Taxonomy proposal 2021: Create 3 new orders and 14 new families in the class Caudoviricetes (Duplodnaviria, Uroviricota) for classification of tailed archaeal viruses

Liu, Ying

International Committee on Taxonomy of Viruses (ICTV)
2022-03

Liu , Y , Demina , T , Roux , S , Aiewsakun , P , Kazlauskas , D M , Simmonds , P , Prangishvili , D , Oksanen , H M & Krupovic , M 2022 , Taxonomy proposal 2021: Create 3 new orders and 14 new families in the class Caudoviricetes (Duplodnaviria, Uroviricota) for classification of tailed archaeal viruses . in ICTV Online: International Committee on Taxonomy of Viruses (ICTV). . International Committee on Taxonomy of Viruses (ICTV) . < https://talk.ictvonline.org/files/ictv_official_taxonomy_updates_since_the_8th_report/m/archaeal-viruses/13417 >

http://hdl.handle.net/10138/343376

unspecified
publishedVersion

October 2020

Part 1: TITLE, AUTHORS, APPROVALS, etc

Code assigned:	**2021.001A**
Short title: Create three new orders and 14 new families in the class *Caudoviricetes* (*Duplodnaviria*, *Uroviricota*) for classification of archaeal tailed viruses	

Author(s) and email address(es)

Liu Y, Demina TA, Roux S, Aiewsakun P, Kazlauskas D, Simmonds P, Prangishvili D, Oksanen HM, Krupovic M	ying.liu@pasteur.fr; tatiana.demina@helsinki.fi; siroux1@gmail.com; pakorn.aie@mahidol.ac.th; kazlauskas.darius@gmail.com; peter.simmonds@ndm.ox.ac.uk; david.prangishvili@pasteur.fr; hanna.oksanen@helsinki.fi; mart.krupovic@pasteur.fr

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

Corresponding author

Mart Krupovic (mart.krupovic@pasteur.fr)

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

Archaeal Viruses Subcommittee

ICTV study group comments and response of proposer

Authority to use the name of a living person

Is any taxon name used here derived from that of a living person (Y/N)	Y

Taxon name	**Person from whom the name is derived**	**Permission attached (Y/N)**
Leisingerviridae	Thomas Leisinger	Y

Submission dates

Date first submitted to SC Chair	
Date of this revision (if different to above)	

ICTV-EC comments and response of the proposer

Part 3: TAXONOMIC PROPOSAL

Name of accompanying Excel module

2021.001A.R.Archaeal_Caudoviricetes.xlsx

Abstract

Here we propose creating 3 new orders and 14 new families for classification of archaeal tailed viruses, related to bacteriophages of the class *Caudoviricetes*. The new orders and families would be included in the class *Caudoviricetes*. In addition, we propose renaming six existing species to conform to the binomial species nomenclature format.

Text of proposal

Viruses with helical tails and icosahedral capsids (tailed viruses), classified into the class *Caudoviricetes* [7], represent the most widespread, abundant and diverse group of viruses on our planet [5]. Currently classified members of the *Caudoviricetes* nearly exclusively infect bacteria. Out of 2,814 species included in the class *Caudoviricetes,* only 9 correspond to archaeal viruses. The latter constitute 3 genera, *Haloferacalesvirus* (5 species), *Myohalovirus* (3 species) and *Psimunavirus* (1 species), which are included into the obsolete families *Myoviridae* and *Siphoviridae*, destined to be abolished in the near future. More than 60 archaeal tailed viruses (arTVs) have been isolated and genomically characterized, but the majority of them have not been classified. Here we propose to classify the 63 sequenced arTVs into 1 new order, 14 new families, 23 genera (20 new) and 32 species (23 new).

To assess the global diversity of arTVs and analyze their relationship to bacterial members of the class *Caudoviricetes*, we analyzed 63 complete arTV genomes [9] using GRAViTy [1, 2] and vConTACT v2.0 [4]. The GRAViTy tool classifies viruses into family-level taxonomic groupings according to homology between viral genes and similarities in genome organizations, which are expressed using composite generalized Jaccard (CGJ) distances [1, 2]. We used a CGJ distance of over 0.8 as the threshold for family-level assignment, consistent with the family-level classification for eukaryotic viruses and the recently created families of bacterial viruses [1-3]. GRAViTy analysis of the global prokaryotic virome classified arTVs into two large assemblages, which could be further subdivided into 14 family-level groupings (CGJ distance ≥ 0.8) (Fig. 1). To reveal a finer taxonomic structure within the arTV assemblage, we relied on the network analytics implemented in vConTACT v2.0, which has been specifically developed and calibrated to identify genus-level groupings of prokaryotic viruses [4]. Consistent with the GRAViTy results, the network analysis revealed two assemblages of arTVs, which were disconnected from all known bacteriophages and non-tailed archaeal viruses (Fig. 2). Viruses within the two clades formed 23 genus-level and 14 family-level groups, many containing just one or two members, indicating that genetic diversity of archaeal viruses remains largely undersampled.

Description of the proposed virus families

Clade I in the GRAViTy analysis forms a sister branch to several groups of bacterial myoviruses, including families *Ackermannviridae*, *Herelleviridae* and T4-like bacteriophages, and consists of four family-level groups (Fig. 1A). Among these, family (F) 1 is the largest, with 39 members, which can be further divided into four genus-level (G) subgroups (Fig. 2).

Family *Hafunaviridae*

F1G1 (virus cluster [VC] 77) consists of 31 viruses, including viruses HF1, HF2, HRTV-5, HRTV-8, Serpecor1 and Hardycor2 (Fig. 1B). Using the 95% nucleotide sequence identity threshold as a species demarcation criterion [8], the 31 viruses were assigned to seven species. Notably, HF1, HF2, HRTV-5, HRTV-7 and HRTV-8 are currently classified into separate species within a single genus, *Haloferacalesvirus*. Our analysis suggests that HF2 belongs to the same species with HF1, whereas HRTV-7 falls into a separate genus (F1G2). Thus, in addition to the 3 remaining species, we propose creating 4 new species in the genus *Haloferacalesvirus* (Table 1).

F1G2 (VC_80) consists of a single species, including six closely related viruses, namely, HRTV-2, HRTV-11, HCTV-6, HCTV-15, HSTV-2 and HRTV-7. F1G3 and F1G4 each contains a single virus, HRTV-25 and HRTV-27, respectively, which were identified as outliers tightly linked to the other F1 virus genomes (Fig. 2). The viruses from the four genera in F1 share the virion morphogenesis and genome replication modules. The morphogenetic module consists of 18 orthologous genes responsible for the formation of virions with myovirus morphology, whereas replication module includes family B DNA polymerase (PolB), archaeo-eukaryotic primase (AEP) and replicative minichromosome maintenance (MCM) helicase (Fig. 3A). The genus-specific genes constitute almost half of the gene contents.

