

The International Committee on Taxonomy of Viruses

Taxonomy Proposal Form, 2025

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

|  |  |
| --- | --- |
| **Title:** | Eight new families of archaeal viruses within the class *Caudoviricetes* |
| **Code assigned:** | 2025.002A.N.v2.Archaeal\_Caudoviricetes\_8nf | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Author(s), affiliation and email address(es):** | | | | |
| **Given name (+middle initial(s))** | **Surname** | **Affiliation** | **Email address** | **Corr. author(s)** |
| Yifan | Zhou | Institut Pasteur, Université Paris Cité, CNRS UMR6047, Archaeal Virology Unit, Paris, France | [yifan.zhou@pasteur.fr](mailto:yifan.zhou@pasteur.fr) |  |
| Mart | Krupovic | Institut Pasteur, Université Paris Cité, CNRS UMR6047, Archaeal Virology Unit, Paris, France | [mart.krupovic@pasteur.fr](mailto:mart.krupovic@pasteur.fr) | 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 |  |
| Animal positive-strand RNA viruses |  | Plant viruses |  |
| Archaeal viruses | **X** | 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:** | 20/06/2025 |

**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| W - Withdrawn |  |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**

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

|  |  |
| --- | --- |
| **Etymology (origin) of proposed taxonomic names:** | |
| **Taxon name** | **Etymology of the term** |
| *Danacaviridae* | after **Dana**kil depression (source of isolation/sequencing) + **ca**udo |
| *Dallocaviridae* | after **Dallo**l (source of isolation/sequencing) + **ca**udo |
| *Assalcaviridae* | after lake **Assal**e (source of isolation/sequencing) + **ca**udo |
| *Quasboviridae* | after **quasbo**, salt in Afar language |
| *Gulliviridae* | after **Gulli**ver, the protagonist of Jonathan Swift's novel Gulliver's Travels, referring to the large size of the genome |
| *Lilliviridae* | after fictional **Lilli**put Island from Jonathan Swift's novel Gulliver's Travels |
| *Blefusviridae* | after fictional **Blefus**cu Island from Jonathan Swift's novel Gulliver's Travels |
| *Saladoviridae* | after **salado**, salty in Spanish |
| *Gablavirus* | after **gabla** for cave in Afar language (place of isolation/sequencing) |
| *Kalovirus* | after kalo for lake in Afar language |
| *Ethicavirus* | abbreviation of **Ethi**opia (place of isolation/sequencing) + **ca**udovirus |
| *Morpovirus* | referring to a larger portal gene in this genus (**mor**e **po**rtal) |
| *Traglyvirus* | reference to Queuine tRNA-guanine **tra**ns**gly**cosylase specific to this genus within *Madisaviridae* |
| *Karumvirus* | after lake **Karum** (also known as Assale; place of isolation/sequencing) |
| *Haroovirus* | after **haroo** for lake in Oromo language |
| *Cusbovirus* | after **cusbo** for salt in Somali language |
| *Lemuelvirus* | after **Lemuel**, the first name of Gulliver, the protagonist of Jonathan Swift's novel Gulliver's Travels |
| *Mildendovirus* | after **Mildendo** fictional capital of Lilliput island |
| *Wecalvirus* | after **We**st **Ca**nyon **L**ake (source of isolation/sequencing) |
| *Crypovirus* | from **cry**stallizer **po**nd, source of isolation/sequencing |
| *Latyvirus* | from **La**ke **Ty**rrell, Australia, source of isolation/sequencing |
| *danakilense* | Latinized species epithet is derived from **Danakil** depression, source of isolation/sequencing |
| *alicantense* | Latinized species epithet is derived from **Alicante**, source of isolation/sequencing |
| *chilense* | Latinized species epithet is derived from **Chile**, source of isolation/sequencing |
| *nanohalovivens* | from Nanohaloarchaea (host group) + vivens (Latin for "living/dwelling"), indicating the associated with nanohaloarchaea |

|  |  |  |
| --- | --- | --- |
| **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*:  *Duplodnaviria, Heunggongvirae, Uroviricota, Caudoviricetes*  *Description of current taxonomy*:  Head-tailed viruses infecting halophilic archaea (class Halobacteria) are currently classified into 12 families within class *Caudoviricetes*, whereas viruses infecting nanohaloarchaeal hosts (phylum Nanohaloarchaeota) remain unclassified.  *Proposed* *taxonomic change(s):*  We propose classifying 14 head-tailed viruses associated with haloarchaeal and nanohaloarchaeal hosts into 14 new species within 13 new genera, which are assigned to 8 new families and 3 existing families within the class *Caudoviricetes*.  *Justification*:  The classification is based on the proteome-wide phylogenomic analysis implemented in VipTree and further supported by comparative genomics analysis. |

