This form should be used for all taxonomic proposals. Please complete all those modules that are applicable (and then delete the unwanted sections). For guidance, see the notes written in blue and the separate document "Help with completing a taxonomic proposal"

Please try to keep related proposals within a single document; you can copy the modules to create more than one genus within a new family, for example.

MODULE 1: **TITLE, AUTHORS, etc**

Code assigned:	***2013.001a-oB***	(to be completed by ICTV officers)

Short title: Create the family *Sphaerolipoviridae*, comprising three new genera and 6 new species
(e.g. 6 new species in the genus *Zetavirus*)

Modules attached (modules 1 and 9 are required)	**1** ☒	**2** ☒	**3** ☒	**4** ☐	**5** ☐
	6 ☐	**7** ☐	**8** ☐	**9** ☒	

Author(s) with e-mail address(es) of the proposer:

Mike Dyall-Smith, mike.dyallsmith@gmail.com
Kate Porter, K.Porter@biota.com.au
Sen-Lin Tang, sltang@gate.sinica.edu.tw
Alice Pawlowski, alice.pawlowski@jyu.fi
Ilona Rissanen, ilona.a.rissanen@jyu.fi
Jaana K.H. Bamford, jaana.bamford@jyu.fi
Mart Krupovic, krupovic@pasteur.fr
Matti Jalasvuori, matti.jalasvuori@jyu.fi

List the ICTV study group(s) that have seen this proposal:

A list of study groups and contacts is provided at http://www.ictvonline.org/subcommittees.asp . If in doubt, contact the appropriate subcommittee chair (fungal, invertebrate, plant, prokaryote or vertebrate viruses)	Prokaryote Virus Subcommittee, Chair Rob Lavigne, rob.lavigne@biw.kuleuven.be; Archaeal viruses Study Group, Chair David Prangishvili, david.prangishvili@pasteur.fr

ICTV-EC or Study Group comments and response of the proposer:

Date first submitted to ICTV:
Date of this revision (if different to above): June 2014

MODULE 2a: **NEW SPECIES**

creating and naming one or more new species.

If more than one, they should be a group of related species belonging to the same genus. All new species must be placed in a higher taxon. This is usually a genus although it is also permissible for species to be "unassigned" within a subfamily or family. Wherever possible, provide sequence accession number(s) for one isolate of each new species proposed.

Code	***2013.001aB***	(assigned by ICTV officers)

To create 3 new species within:

Genus:	***Alphasphaerolipovirus* (new)**
Subfamily:	
Family:	***Sphaerolipoviridae* (new)**
Order:	

Fill in all that apply.

- If the higher taxon has yet to be created (in a later module, below) write "**(new)**" after its proposed name.
- If no genus is specified, enter "**unassigned**" in the genus box.

And name the new species:	**GenBank sequence accession number(s) of reference isolate:**
Haloarcula hispanica virus SH1 *Haloarcula hispanica virus PH1* *Haloarcula hispanica icosahedral virus 2*	SH1: AY950802 PH1: KC252997 HHIV-2: JN968479

Reasons to justify the creation and assignment of the new species:

- Explain how the proposed species differ(s) from all existing species.
  - If species demarcation criteria (see module 3) have previously been defined for the genus, **explain how the new species meet these criteria**.
  - If criteria for demarcating species need to be defined (because there will now be more than one species in the genus), please state the proposed criteria.
- Further material in support of this proposal may be presented in the Appendix, Module 9

These three viruses of *Haloarcula hispanica* (Family *Halobacteriaceae*, Phylum *Euryarchaeota*), share similar virus morphology including an internal membrane layer, genomic and protein sequences (particularly the major capsid protein genes), gene synteny, and genomic structure (dsDNA with terminal proteins). They differ from other described viruses that infect members of the *Halobacteriaceae*, or other *Archaea*, *Bacteria* or *Eukarya*.

MODULE 3a: **NEW GENUS -** ***Alphasphaerolipovirus***

creating a new genus

Ideally, a genus should be placed within a higher taxon.

Code	***2013.001bB***	(assigned by ICTV officers)

To create a new genus within:

Subfamily:	
Family:	***Sphaerolipoviridae* (new)**
Order:	

Fill in all that apply.

