded in grey; within-lineage distances below the inter-lineage threshold shown in black.

**Fig. 8.** Diagrammatic summary of proposed changes to flavivirid taxonomy