|  |
| --- |
| **Text of Taxonomy proposal:** |
| *Taxonomic rank(s) affected*:  *Duplodnaviria, Heunggongvirae, Uroviricota, Caudoviricetes*  *Description of current taxonomy*:  Head-tailed viruses infecting halophilic archaea (class Halobacteria) are currently classified into 12 families within class *Caudoviricetes*, namely, *Druskaviridae*, *Graaviviridae*, *Hafunaviridae*, *Haloferuviridae*, *Halomagnusviridae*, *Madisaviridae*, *Pyrstoviridae*, *Saparoviridae*, *Shortaselviridae*, *Soleiviridae*, *Suolaviridae*, *Vertoviridae*. By contrast, no viruses infecting nanohaloarchaeal hosts (phylum Nanohaloarchaeota) have been classified thus far.  *Proposed* *taxonomic change(s)*:  We propose classifying 14 head-tailed viruses associated with haloarchaeal and nanohaloarchaeal hosts into 14 new species within 13 new genera, and to assign them to 8 new families and 3 existing families within the class *Caudoviricetes*.  *Demarcation criteria:*  We propose using 95% sequence identity as a species demarcation threshold, to be consistent with the classification of other bacterial and archaeal viruses in the class *Caudoviricetes*. For family demarcation, we used proteome-wide phylogenomic analysis with VipTree, where branch lengths of around 0.05 were previously concluded to correspond to family-level groupings for tailed archaeal viruses.  *Justification*:  Geothermally influenced salt lakes in the Danakil Depression, Ethiopia, some of the most extreme ecosystems known, are dominated by microbial communities consisting of haloarchaea and nanohaloarchaea (Belilla et al., 2021; Gutiérrez-Preciado et al., 2024). Analysis of metagenomes from Lake Assale or Karum (samples Ass and 9Ass collected during different years), cave reservoir at the Dallol proto-volcano salt canyons (9Gt) and two of the Western-Canyon Lakes (WCL2 and WCL3) (Gutiérrez-Preciado et al., 2024) using geNomad (Camargo et al., 2024) and VirSorter2 (Guo et al., 2021) yielded 2,085 viral contigs (≥5 kb). Among these, 11 and 7 complete virus genomes (circular contigs, i.e., with direct terminal redundancies) could be assigned to haloarchaeal and nanohaloarchaeal hosts, respectively, based on specific CRISPR spacer-protospacer matches and BlastP analysis (Zhou et al., 2025). Based on the presence of signature genes involved in virion morphogenesis, these viruses could be broadly assigned to four realm-level groups, namely, (i) head-tailed viruses of the realm *Duplodnaviria* (class *Caudoviricetes*; 8 haloarchaeal viruses [HVs] and 3 nanohaloarchaeal viruses [NHVs]), (ii) tailless icosahedral viruses of the realm *Singelaviria* (1 HVs and 1 NHVs), (iii) pleomorphic viruses of the realm *Monodnaviria* (family *Pleolipoviridae*; 1 HVs and 1 NHVs), and (iv) spindle-shaped viruses which are currently not assigned to any realm (1 HVs and 2 NHVs). In this proposal, we propose the classification of the 11 head-tailed HVs and NHVs from the Danakil Depression as well as 3 additional NHVs identified in GenBank into 14 new species within 13 new genera, and to assign them to 8 new families and 3 existing families within the class *Caudoviricetes* (Table 1).  Comparison of the inferred proteomes of HVs and NHVs showed that the corresponding virus groups are only distantly related to each other (Fig. 1), arguing against recent horizontal virus transfer between haloarchaeal and nanohaloarchaeal hosts. In particular, proteome-wide comparison of HVs and NHVs in the context of the reference archaeal viruses representing diverse *Caudoviricetes* families showed that HVs and NHVs formed distinct families (Fig. 1). Based on the established demarcation criteria, i.e., archaeal viruses of the same *Caudoviricetes* family typically share about 20-50% of orthologous proteins, whereas viruses from different families share less than 10% of proteins (Liu et al., 2021), the 8 HVs could be classified into 7 different families, including 4 new families, whereas the NHVs and their relatives formed 4 new families (Fig. 2). These assignments were consistent with the proteomic tree analysis, where the branch length demarcation for different archaeal *Caudoviricetes* families is ~0.05 (Fig. 1). Notably, members of the same family exhibited similar genome lengths and GC% content (Fig. 1), further supporting the taxonomic assignments. Below we briefly describe each of the families proposed to be created or expanded with new species.  **Families “*Danacaviridae*” and “*Dallocaviridae*”**  The new proposed families “*Danacaviridae*” and and “*Dallocaviridae*” each will include a single representative, Danakil haloarchaeal tailed virus 1 (DHTV1) and DHTV2, respectively. Both viruses have large genomes, DHTV1 (109.7 kb) and DHTV2 (109.5 kb), and encode nearly complete replisomes, including replicative MCM helicase, DNA primase, family B DNA polymerase, DNA polymerase sliding clamp (PCNA), sliding clamp loader (replication factor C) and DNA ligase (Fig. 3), suggesting that these viruses have the capacity to semi-autonomously replicate their genomes.  