For classification of the F1 group, we propose creating a new family and naming it *Hafunaviridae*, after HF1 virus, the first isolated representative of the family. F1G1 group corresponds to the existing genus *Haloferacalesvirus*. F1G2 is proposed to be named *Mincapvirus* referring to the minor capsid protein identified in the lattice of HSTV-2. Notably, we chose Halorubrum sodomense head-tail virus 2 as the representative virus of this species, rather than the previously classified Halorubrum head-tail virus 7, because the former has been more extensively studied and structurally characterized [10]. F1G3 is proposed to be named *Laminvirus*, referring to the gene encoding Laminin G, which is one

of the distinguishing features of HRTV-25 compared to other members of the family. F1G4 is proposed to be named *Minorvirus*, referring to the fact that HRTV-27 has the smallest genome in this virus family (Fig. 3A).

Family *Queuoviridae*

The F2 consists of 5 members divided into two genera, consistent with the network analysis (Fig. 1, Fig. 2). F2G1 (VC_226_1) includes HVTV-2, HVTV-1 and HCTV-5, whereas F2G2 (VC_226_0) includes HCTV-16 and HCTV-1. Viruses in the two genera contain largely syntenic genomes, but differ within the ~14 kb-long tail module (Fig. 3B). Notably, viruses from F2G1 (but not F2G2) share most of the tail genes with viruses from F11 (HCTV-2 and HHTV-2), despite the fact that F11 viruses have genomes half as long as those in F2. Most likely, the tail module has been acquired by the ancestor of HVTV-1 through recombination from an F11 member following its divergence from the common ancestor with the HCTV-1 (Fig. 3B). Consistent with this possibility, viruses from F2 and F11 infect *Haloarcula* strains.

For classification of the F2 group, we propose creating a new family and naming it *Queuoviridae*, referring to the distinguishing feature of viruses in this family to encode enzymes required for synthesis of queuosine. F2G1 is proposed to be named *Tredecimvirus*, from *tredecim* meaning 13 in Latin, referring to the T=13 symmetry of the HVTV-1 capsid [10]. F2G2 is proposed to be named *Hacavirus*, referring to the host (truncation of *Haloarcula californiae*).

Families *Soleiviridae* and *Halomagnusviridae*

HATV-2 and HGTV-1 are singletons in F3 and F4, respectively (Fig. 1). In the network analysis, the two genomes formed a single cluster, VC_66, albeit with a low topology confidence score (Fig. 2). However, given the difference in genome sizes (HGTV-1 is more than twice larger than HATV-2; Fig. 3C), different mechanisms of genome packaging (HATV-2 has direct terminal repeats, whereas HGTV-1 uses the headful mechanism), we propose classifying them into separate families. The two viruses share 12 genes (~4% and ~10% of the HGTV-1 and HATV-2 genomes, respectively), including those for several structural proteins and genome replication proteins (Fig. 3C).

For classification of HATV-2 (F3 group), we propose creating a new genus named *Eilatmyovirus*, referring to the source of HATV2 isolation and placing it into a new family *Soleiviridae* (truncation of solar (saltern) and Eilat, referring to the source of HATV-2 isolation).

For classification of HGTV-1 (F4 group), we propose creating a new genus named *Hagravirus*, referring to the host (***Ha****lo****gra****num*), and placing it into a new family *Halomagnusviridae* (*Halo* for halophilic + *magnus*, from Latin for large, referring to the fact that HGTV-1 has by far the largest genome among known haloarchaeal viruses).

Order *Thumleimavirales*

More generally, viruses forming Clade I in the GRAViTy analysis represent a cohesive assemblage held together by 34 protein clusters (PCs), including those involved in virion morphogenesis (major capsid protein, baseplate wedge J-like protein, tail sheath, prohead protease, etc.), genome replication (PolB and MCM), nucleotide metabolism/DNA repair (thymidylate synthase thyX, dUTPase, thymidylate kinase, DNA methyltransferase, ribonucleotide diphosphate reductase, dCTP tRNA splicing ligase RtcB, holiday junction resolvase), and several proteins of unclear functions, such as SprT-like metalloprotease, dual specificity protein phosphatase (DUSP), various nucleases and proteins with SPFH and ATPase domains. Thus, we propose unifying the four families, *Hafunaviridae*, *Queuoviridae*, *Soleiviridae* and *Halomagnusviridae*, into a new virus order *Thumleimavirales* (after Thumleima, the goddess and the female personification of the salt and the natural salt brines in Meitei mythology).

Clade II consists of 10 family-level groupings, half of which consist of singletons (Fig. 1A and 1C).

Order *Kirjokansivirales*: families *Haloferuviridae, Pyrstoviridae*, *Shortaselviridae* and *Suolaviridae*

F5 contains one genus (VC_238) which consists of three viruses: HRTV-29, HFTV1 and HRTV-4 (Fig. 1C, Fig. 2). Pairwise genomic comparison showed an overall similar genomic organization of the three viruses, including the genes in the morphogenesis module. The replication related genes, however, are distinct in these viruses: HFTV1 encodes MCM and PCNA, HRTV-29 has MCM only, whereas HRTV-4 has neither of the two (Fig. 3D). We thus propose to classify HFTV1, HRTV-29 and HRTV-4 into separate genera, namely, *Retbasiphovirus* (referring to the source of isolation, i.e., Lake Retba), *Dpdavirus* (referring to the *dpdA* gene specific to HRTV-29), *Saldibavirus* (truncation of Saline di Barletta, referring to the source of isolation), within a new family *Haloferuviridae* (referring to the hosts; truncation of *Haloferax* and *Halorubrum*).

Viruses HATV-3, HRTV-28 and HSTV-1 are singletons in the F6, F7 and F10, respectively. The three viruses share only a handful of genes with each other as well as with HFTV1 from *Haloferuviridae* (Fig. 3E).

For classification of HATV-3 (F6 group), we propose creating a new genus named *Hatrivirus* (name derived by transliteration of the abbreviated virus name), and placing it into a new family *Pyrstoviridae* (from Finnish *pyrstö*, for fish tail).

For classification of HRTV-28 (F7 group), we propose creating a new genus named *Lonfivirus* (referring to the long fibers of this virus), and placing it into a new family *Shortaselviridae* (referring to short tail and *sel* from French for salt).

For classification of HSTV-1 (F10 group), we propose creating a new genus named *Pormufvirus* (for fusion of the portal and Mu gpF proteins), and placing it into a new family *Suolaviridae* (from Finnish *suola*, for salt).

Viruses from the proposed families *Pyrstoviridae*, *Shortaselviridae* and *Suolaviridae* are connected to members of the *Haloferuviridae* as outliers in the network (Fig. 2) and share several genes involved in virion morphogenesis (TerS, portal, MCP and baseplate hub), genome replication (PCNA and MCM), and other functions (Rad52, HNH and GIY-YIG endonucleases etc.) (Fig. 3E). Thus, we propose unifying the four families within a new order *Kirjokansivirales* (referring to *Kirjokansi* [from Finnish *kirjo* for versatility or variety; from Finnish *kansi* for cover], a magical object in Finnish mythology Kalevala which brought riches and happiness including salt to its holder).