- If the higher taxon has yet to be created (in a later module, below) write "**(new)**" after its proposed name.
- If no family is specified, enter "**unassigned**" in the family box

naming a new genus

Code	***2013.001cB***	(assigned by ICTV officers)

To name the new genus: *Alphasphaerolipovirus*

Assigning the type species and other species to a new genus

Code	***2013.001dB***	(assigned by ICTV officers)

To designate the following as the type species of the new genus

Haloarcula hispanica virus SH1

Every genus must have a type species. This should be a well characterized species although not necessarily the first to be discovered

The new genus will also contain any other new species created and assigned to it (Module 2) and any that are being moved from elsewhere (Module 7b). **Please enter here the TOTAL number of species (including the type species) that the genus will contain:**

3

Reasons to justify the creation of a new genus:

Additional material in support of this proposal may be presented in the Appendix, Module 9

These viruses show strong similarity to each other in the sequences of their major capsid proteins, in their DNA genomes, and in their particle morphology, but they differ significantly from other described viruses in all of these properties. They have a similar particle structure, including an icosahedral protein capsid (50 to 80 nm) that contains an internal lipid membrane. The capsid geometry of one member (SH1) has been determined and found to be novel (T=28 *dextro*). Their genomes are linear dsDNA (ranging from approximately 28 to 31 kbp) and have terminal inverted repeats and the termini have attached proteins. (For more detail see the annex in module 9)

Origin of the new genus name:

Sphaerolipovirus; from the Latin *sphaero*, for "sphere", and the Greek *lipos*, for "fat".

Reasons to justify the choice of type species:

SH1 was the first described member of this genus

Species demarcation criteria in the new genus:

If there will be more than one species in the new genus, list the criteria being used for species demarcation and explain how the proposed members meet these criteria.

All members infect haloarchaeal hosts. Virions have isometric (icosahedral) capsids with internal lipid membranes, and contain dsDNA with inverted terminal repeat sequences and terminal

proteins. The capsid structure has been resolved for SH1, with a novel geometry of T=28 *dextro* (Jäälinoja et al., 2008). SH1, PH1 and HHIV-2 encode virus capsid proteins that are highly similar in sequence, particularly the major capsid proteins (see appendix). Species demarcation within this genus is based on nucleotide sequence similarity (80% threshold) of their genomes, and by differences in the amino acid sequence of their major capsid proteins. The three proposed species share less than 75% nucleotide similarity across their aligned genomes. Regarding the amino acid sequences of the two major capsid proteins, VP4 and VP7, all species show 14 or more, and 3 or more amino acid differences, respectively.

MODULE 2b: **NEW SPECIES**

creating and naming one or more new species.

If more than one, they should be a group of related species belonging to the same genus. All new species must be placed in a higher taxon. This is usually a genus although it is also permissible for species to be "unassigned" within a subfamily or family. Wherever possible, provide sequence accession number(s) for one isolate of each new species proposed.

Code	***2013.001eB***	(assigned by ICTV officers)

To create 1 new species within:

Genus:	***Betasphaerolipovirus* (new)**
Subfamily:	
Family:	***Sphaerolipoviridae* (new)**
Order:	

Fill in all that apply.

- If the higher taxon has yet to be created (in a later module, below) write "**(new)**" after its proposed name.
- If no genus is specified, enter "**unassigned**" in the genus box.

And name the new species:	**GenBank sequence accession number(s) of reference isolate:**
Natrinema virus SNJ1	AY048850

Reasons to justify the creation and assignment of the new species:

- Explain how the proposed species differ(s) from all existing species.
  - If species demarcation criteria (see module 3) have previously been defined for the genus, **explain how the new species meet these criteria**.
  - If criteria for demarcating species need to be defined (because there will now be more than one species in the genus), please state the proposed criteria.
- Further material in support of this proposal may be presented in the Appendix, Module 9

This virus of *Natrinema sp. J7-1* (Family *Halobacteriaceae*, Phylum *Euryarchaeota*), differs significantly from other known viruses, including its genome sequence and predicted proteins.

MODULE 3b: **NEW GENUS - *Betasphaerolipovirus***

creating a new genus

Ideally, a genus should be placed within a higher taxon.

Code	***2013.001fB***	(assigned by ICTV officers)

To create a new genus within:

Subfamily:	
Family:	***Sphaerolipoviridae* (new)**
Order:	

Fill in all that apply.

- If the higher taxon has yet to be created (in a later module, below) write "**(new)**" after its proposed name.
- If no family is specified, enter "**unassigned**" in the family box

naming a new genus

Code	***2013.001gB***	(assigned by ICTV officers)

To name the new genus: *Betasphaerolipovirus*

Assigning the type species and other species to a new genus

Code	***2013.001hB***	(assigned by ICTV officers)

To designate the following as the type species of the new genus

Natrinema virus SNJ1

Every genus must have a type species. This should be a well characterized species although not necessarily the first to be discovered

The new genus will also contain any other new species created and assigned to it (Module 2) and any that are being moved from elsewhere (Module 7b). **Please enter here the TOTAL number of species (including the type species) that the genus will contain:**

1

Reasons to justify the creation of a new genus:

Additional material in support of this proposal may be presented in the Appendix, Module 9

Natrinema virus SNJ1 infects the haloarchaeon, *Natrinema* sp. J7-1, and has virions that are similar in size and structure to *alphasphaerolipoviruses* i.e. are round, approximately 72 nm in diameter, and contain an internal lipid layer. The SNJ1 genome shares little nucleotide similarity with and is significantly smaller than the genomes of alphasphaerolipoviruses. The genome of SNJ1 is circular dsDNA, which differs from the linear dsDNAs of alphasphaerolipoviruses. The proteins specified by the SNJ1 genome show a weak but specific relationship to members of the genus *alphasphaerolipovirus*. See also appendix, Module 9

Origin of the new genus name:

Betasphaerolipovirus; from *beta*, second letter of the Greek alphabet, indicating second genus of this family; from the Latin *sphaero*, for "sphere", and the Greek *lipos*, for "fat".

Reasons to justify the choice of type species:

Natrinema virus SNJ1 is the first described member of this genus (Zhang et al., 2012)

Species demarcation criteria in the new genus:

If there will be more than one species in the new genus, list the criteria being used for species demarcation and explain how the proposed members meet these criteria.

Currently only one species described.

MODULE 2c: **NEW SPECIES**

creating and naming one or more new species.

If more than one, they should be a group of related species belonging to the same genus. All new species must be placed in a higher taxon. This is usually a genus although it is also permissible for species to be "unassigned" within a subfamily or family. Wherever possible, provide sequence accession number(s) for one isolate of each new species proposed.

Code	***2013.001iB***	(assigned by ICTV officers)

To create 2 new species within:

Genus:	***Gammasphaerolipovirus* (new)**
Subfamily:	
Family:	***Sphaerolipoviridae* (new)**
Order:	

Fill in all that apply.

- If the higher taxon has yet to be created (in a later module, below) write "**(new)**" after its proposed name.
- If no genus is specified, enter "**unassigned**" in the genus box.

Name of new species:	**GenBank sequence accession number(s)**
Thermus thermophilus phage P23-77 *Thermus thermophilus* phage IN93	P23-77: GQ403789 IN93: AB063393

Reasons to justify the creation and assignment of the new species:

- Explain how the proposed species differ(s) from all existing species.
  - If species demarcation criteria (see module 3) have previously been defined for the genus, **explain how the new species meet these criteria**.
  - If criteria for demarcating species need to be defined (because there will now be more than one species in the genus), please state the proposed criteria.
- Further material in support of this proposal may be presented in the Appendix, Module 9

The two viruses infecting extreme thermophilic *Thermus thermophilus* (Family *Thermaceae*, Order *Thermales*, Class *Deinococci*, Phylum *Deinococcus-Thermus*) share similar virion organization: the icosahedral protein capsid composed of two major capsid proteins covers the internal lipid membrane, which encloses the circular double-stranded DNA genome. Furthermore, the two viruses display extensive similarity and colinearity along their genome lengths. In their structural and genomic properties, P23-77 and IN93 differ from all other described viruses that infect Bacteria. However, the features described above are shared between the two viruses and members of the proposed family of archaeal viruses, the "Sphaerolipoviridae".

MODULE 3c: **NEW GENUS - *Gammasphaerolipovirus***

creating a new genus

Ideally, a genus should be placed within a higher taxon.

Code	***2013.001jB***	(assigned by ICTV officers)

To create a new genus within:

Subfamily:	
Family:	***Sphaerolipoviridae* (new)**
Order:	

Fill in all that apply.

- If the higher taxon has yet to be created (in a later module, below) write "**(new)**" after its proposed name.
- If no family is specified, enter "**unassigned**" in the family box

naming a new genus

Code	***2013.001kB***	(assigned by ICTV officers)

To name the new genus: *Gammasphaerolipovirus*

Assigning the type species and other species to a new genus

Code	***2013.001lB***	(assigned by ICTV officers)

To designate the following as the type species of the new genus

***Thermus thermophilus* bacteriophage P23-77**

Every genus must have a type species. This should be a well characterized species although not necessarily the first to be discovered

The new genus will also contain any other new species created and assigned to it (Module 2) and any that are being moved from elsewhere (Module 7b). **Please enter here the TOTAL number of species (including the type species) that the genus will contain:**

2

Reasons to justify the creation of a new genus:

Additional material in support of this proposal may be presented in the Appendix, Module 9