DHTV1 and DHTV2 encode a number of proteins implicated in DNA modification and nucleotide metabolism. DHTV1 encodes a thymidylate synthase, whereas DHTV2 encodes a thymidylate kinase, inosine-5'-monophosphate dehydrogenase and nucleoside 2-deoxyribosyltransferase, and both viruses carry genes for dCTP deaminases (Fig. 3). In addition, DHTV1 and DHTV2 carry genes for 7-deazaguanine synthase and multiple DNA methyltransferases, which could confer resistance to type II restriction modification systems.  Despite this shared gene content, the overall fraction of shared proteins is below 10% (Fig. 2). Besides, DHTV1 and DHTV2 sequences are highly divergent, as is evident from the proteomic tree (Fig. 1), and homologous genes might not be necessarily orthologous. Indeed, phylogenetic analysis of the major capsid protein, portal and large terminase subunit sequences showed that the proteins from the two viruses do not form monophyletic groups (Fig. 4). Based on these observations, the two viruses are assigned to different families.  **Families “*Assalcaviridae*” and “*Quasboviridae*”**  The proposed families “*Assalcaviridae*” and “*Quasboviridae*” each will include a single representative, DHTV6 and DHTV8, respectively, predicted to infect haloarchaeal hosts. The two viruses have moderately-sized genomes, DHTV1 (41.9 kb) and DHTV8 (32.1 kb), and predominantly encode viral structural proteins (Fig. 5), with only few encoded proteins implicated in genome replication and nucleotide metabolism. Based on the VipTree (Fig. 1), fraction of shared proteins (Fig. 2) and phylogenetic analysis of the core proteins (Fig. 4), the two viruses are not closely related to each other or to other haloarchaeal viruses and thus are classified into separate families.  **Expansion of the existing families *Graaviviridae*, *Madisaviridae* and *Pyrstoviridae***  We propose assigning DHTV3 to familiy *Graaviviridae*; DHTV4 and DHTV5 to family *Madisaviridae*; and DHTV7 to family *Pyrstoviridae*. Notably, thus far, families *Madisaviridae* and *Pyrstoviridae* were represented by singletons. Based on the VipTree analysis, the four viruses cluster with the classified representatives of the corresponding families with the branch lengthsaround 0.05 (Fig. 1) and share with them 14-26% of the proteins (Fig. 2, Fig. 6). Given that members of different genera within the same family typically share >30% of proteins (Liu et al., 2021), the four viruses should be considered as representatives of four new genera.  **Four new families for classification of nanohaloarchaeal viruses: “*Gulliviridae*”, “*Lilliviridae*”, “*Blefusviridae*” and “*Saladoviridae*”**  Nanohaloarchaea are ultra-small archaea of the candidate phylum Nanohaloarchaeota, which establish obligate symbiotic interactions with extreme halophiles of the class Halobacteria. We propose establishing four new families for the classification of 6 viruses discovered in the metagenomes from the Danakil Depression (Fig. 7) and viruses previously sequenced from other geographical locations (Fig. 8).  “*Lilliviridae*” and “*Blefusviridae*” are each represented by a single virus, Danakil nanohaloarchaeal tailed viruses (DNTV) 2 and 3, respectively, detected in the Danakil Depression metagenomes (Fig. 7a). Several relatives of DNTV2 were identified in the IMG/VR database (Fig. 1), but these will not be classified because GenBank accession numbers are not available for these genomes. DNTV3 remains a singleton.  “*Gulliviridae*” includes DNTV1 with a genome of 103.3 kb (Fig. 7a) and a previously described but unclassified virus, LTV2 (Emerson et al., 2012). The circular genome (76.7 kb) of LTV2 has been assembled from metagenomic data obtained from hypersaline Lake Tyrrell, Australia (Emerson et al., 2012). VipTree analysis suggests that the two viruses belong to the same family (Fig. 1). However, there is considerable different in genome size between the two viruses and they share only 22% of their genes. Thus, we propose to classify them into separate genera, “*Lemuelvirus*” and “*Latyvirus*”, respectively, within the family “*Gulliviridae*”.  “*Saladoviridae*” includes two nanohaloarchaeal viruses for which genomes are available in GenBank, namely, eHP-23, which was sequenced from the saturated NaCl brines in Alicante, Spain (Garcia-Heredia et al., 2012), and GNTV1, which was sequenced from a halite rock sample collected from the Salar Grande, Chile (Crits-Christoph et al., 2016). Related virus genomes were also detected in the IMG/VR database (Fig. 8). Notably, eHP-23 and GNTV1 share 35% of proteins and hence are proposed to be classified as distinct species within the same genus, “*Crypovirus*”.  The four groups of viruses formed distinct clades in the VipTree analysis with the branch lengths compatible with the family level demarcation (Fig. 1). Consistently, representatives of the four putative families share less than 10% of proteins with viruses from other families. |