Order *Methanobavirales*: families *Leisingerviridae* and *Anaerodiviridae*

F13 and F14 include tailed viruses infecting methanogenic archaea. F13 includes virus psiM2, which is related to the previously reported defective provirus psiM100 (VC_287) (Fig. 2), but shows no appreciable sequence similarity to haloarchaeal tailed viruses. psiM2 is currently classified into a genus *Psimunavirus* with the obsolete family *Siphoviridae*.

We propose moving genus *Psimunavirus* (F13 group) into a new family *Leisingerviridae* (after Thomas Leisinger, who isolated the virus psiM2).

.

F14 also includes a single representative, virus Drs3. For classification of Drs3 (F14 group), we propose creating a new genus named *Metforvirus* (referring to the host; truncation of *Methanobacterium formicicum*), and placing it into a new family *Anaerodiviridae* (truncation of anaerobic digester, referring to source of isolation).

Viruses in the proposed families *Leisingerviridae* and *Anaerodiviridae* share genes encoding proteins responsible for capsid formation and genome packaging (TerL, portal, prohead protease and MCP; Fig. 3H). Thus, we propose unifying the two families into a new order *Methanobavirales* (name derived from truncation of Methanobacteriales, referring to the host of the viruses).

Unassigned families in the class *Caudoviricetes*

Family *Flexireviridae*

The F8 consists of two viruses, BJ1 and CGphi46, and they are sufficiently distinct to be classified into separate genera. Although the two viruses formed a single cluster (VC_255) in the network analysis (Fig. 2), comparative genomics revealed that more than half of the genes in the two viruses are unrelated (Fig. 3F). In the two viruses, all genes in the structural module are conserved, except for *terL*, while genes involved in their genome replication are distinct, with Orc1/Cdc6 and MCM encoded by BJ1, whereas a fusion protein of primase-helicase and PCNA encoded by CGphi46 (Fig. 3F).

For classification of the F8 group, we propose creating a new family and naming it *Flexireviridae* (truncation of 'flexible replication', referring to different genome replication modules in viruses of this family). BJ1 is proposed to be classified into a new genus *Beejeyvirus* (name derived from spelling of "BJ"). CGphi46 is proposed to be classified into a new genus *Seejivirus* (name derived from spelling of "CG").

Family *Vertoviridae*

The F9 contains two genera, with phiCh1 and phiH1 forming one genus, and ChaoS9 forming the other one. Although the three viruses formed a single cluster in the network analysis (Fig. 2), the genome alignments showed that ChaoS9 shares with the other two viruses only the tail morphogenesis module (~1/3 of the genome; Fig. 3G). Therefore, we propose extracting ChaoS9 from the genus *Myohalovirus* [6] and assigning it into a separate new genus, *Chaovirus* (referring to the name of the first virus representative in this genus).

For classification of the F9 group, we propose creating a new family and naming it *Vertoviridae* (from Latin *verto*, meaning invert, referring to the invertible tail fiber locus characterized in viruses of this family).

Family *Saparoviridae*

The F11 contains two viruses, HCTV-2 and HHTV-2, which share only the morphogenetic module and are thus proposed to be classified into two separate genera, *Samsavirus* (truncation of Samut Sakhon, referring to the source of isolation) and *Halohivirus* (referring to the host; truncation of *Haloarcula hispanica*). As mentioned above, the two viruses are connected to viruses in F2G1 (genus *Tredecimvirus*, family *Queuoviridae*) in the network analysis due to the shared tail proteins (Fig. 2, Fig. 3B).

For classification of the F11 group, we propose creating a new family and naming it *Saparoviridae* (from Finnish saparo, for pig tail).

Family *Madisaviridae*

F12 contains a previously described virus HHTV-1, which was identified as a singleton in the network analysis, although it shares several proteins (e.g., TerS, MCP, tail tape measure protein, PCNA, etc.) with other archaeal tailed viruses from different families.

For classification of HHTV-1 (F12 group), we propose creating a new genus named *Clampvirus* (referring to the DNA polymerase sliding clamp encoded by the virus), and placing it into a new family *Madisaviridae* (truncation of Margherita di Savoia – place of isolation).

The four families do not display closer affinity to each other or to other families within the class *Caudoviricetes*. Thus, for now, we propose to place them as unassigned families within the class *Caudoviricetes.*

The proposed taxonomy is summarized in Table 1.

Demarcation criteria
We propose using 95% sequence identity as a species demarcation criteria, to be consistent with the classification of related viruses infecting bacteria.

Members of the same proposed genus typically share more than 60% of their proteins and members of the same family share 20-50% of homologous proteins, whereas viruses from different families share less than 10% of proteins (Fig. 4).

Renaming existing species to binomial format
To conform to the mandated binomial species nomenclature, we propose renaming six existing species of arTVs. The proposed changes are listed in Table 2.

Supporting evidence

Table 1. Proposed taxonomy of archaeal tailed viruses.

Order/Family	Genus	Species	Virus isolate	Virus abbreviation	Morphotype	Genome length (bp)	Accession number
Order *Thumleimavirales* *Hafunaviridae* (F1)	*Haloferacalesvirus**	*Haloferacalesvirus HRTV8**	Halorubrum tailed virus 8	HRTV-8	myo	74519	KC292020
			Halorubrum tailed virus 14	HRTV-14	myo	74355	MZ334492
			Halorubrum tailed virus 17	HRTV-17	myo	74754	MZ334493
			Halorubrum tailed virus 19	HRTV-19	myo	77739	MZ334494
			Halorubrum tailed virus 23	HRTV-23	myo	77739	MZ334495
		Haloferacalesvirus HRTV10	Halorubrum tailed virus 10	HRTV-10	myo	76759	MZ334496
			Halorubrum tailed virus 18	HRTV-18	myo	76361	MZ334497
			Halorubrum tailed virus 20	HRTV-20	myo	76259	MZ334498
			Halorubrum tailed virus 22	HRTV-22	myo	76812	MZ334499
			Halorubrum tailed virus 26	HRTV-26	myo	77578	MZ334500
		Haloferacalesvirus HSTV4	Halorubrum sodomense tailed virus 4	HSTV-4	myo	75181	MZ334501
		*Haloferacalesvirus HRTV5**	Halorubrum tailed virus 5	HRTV-5	myo	76134	KC292022
			Haloarcula californiae tailed virus 7 (12)	HCTV-7	myo	76008	MZ334502
			Haloarcula californiae tailed virus 9	HCTV-9	myo	76008	MZ334503
			Haloarcula californiae tailed virus 11	HCTV-11	myo	76008	MZ334504
			Halorubrum tailed virus 9	HRTV-9	myo	75429	MZ334505
			Halorubrum tailed virus 16	HRTV-16	myo	77109	MZ334506
			Haloarcula californiae tailed virus 8	HCTV-8	myo	75019	MZ334507
			Haloarcula californiae tailed virus 10	HCTV-10	myo	75019	MZ334508