|  |
| --- |
| **References:** |
| Belilla J, Iniesto M, Moreira D, Benzerara K, López-García JM, López-Archilla AI, Reboul G, Deschamps P, Gérard E, López-García P. Archaeal overdominance close to life-limiting conditions in geothermally influenced hypersaline lakes at the Danakil Depression, Ethiopia. Environ Microbiol. 2021; 23(11):7168-7182. doi: 10.1111/1462-2920.15771. PMID: 34519149  Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, Hu B, Chain PSG, Nayfach S, Kyrpides NC. Identification of mobile genetic elements with geNomad. Nat Biotechnol. 2024; 42(8):1303-1312. doi: 10.1038/s41587-023-01953-y. PMID: 37735266  Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009; 25(15):1972-3. doi: 10.1093/bioinformatics/btp348. PMID: 19505945  Crits-Christoph A, Gelsinger DR, Ma B, Wierzchos J, Ravel J, Davila A, Casero MC, DiRuggiero J. Functional interactions of archaea, bacteria and viruses in a hypersaline endolithic community. Environ Microbiol. 2016; 18(6):2064-77. doi: 10.1111/1462-2920.13259. PMID: 26914534  Edgar RC. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. Nat Commun. 2022; 13(1):6968. doi: 10.1038/s41467-022-34630-w. PMID: 36379955  Emerson JB, Thomas BC, Andrade K, Allen EE, Heidelberg KB, Banfield JF. Dynamic viral populations in hypersaline systems as revealed by metagenomic assembly. Appl Environ Microbiol. 2012; 78(17):6309-20. doi: 10.1128/AEM.01212-12. PMID: 22773627  Garcia-Heredia I, Martin-Cuadrado AB, Mojica FJ, Santos F, Mira A, Antón J, Rodriguez-Valera F. Reconstructing viral genomes from the environment using fosmid clones: the case of haloviruses. PLoS One. 2012; 7(3):e33802. doi: 10.1371/journal.pone.0033802. PMID: 22479446  Guo J, Bolduc B, Zayed AA, Varsani A, Dominguez-Huerta G, Delmont TO, Pratama AA, Gazitúa MC, Vik D, Sullivan MB, Roux S. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. Microbiome. 2021; 9(1):37. doi: 10.1186/s40168-020-00990-y. PMID: 33522966  Gutiérrez-Preciado A, Dede B, Baker BA, Eme L, Moreira D, López-García P. Extremely acidic proteomes and metabolic flexibility in bacteria and highly diversified archaea thriving in geothermal chaotropic brines. Nat Ecol Evol. 2024 Oct;8(10):1856-1869. doi: 10.1038/s41559-024-02505-6. PMID: 39134651  Liu Y, Demina TA, Roux S, Aiewsakun P, Kazlauskas D, Simmonds P, Prangishvili D, Oksanen HM, Krupovic M. Diversity, taxonomy, and evolution of archaeal viruses of the class *Caudoviricetes*. PLoS Biol. 2021; 19(11):e3001442. doi: 10.1371/journal.pbio.3001442. PMID: 34752450  Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020; 37(5):1530-1534. doi: 10.1093/molbev/msaa015. PMID: 32011700  Zhou Y, Gutiérrez-Preciado A, Liu Y, Moreira D, Yakimov MM, López-García P, Krupovic M. Nested parasitism in hypersaline environments: viruses and virus satellites of haloarchaea and their nanosized cellular symbionts. bioRxiv. 2025; doi: https://doi.org/10.1101/2025.02.15.638363.  Zhou Y, Wang Y, Prangishvili D, Krupovic M. Exploring the archaeal virosphere by metagenomics. Methods Mol Biol. 2024;2732:1-22. doi: 10.1007/978-1-0716-3515-5\_1. PMID: 38060114 |

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

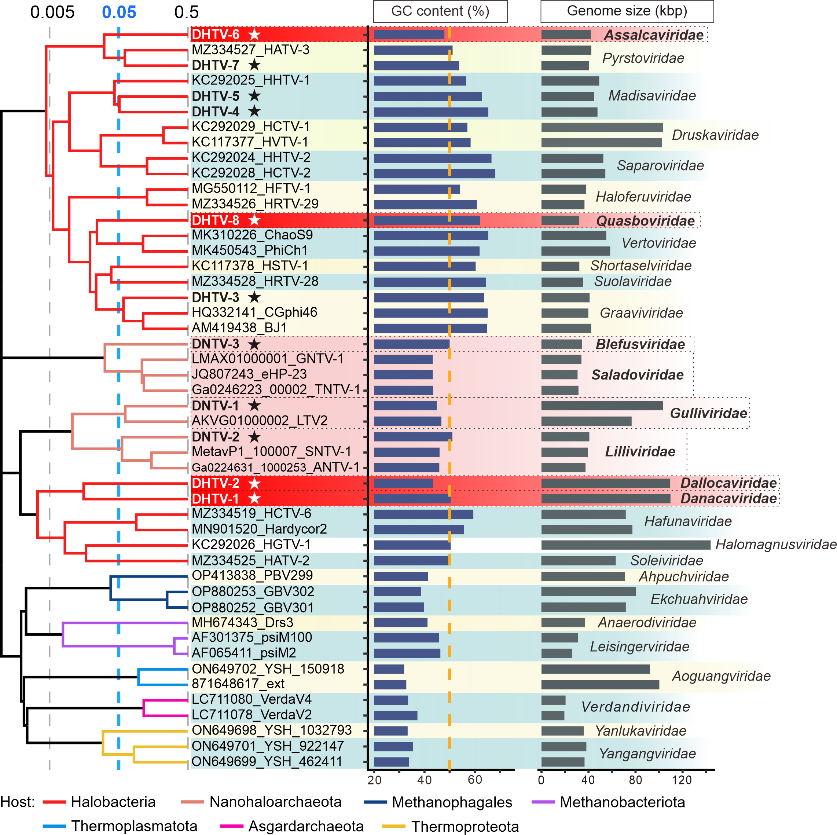
|  |
| --- |
| **Tables, Figures** |

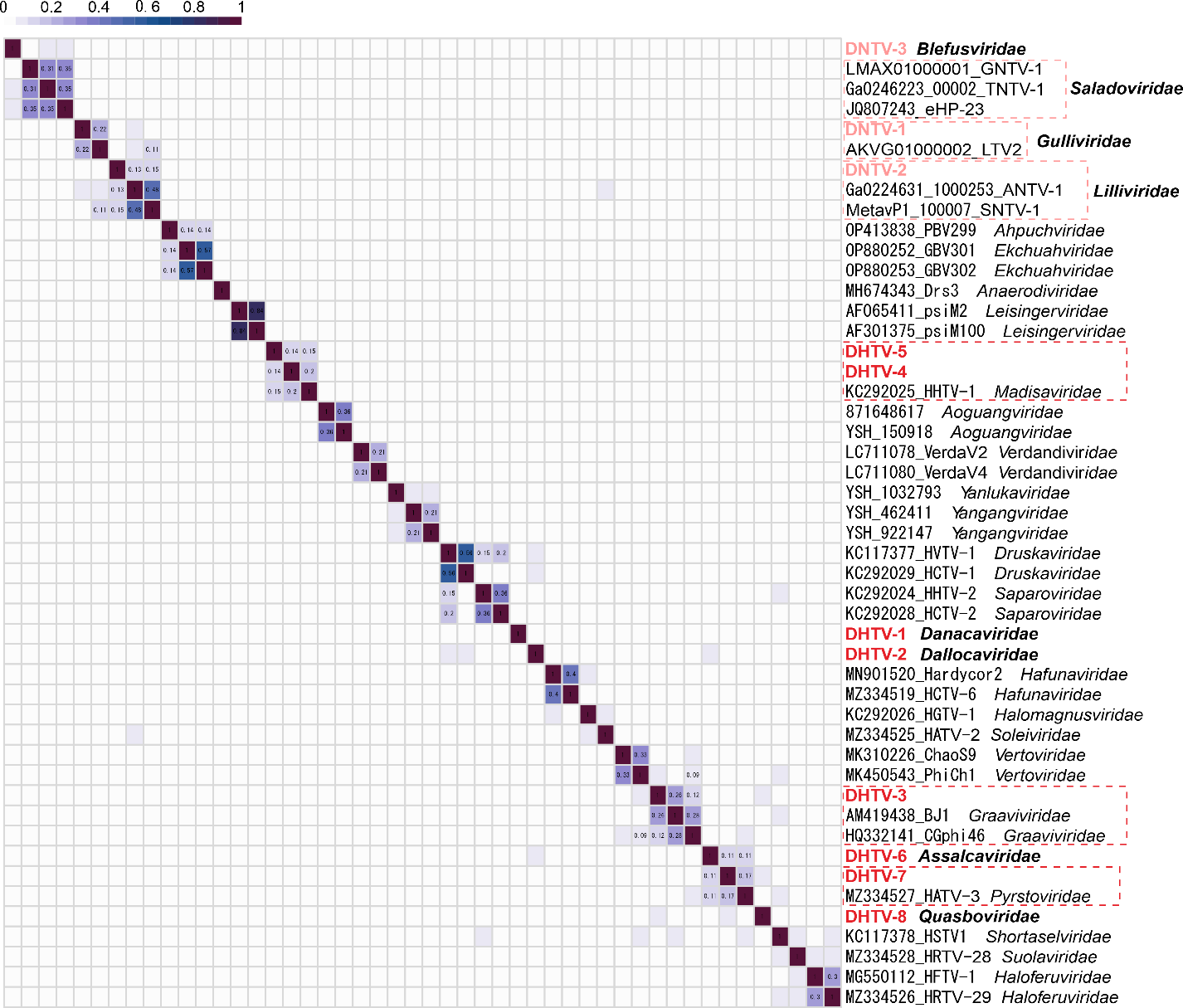
<Start here>

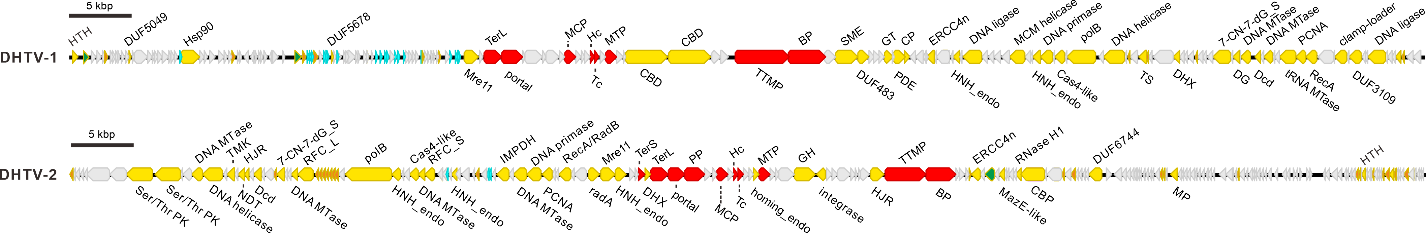
**Table 1.** Proposed taxonomy and classified viruses.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Family*** | ***Genus*** | ***Species*** | **Virus name** | **GenBank accession** | **Genome size, bp** |
| *Danacaviridae* | *Gablavirus* | *Gablavirus danakilense* | DHTV1 | PQ827553 | 109662 |
| *Dallocaviridae* | *Kalovirus* | *Kalovirus danakilense* | DHTV2 | PQ827554 | 109518 |
| *Assalcaviridae* | *Karumvirus* | *Karumvirus danakilense* | DHTV6 | PQ827558 | 41967 |
| *Quasboviridae* | *Cusbovirus* | *Cusbovirus danakilense* | DHTV8 | PQ827560 | 32100 |
| *Graaviviridae\** | *Ethicavirus* | *Ethicavirus danakilense* | DHTV3 | PQ827555 | 47583 |
| *Madisaviridae\** | *Morpovirus* | *Morpovirus danakilense* | DHTV4 | PQ827556 | 40947 |
| *Madisaviridae\** | *Traglyvirus* | *Traglyvirus danakilense* | DHTV5 | PQ827557 | 44928 |
| *Pyrstoviridae\** | *Haroovirus* | *Haroovirus danakilense* | DHTV7 | PQ827559 | 40430 |
| *Gulliviridae* | *Lemuelvirus* | *Lemuelvirus danakilense* | DNTV1 | PQ827565 | 103371 |
| *Gulliviridae* | *Latyvirus* | *Latyvirus* *nanohalovivens* | LTV2 | AKVG01000002 | 76707 |
| *Lilliviridae* | *Mildendovirus* | *Mildendovirus danakilense* | DNTV2 | PQ827566 | 40822 |
| *Blefusviridae* | *Wecalvirus* | *Wecalvirus danakilense* | DNTV3 | PQ827567 | 34579 |
| *Saladoviridae* | *Crypovirus* | *Crypovirus alicantense* | eHP-23 | JQ807243 | 30693 |
| *Saladoviridae* | *Crypovirus* | *Crypovirus chilense* | GNTV1 | LMAX01000001 | 34000 |