			Haloarcula japonica tailed virus 1	HJTV-1	myo	78012	MZ334509
			Halorubrum tailed virus 13	HRTV-13	myo	76666	MZ334510
			Halorubrum tailed virus 21	HRTV-21	myo	76556	MZ334511
		*Haloferacalesvirus HF1**	Halovirus HF1	HF1	myo	75898	AY190604
			Halorubrum virus HF2	HF2	myo	77670	AF222060
			Halorubrum coriense virus Hardycor2	Hardycor2	myo	77342	MN901520
		Haloferacalesvirus HJTV2	Halorubrum tailed virus 24	HRTV-24	myo	77537	MZ334512
			Haloarcula japonica tailed virus 2	HJTV-2	myo	76821	MZ334513
			Halorubrum sodomense tailed virus 3	HSTV-3	myo	76908	MZ334514
			Haloarcula japonica tailed virus 3	HJTV-3	myo	77353	MZ334515
			Halorubrum tailed virus 15	HRTV-15	myo	76242	MZ334516
		Haloferacalesvirus Serpecor1	Halorubrum coriense virus Serpecor1	Serpecor1	myo	74196	MN901521
	Mincapvirus	*Mincapvirus HSTV2*	Halorubrum sodomense tailed virus 2	HSTV-2	myo	68527	KC117376
			Halorubrum tailed virus 7	HRTV-7	myo	69048	KC292021
			Halorubrum tailed virus 2	HRTV-2	myo	68923	MZ334517
			Halorubrum tailed virus 11	HRTV-11	myo	71449	MZ334518
			Haloarcula californiae tailed virus 6 (13)	HCTV-6	myo	71672	MZ334519
			Haloarcula californiae tailed virus 15	HCTV-15	myo	71672	MZ334520
	Laminvirus	*Laminvirus HRTV25*	Halorubrum tailed virus 25	HRTV-25	myo	61934	MZ334521
	Minorvirus	*Minorvirus HRTV27*	Halorubrum tailed virus 27	HRTV-27	myo	56593	MZ334522
Queuoviridae (F2)	*Tredecimvirus*	*Tredecimvirus HVTV1*	Haloarcula vallismortis tailed virus 1	HVTV-1	sipho	102319	KC117377
			Haloarcula vallismortis tailed virus 2	HVTV-2	sipho	102319	MZ334523
			Haloarcula californiae tailed virus 5	HCTV-5	sipho	102105	KC292027
	Hacavirus	*Hacavirus HCTV1*	Haloarcula californiae tailed virus 1	HCTV-1	sipho	103257	KC292029
			Haloarcula californiae tailed virus 16	HCTV-16	sipho	104681	MZ334524
Soleiviridae (F3)	*Eilatmyovirus*	*Eilatmyovirus HATV2*	Haloarcula tailed virus 2	HATV-2	myo	63301	MZ334525
Halomagnusviridae (F4)	*Hagravirus*	*Hagravirus HGTV1*	Halogranum tailed virus 1	HGTV-1	myo	143855	KC292026
Order *Kirjokansivirales*							
Haloferuviridae (F5)	*Retbasiphovirus*	*Retbasiphovirus HFTV1*	Haloferax tailed virus 1	HFTV1	sipho	38059	MG550112
	Dpdavirus	*Dpdavirus HRTV29*	Halorubrum tailed virus 29	HRTV-29	sipho	36603	MZ334526
	Saldibavirus	*Saldibavirus HRTV4*	Halorubrum tailed virus 4	HRTV-4	sipho	35722	KC292023
Pyrstoviridae (F6)	*Hatrivirus*	*Hatrivirus HATV3*	Haloarcula tailed virus 3	HATV-3	sipho	42293	MZ334527
Shortaselviridae (F7)	*Lonfivirus*	*Lonfivirus HSTV1*	Haloarcula sinaiiensis tailed virus 1	HSTV-1	podo	32189	KC117378
Flexireviridae (F8)	*Beejeyvirus*	*Beejeyvirus BJ1*	Archaeal BJ1 virus	BJ1	sipho	42271	AM419438
	Seejivirus	*Seejivirus CGphi46*	Halorubrum virus CGphi46	CGphi46		39784	HQ332141
Vertoviridae (F9)	*Myohalovirus**	*Myohalovirus phiCh1**	Natrialba phage PhiCh1	phiCh1	myo	58487	MK450543
		*Myohalovirus phiH**	Halobacterium virus phiH variant phiH1	phiH	myo	58072	MK002701
	Chaovirus	*Chaovirus chaoS9**	Halobacterium virus ChaoS9	ChaoS9	myo	55145	MK310226
Suolaviridae (F10)	*Pormufvirus*	*Pormufvirus HRTV28*	Halorubrum tailed virus 28	HRTV-28	sipho	35270	MZ334528
Saparoviridae (F11)	*Samsavirus*	*Samsavirus HCTV2*	Haloarcula californiae tailed virus 2	HCTV-2	sipho	54291	KC292028
	Halohivirus	*Halohivirus HHTV2*	Haloarcula hispanica tailed virus 2	HHTV-2	sipho	52643	KC292024
Madisaviridae (F12)	*Clampvirus*	*Clampvirus HHTV1*	Haloarcula hispanica tailed virus 1	HHTV-1	sipho	49107	KC292025

Order *Methanobavirales*							
Leisingerviridae (F13)	*Psimunavirus**	*Psimunavirus psiM2**	Methanobacterium virus psiM2	psiM2	sipho	26111	AF065411
Anaerodiviridae (F14)	*Metforvirus*	*Metforvirus drs3*	Methanobacterium virus Drs3	Drs3	sipho	37129	MH674343

* - existing, ICTV-approved taxa.

Table 2. Changes to the existing species names.

Genus	**Current species name**	**Proposed species name**
Haloferacalesvirus	*Haloferax virus HF1*	*Haloferacalesvirus HF1*
Haloferacalesvirus	*Halorubrum Tailed Virus 5*	*Haloferacalesvirus HRTV5*
Haloferacalesvirus	*Halorubrum Tailed Virus 8*	*Haloferacalesvirus HRTV8*
Chaovirus	*Halobacterium virus ChaoS9*	*Chaovirus chaoS9*
Myohalovirus	*Halobacterium virus phiH*	*Myohalovirus phiH*
Myohalovirus	*Natrialba virus PhiCh1*	*Myohalovirus phiCh1*