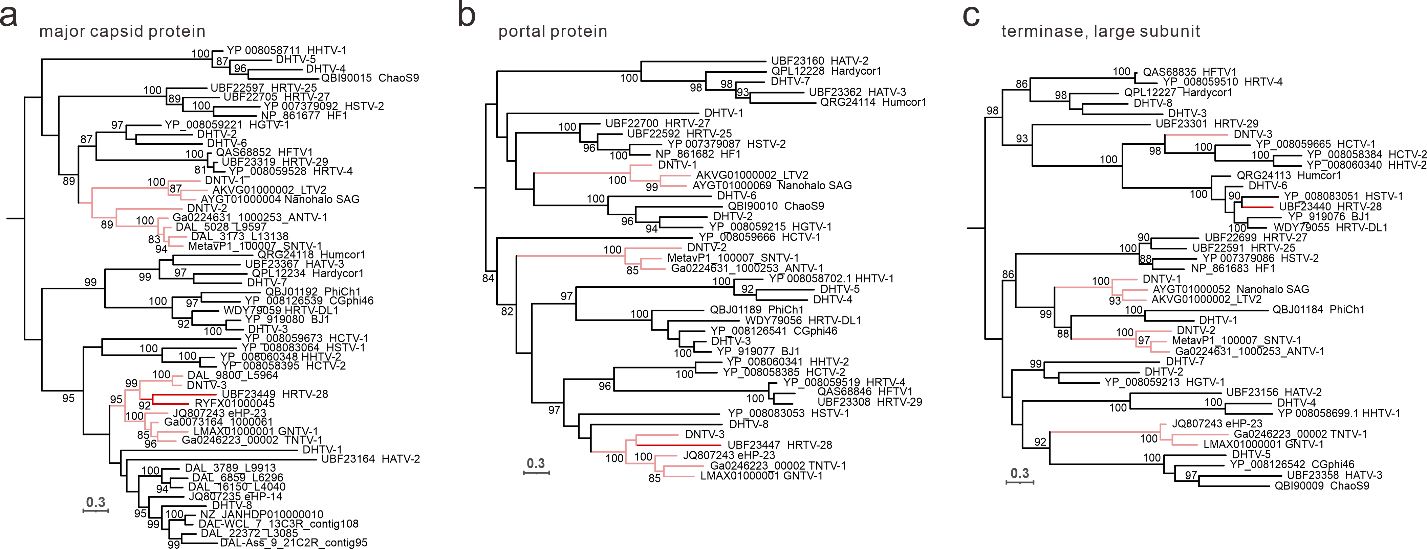
\* - existing families. DHTV, Danakil haloarchaeal virus; DNTV, Danakil nanohaloarchaeal virus; eHP-23, environmental halophage 23; GNTV1, Grande nanohaloarchaeal virus 1.

**Fig. 1.** Genome-wide proteomic tree of representative archaeal head-tailed viruses. The bars next to each genome indicate the GC% content and genome length, respectively. Viruses from the Danakil Depression are indicated with stars. Proposed new families are highlighted in red (HVs) or pink (NHVs). The proteomic trees are based on all-versus-all proteomic similarity matrix and are mid-point rooted. Branch lengths are log-scaled and the branch length for family-level demarcation is around 0.05. Branches are colored based on the viral host groups, with the key provided at the bottom of the figure.

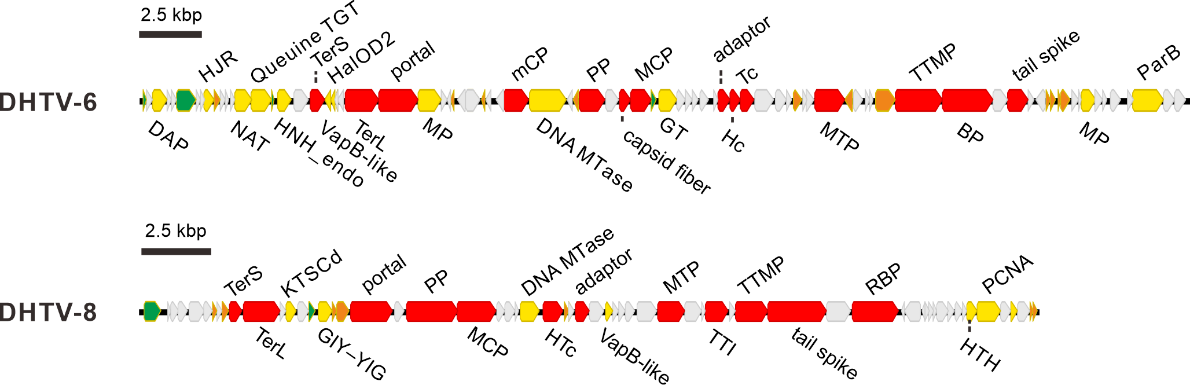
**Fig. 2.** The heatmap of orthologous fraction among archaeal tailed viruses. HVs and NHVs considered in this TaxoProp are shown in red. Family-level groups (values >0.08) of viruses including representatives from the Danakil Depression are boxed. The proportion of orthologous fraction between the viral genomes was estimated as previously described (Zhou et al., 2024).