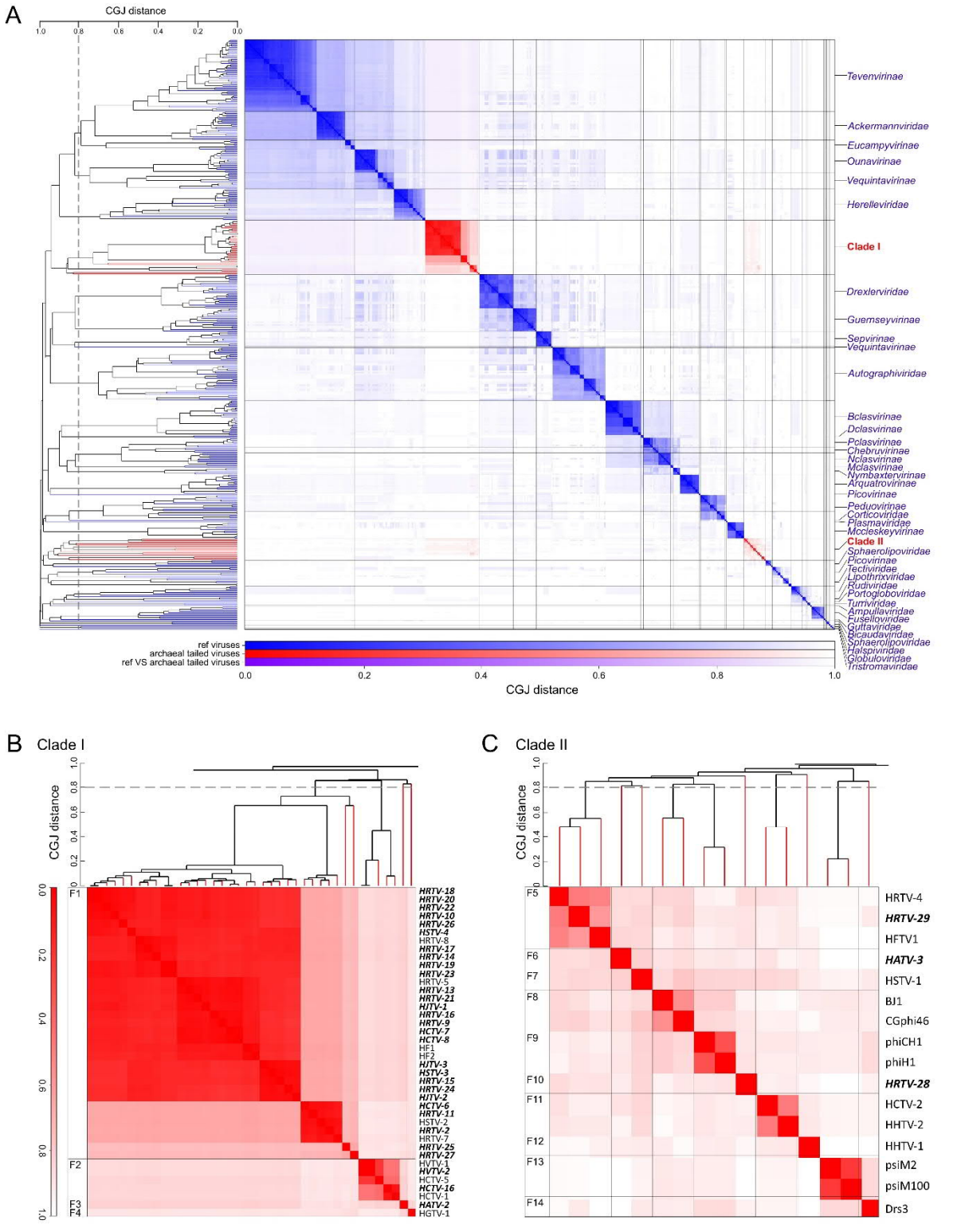


Fig. 1. Genome relationships between prokaryotic dsDNA viruses. (A) Heat map and dendrogram of composite generalized Jaccard (CGJ) distances for classified bacteriophages and archaeal dsDNA viruses. Branches and clusters corresponding to arTVs are shown in red, whereas those of other viruses are in blue. (B) Zoom in on the arTV Clade I. (C) Zoom in on the arTV Clade II. Viruses sequenced in this study are highlighted in bold. CGJ distance of 0.8, chosen as a family-level threshold, is indicated with a broken line, with the family-level groups (F1-F14) indicated on the left of the heatmap.

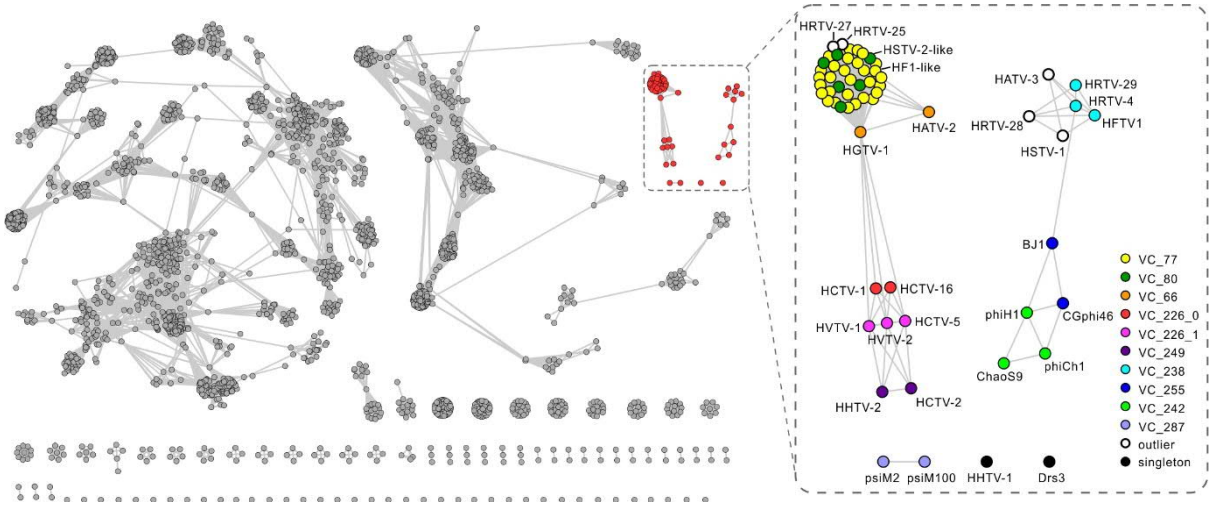
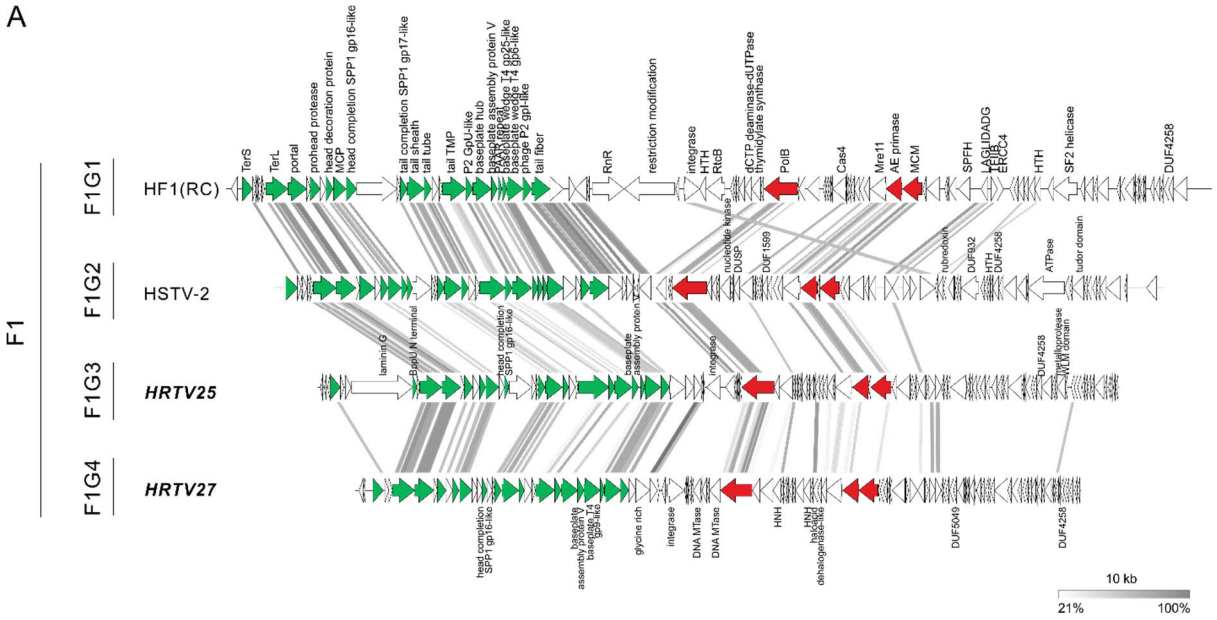
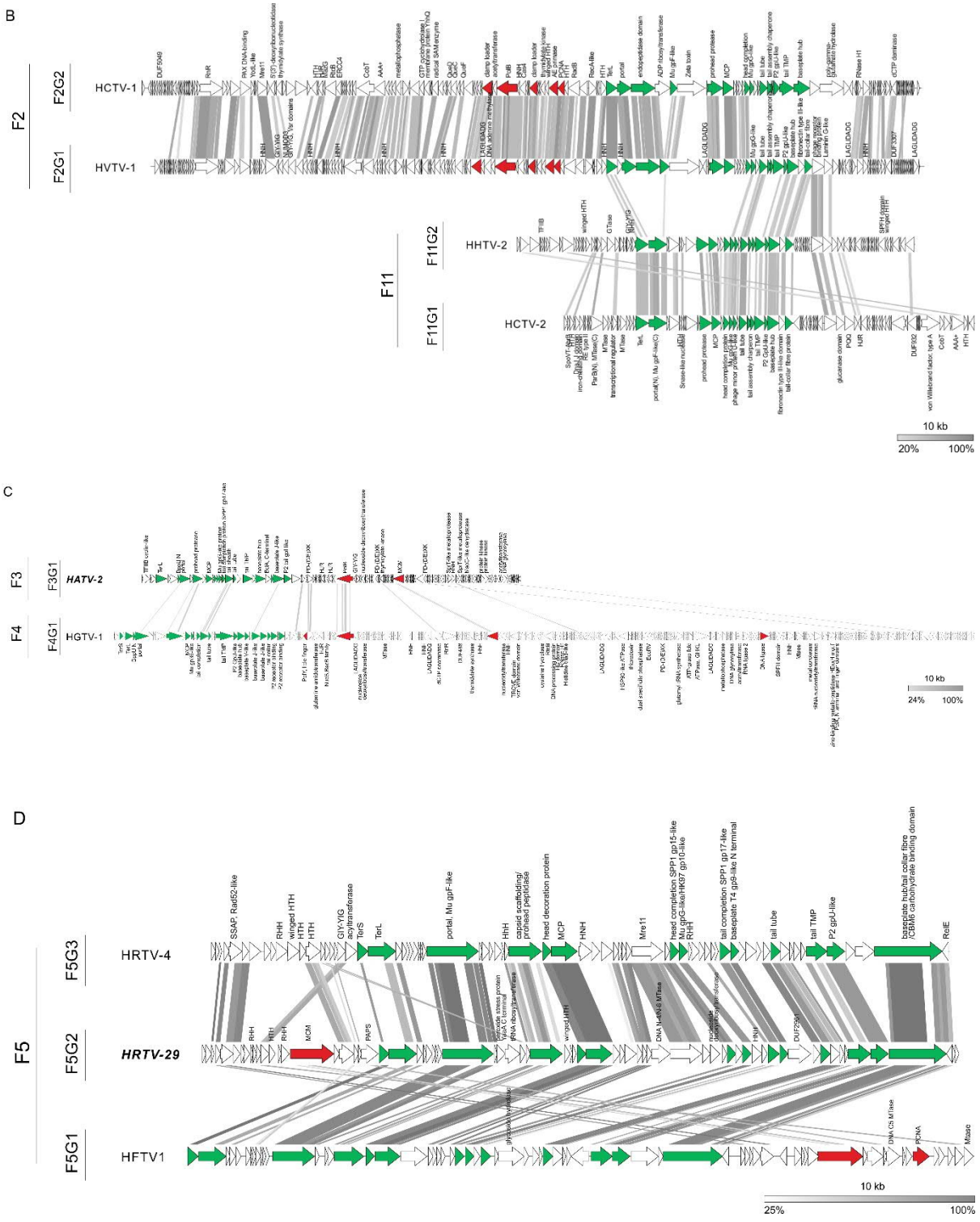
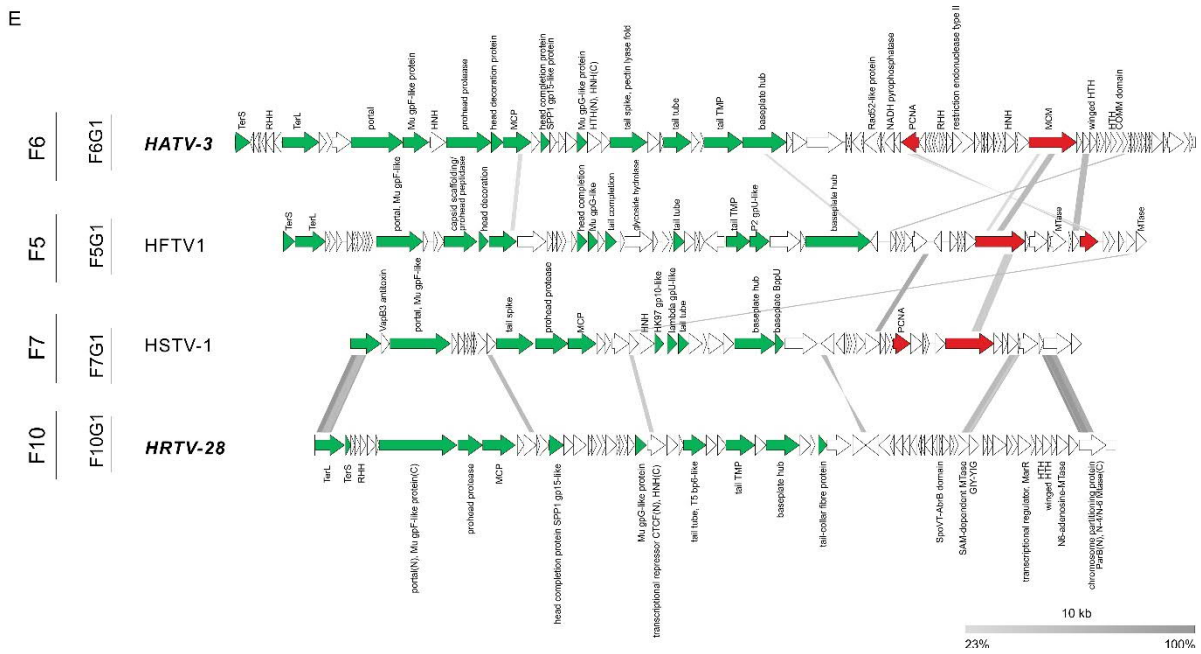
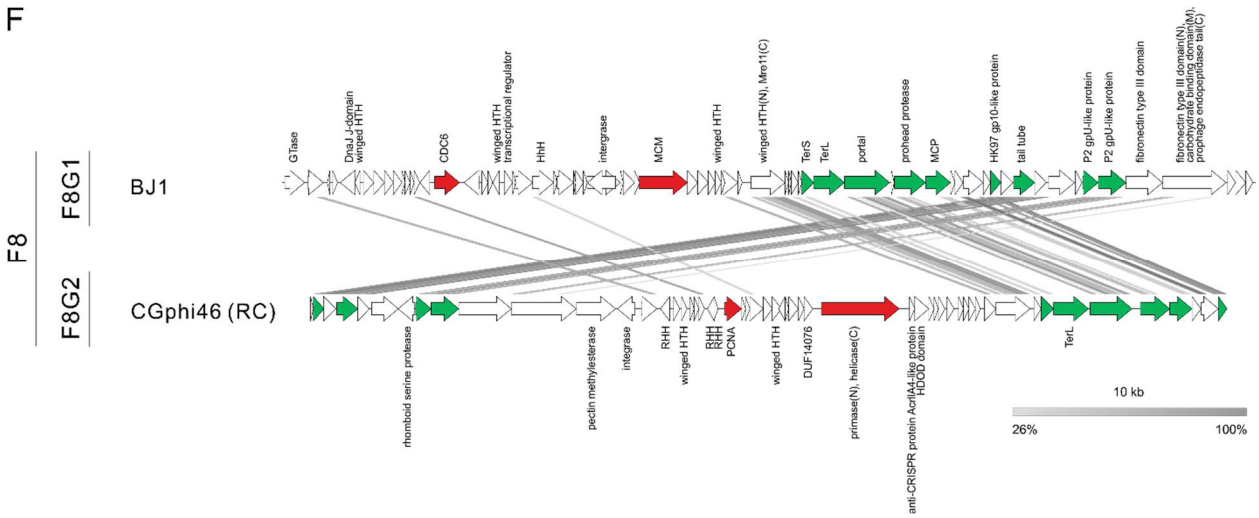