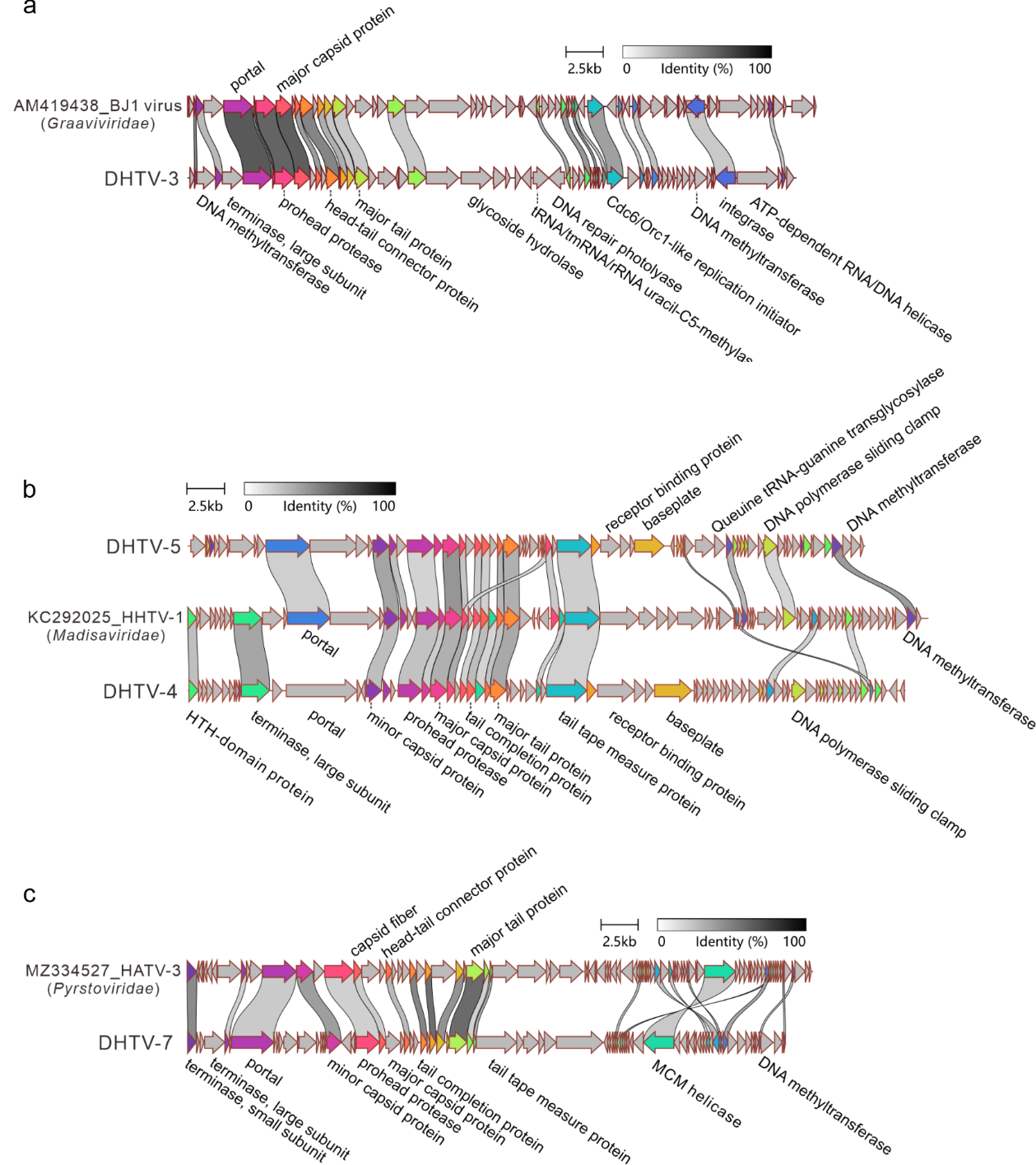
**Fig. 3.** Genome maps of DHTV1 and DHTV2. Abbreviations: HTH, helix-turn-helix; DUF, domain of unknown function; Hsp90, ATP-dependent chaperone Hsp90; Mre11, nuclease Mre11; TerS and TerL, small and large subunits of the terminase, respectively; MCP, major capsid protein; Hc, head-closure protein; Tc, tail-completion protein; MTP, major tail protein; CBD, carbohydrate binding domain; TTMP, tail tape measure protein; BP, baseplate protein; SME, sulfatase-maturating enzyme; GT; glycosyltransferase; PDE, phosphodiesterase; CP, cysteine protease; ERCC4n, ERCC4-type nuclease; HNH\_endo, HNH family endonuclease; polB, family B DNA polymerase; DHX, DEAD/DEAH-box helicase; 7-CN-7-dG\_S, 7-cyano-7-deazaguanine synthase; DNA MTase, DNA methyltransferase; DG, DNA glycosylase; Dcd, dCTP deaminase; tRNA MTase, tRNA methyltransferase; PCNA, DNA polymerase sliding clamp; RecA, RecA ATPase; Ser/Thr PK, serine/threonine protein kinase; TS, thymidylate synthase; TMK, thymidylate kinase; TK, thymidine kinase; HJR, Holliday junction resolvase; NDT, nucleoside 2-deoxyribosyltransferase; RCF\_L, replication factor C large subunit; RCF\_S, replication factor C small subunit; IMPDH, inosine-5'-monophosphate dehydrogenase; RecA/RadB-like, RecA/RadB-like recombination protein; RadA, DNA repair and recombination protein RadA; PP, prohead protease; homing\_endo, homing endonuclease; GH, glycoside hydrolase; RNase H1, ribonuclease HI; MazE-like, MazE-like antitoxin; CBP, cobalamin biosynthesis protein; MP, metalloprotease.