Fig. 2. The network-based analysis of shared protein clusters (PCs) among arTVs and the prokaryotic dsDNA viruses. The nodes represent viral genomes, and the edges represent the strength of connectivity between each genome based on shared PCs. Nodes representing genomes of arTVs are in red, whereas other dsDNA genomes are in grey (left panel). The viral clusters (VCs) of arTVs are enlarged and labeled in the right panel.




B
F2
F2G2
HCTV-1
F2G1
HVTV-1
F11
F11G2
HHTV-2
F11G1
HCTV-2
10 kb
20%
100%
C
F3
F3G1
HATV-2
F4
F4G1
HGTV-1
10 kb
24%
100%
D
F5
F5G3
HRTV-4
F5G2
HRTV-29
F5G1
HFTV1
10 kb
25%
100%


E
F6
F6G1
HATV-3
F5
F5G1
HFTV1
F7
F7G1
HSTV-1
F10
F10G1
HRTV-28
10 kb
23%
100%


F
F8
F8G1
BJ1
F8G2
CGphi46 (RC)
10 kb
26%
100%

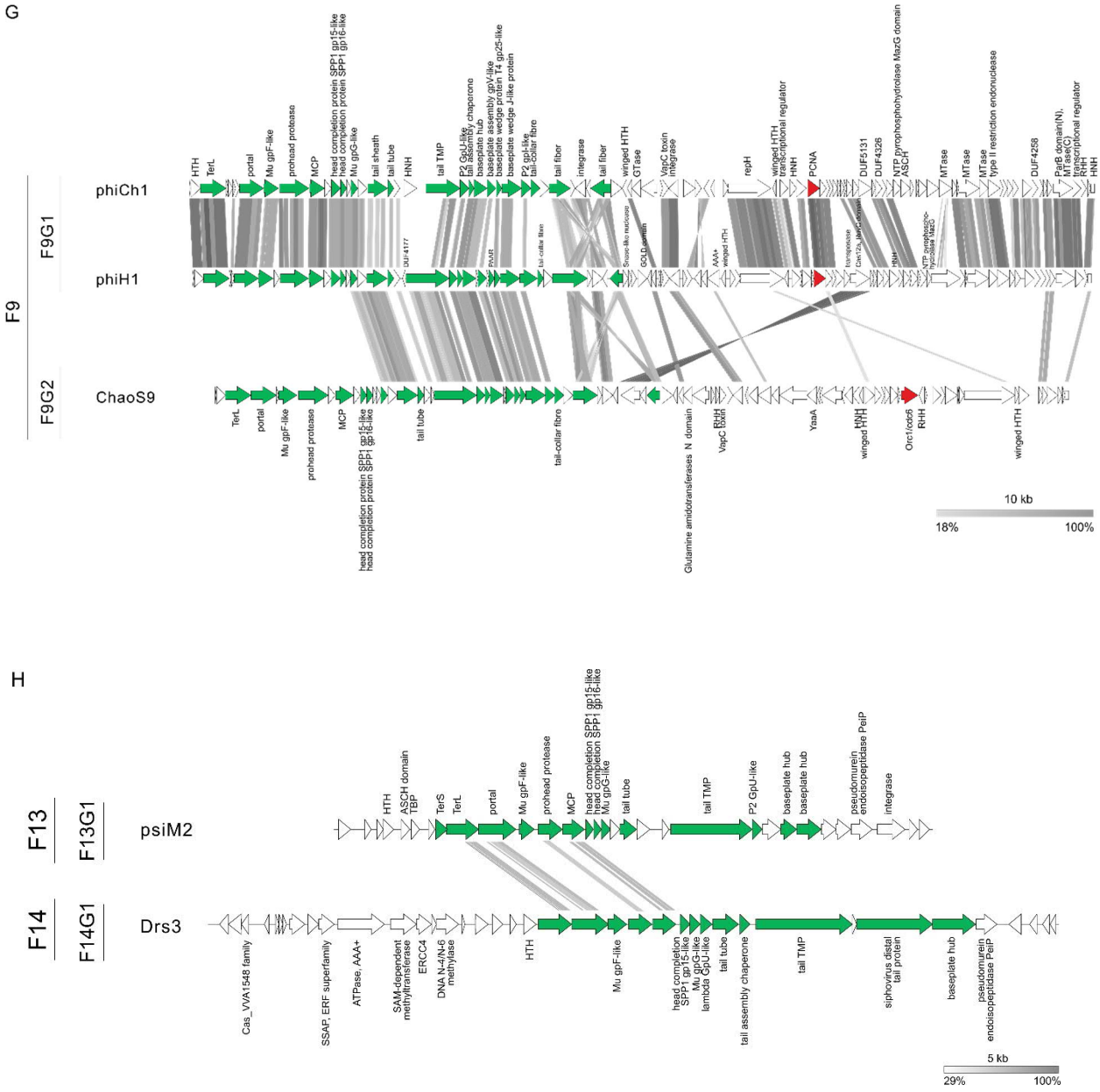


Fig. 3. Genome alignment of arTVs. Comparisons of genomes of (A) representative viruses from four genera in GRAViTy grouping family 1 (F1), (B) viruses from two genera in F2 and viruses from F11, (C) HATV-2 from F3 and HGTV-1 from F4, (D) viruses from F5, (E) HATV-3 from F6, HRTV-28 from F10 and HFTV1 from F5, (F) two members from F8, (G) three viruses from F9, (H) psiM2 from F13 and Drs3 from F14. Putative protein functions are indicated above or below the corresponding ORFs. Genes encoding virus morphogenesis related proteins are colored in green, whereas replication related genes are colored in red. Homologous genes shared between viruses are connected by shadings of different degrees of grey based on the aa-sequence identity. See Table 1 for more information on families (F) and genera (G).