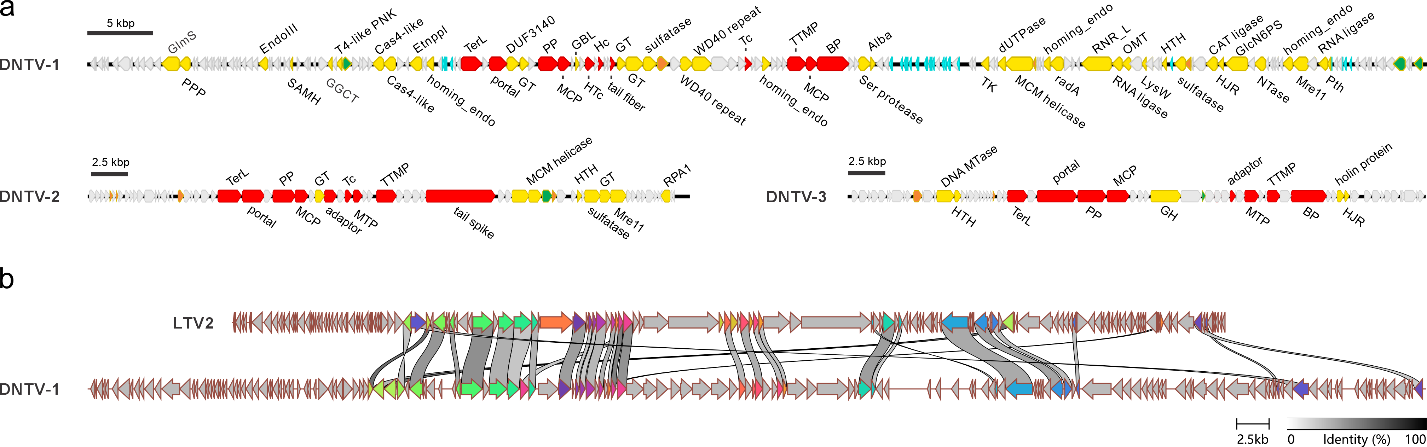
**Fig. 4.** Maximum likelihood phylogenies of the hallmark proteins of head-tailed viruses. a. Major capsid protein of head-tailed HVs and NHVs. b. Portal protein of head-tailed HVs and NHVs. c. Terminase large subunit of head-tailed HVs and NHVs. NHVs are indicated with pink branches. The sequences were aligned using Muscle5 with default parameters (Edgar, 2022), and non-informative columns were removed from the alignment using trimAl v1.2 with option -gt 0.2 (Capella-Gutiérrez et al., 2009). Next, a phylogenetic tree was constructed based on the trimmed alignment using IQ-TREE v2.2.2.2 with the following parameters: -m MFP, -alrt 1000 (Minh et al., 2020). The best fitting models for phylogenetic reconstructions were Blosum62+F+I+R3 (a), Q.pfam+F+I+G4 (b) and Q.pfam+F+R5 (c), respectively.



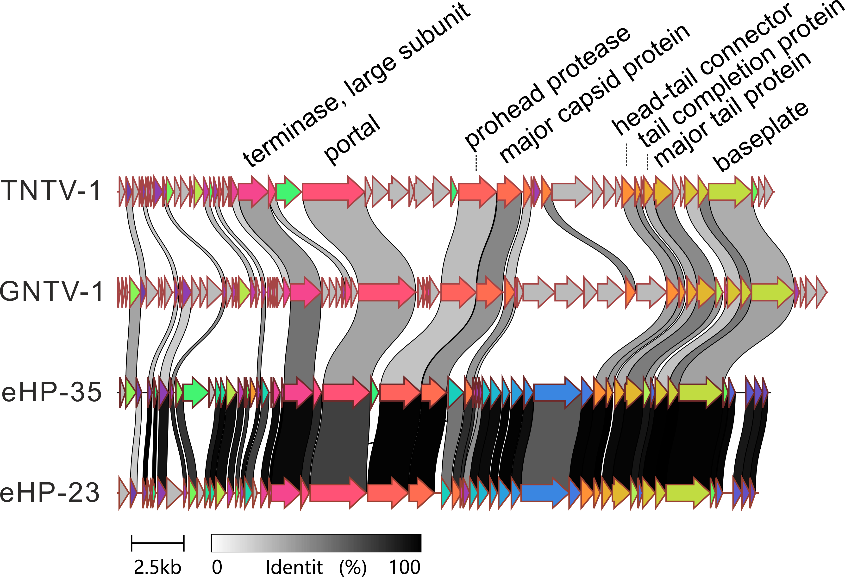
**Fig. 5.** Genome maps of DHTV6 and DHTV8. Abbreviations: DAP, DNA annealing protein; NAT, N-acetyltransferase; Queuine TGT, queuine tRNA-guanine transglycosylase; VapB-like, VapB-like antitoxin; HalOD2, haloarchaeal output domain 2; mCP, minor capsid protein; ParB, ParB family DNA-binding protein; KTSCd, lysine (K) tRNA synthetase C-terminal domain; GIY-YIG, GIY-YIG family nuclease; HTc, head-tail connector protein; TTI, tail tube initiator; RBP, receptor binding protein. Other abbreviations as in Fig. 3 legend.



**Fig. 6.** Genome maps showing the relationships among head-tailed HVs. a. *Graaviviridae*; b. *Madisaviridae*; c. *Pyrstoviridae*.



**Fig. 7.** Genome maps of members of the proposed families “*Gulliviridae*”, “*Lilliviridae*” and “*Blefusviridae*”. a. One genome per family is shown. b. Comparison between DNTV1 and LTV2. Abbreviations: GlmS, glucosamine 6-phosphate synthase; PPP, phosphotyrosine protein phosphatase; EndoIII, endonuclease III; SAMH, S-adenosyl-L-methionine hydrolase; GGCT, gamma-glutamyl cyclotransferase; T4-like PNK, T4-like polynucleotide kinase; Etnppl, ethanolamine phosphate transferase; GBL, galactose-binding lectin; WD40 repeat, WD40 repeat-containing protein; Alba, DNA/RNA-binding protein Alba; RNR\_L, ribonucleotide reductase large subunit; OMT, O-methyltransferase; LysW, amino group carrier protein LysW; GlcN6PS, glucosamine 6-phosphate synthetase; CAT ligase, carboxylate-amine/thiol ligase; Pth, peptidyl-tRNA hydrolase. Other abbreviations as in Fig. 3 legend.



**Fig. 8.** Genome maps showing the relationships among head-tailed NHVs proposed to be classified into family “*Saladoviridae*”.