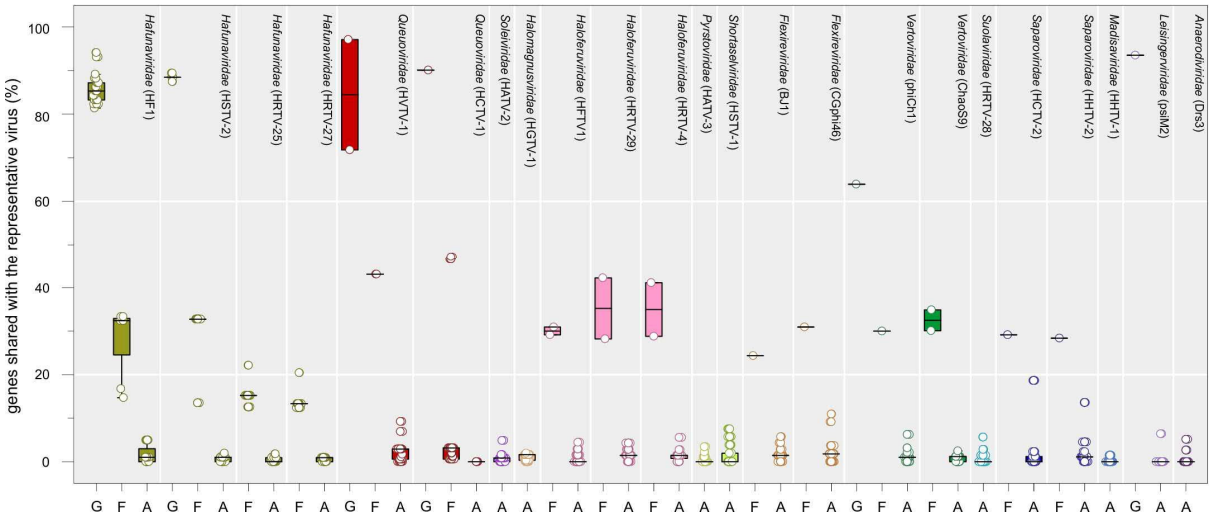


Fig. 4. The box plot shows the percentage of genes shared by arTVs. The percentage of genes of a representative virus from each genus that shared with members of the proposed genus (G) and family (F), as well as arTVs from other families (A) are shown. Each box represents the middle 50th percentile of the data set and is derived using the lower and upper quartile values. The median value is displayed by a horizontal line. Whiskers represent the maximum and minimum values with the range of 1.5 IQR. Each virus is represented by dots. Proteins with over 30% amino acid sequence identity and E-value < 1×10^{-25} in the local arTV database are counted as homologous proteins.

References

1. Aiewsakun P, Adriaenssens EM, Lavigne R, Kropinski AM, Simmonds P (2018) Evaluation of the genomic diversity of viruses infecting bacteria, archaea and eukaryotes using a common bioinformatic platform: steps towards a unified taxonomy. J Gen Virol 99:1331-1343. doi:10.1099/jgv.0.001110. PMID:30016225
2. Aiewsakun P, Simmonds P (2018) The genomic underpinnings of eukaryotic virus taxonomy: creating a sequence-based framework for family-level virus classification. Microbiome 6:38. doi:10.1186/s40168-018-0422-7. PMID:29458427
3. Barylski J, Enault F, Dutilh BE, Schuller MB, Edwards RA, Gillis A, Klumpp J, Knezevic P, Krupovic M, Kuhn JH, Lavigne R, Oksanen HM, Sullivan MB, Jang HB, Simmonds P, Aiewsakun P, Wittmann J, Tolstoy I, Brister JR, Kropinski AM, Adriaenssens EM (2020) Analysis of Spounaviruses as a Case Study for the Overdue Reclassification of Tailed Phages. Syst Biol 69:110-123. doi:10.1093/sysbio/syz036. PMID:31127947
4. Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM, Brister JR, Kropinski AM, Krupovic M, Lavigne R, Turner D, Sullivan MB (2019) Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat Biotechnol 37:632-639. doi:10.1038/s41587-019-0100-8. PMID:31061483
5. Dion MB, Oechslin F, Moineau S (2020) Phage diversity, genomics and phylogeny. Nature reviews Microbiology 18:125-138. doi:10.1038/s41579-019-0311-5. PMID:32015529

6. Dyall-Smith M, Palm P, Wanner G, Witte A, Oesterhelt D, Pfeiffer F (2019) Halobacterium salinarum virus ChaoS9, a Novel Halovirus Related to PhiH1 and PhiCh1. Genes (Basel) 10doi:10.3390/genes10030194. PMID:30832293
7. Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, Yutin N, Zerbini FM, Kuhn JH (2020) Global Organization and Proposed Megataxonomy of the Virus World. Microbiol Mol Biol Rev 84:e00061-19. doi:10.1128/MMBR.00061-19. PMID:32132243
8. Krupovic M, Dutilh BE, Adriaenssens EM, Wittmann J, Vogensen FK, Sullivan MB, Rumnieks J, Prangishvili D, Lavigne R, Kropinski AM, Klumpp J, Gillis A, Enault F, Edwards RA, Duffy S, Clokie MR, Barylski J, Ackermann HW, Kuhn JH (2016) Taxonomy of prokaryotic viruses: update from the ICTV bacterial and archaeal viruses subcommittee. Archives of virology 161:1095-9. doi:10.1007/s00705-015-2728-0. PMID:26733293
9. Liu Y, Demina TA, Roux S, Aiewsakun P, Kazkauskas D, Simmonds P, Prangishvili D, Oksanen HM, Krupovic M (2021) Diversity, taxonomy and evolution of archaeal viruses of the class *Caudoviricetes*. bioRxiv doi:10.1101/2021.05.26.445827.
10. Pietilä MK, Laurinmäki P, Russell DA, Ko CC, Jacobs-Sera D, Butcher SJ, Bamford DH, Hendrix RW (2013) Insights into head-tailed viruses infecting extremely halophilic archaea. Journal of virology 87:3248-60. doi:10.1128/JVI.03397-12. PMID:23283946