P23-77 was isolated from an alkaline hot spring on the North Island of New Zealand. The strictly lytic phage infects *Thermus thermophilus* ATCC 33923 and ATCC 27978 (Yu et al, 2006). IN93 was isolated from hot spring soil in Japan (Matsushita et al., 1995). It is a temperate phage that could be induced from lysogenic host, *T. thermophilus* TZ2. As P23-77, IN93 has a narrow host range, infecting only *T. thermophilus* HB8 besides its original host (Matsushita and Yanase, 2009). P23-77 and IN93 virus particles are spherical, tailless and have an average diameter of ~80 nm. An inner lipid membrane is located between the capsid and the circular dsDNA genome (Module 9, Annex, Figure 8). P23-77 shares 87% of its genes with IN93, yet 78% of the gene products lack similarity to any other protein sequences in public databases. The gene order is highly conserved in the two genomes (Module 9, Annex, Figure 9). The viral core proteins – the putative genome packaging ATPase and the small and large major capsid proteins – are among the most conserved proteins in P23-77 and IN93 with sequence identities of 79, 74 and 79%, respectively. The recently solved structure of the two major capsid proteins revealed a single beta-barrel core fold not found in any of the previously described bacterial viruses with dsDNA genomes (Module 9, Annex, Figure 10). With other members of the Sphaerolipoviridae gammasphaerolipoviruses share several features, including novel capsid geometry (T=28, dextro), capsomer structure and conserved block of viral core genes. Phylogenetic analysis of the three core proteins recapitulates the division of the

Sphaerolipoviridae into three genera (Module 9, Annex, Figure 1)

Origin of the new genus name:

Gammasphaerolipovirus; from *gamma*, third letter of the Greek alphabet, indicating the third genus of this family; from the Latin *sphaero*, for “sphere”, and the Greek *lipos*, for “fat”.

Reasons to justify the choice of type species:

P23-77 is designated as the type species of the suggested genus on the basis of comprehensive analysis of its genome, capsid architecture and high resolution structure of capsid proteins.

Species demarcation criteria in the new genus:

If there will be more than one species in the new genus, list the criteria being used for species demarcation and explain how the proposed members meet these criteria.

Species demarcation is based on genome size, gene content and sequence similarity. Genome sizes are 17,036 bp with 37 assigned ORFs for P23-77 and 19,604 bp with 43 assigned ORFs for IN93 with less than 60% similarity on nucleotide sequence level. The main difference is the presence of an integration cassette in the genome of IN93 required for the lysogenic cycle (Module 9, Annex, Figure 9).

MODULE 5: **NEW FAMILY**

creating and naming a new family

Code	***2013.001mB***	(assigned by ICTV officers)

To create a new family containing the subfamilies and/or genera listed below within the Order: *unassigned*

If there is no Order, write "**unassigned**" here.
If the Order has yet to be created (in Module 6) please write **"(new)"** after the proposed name.

Code	***2013.001nB***	(assigned by ICTV officers)

To name the new family: *Sphaerolipoviridae*

assigning subfamilies, genera and unassigned species to a new family

Code		(assigned by ICTV officers)

To assign the following subfamilies (if any) to the new family:
You may list several subfamilies here. For each subfamily, please state whether it is new or existing.

- If the subfamily is new, it must be created in Module 4
- If the subfamily already exists, please complete Module 7 to 'REMOVE' it from its existing family

Code	***2013.001oB***	(assigned by ICTV officers)

To assign the following genera to the new family:
You may list several genera here. For each genus, please state whether it is new or existing.

- If the genus is new, it must be created in Module 3
- If the genus already exists, please state whether it is currently unassigned or is to be removed from another family. If the latter, complete Module 7 to 'REMOVE' it from that family

Alphasphaerolipovirus

Betasphaerolipovirus

Gammasphaerolipovirus

The new family will also contain any other new species created and assigned to it (Module 3) and any that are being moved from elsewhere (Module 7b). **Please enter here the TOTAL number of unassigned species that the family will contain (those NOT within any of the genera or subfamilies listed above):**

Reasons to justify the creation of the new family:
Additional material in support of this proposal may be presented in the Appendix, Module 9

These viruses share similar particle morphology (icosahedral with an internal membrane layer) but differ significantly from other known viruses. Members fall into three, distantly related genera. The genome type can vary between genera, e.g. linear dsDNA (with terminal proteins) or circular dsDNA.

Origin of the new family name:

Sphaerolipoviridae, *sphaero* (Latin *sphaero*-, from Greek sphairo- for sphere); *lipo* (the Greek lipos, for "fat"); *viridae,* family level suffix.

References:

Jäälinoja HT, Roine E, Laurinmäki P, Kivelä HM, Bamford DH, Butcher SJ: Structure and host-cell interaction of SH1, a membrane-containing, halophilic euryarchaeal virus. Proc Natl Acad Sci U S A 2008, 105(23):8008-8013.

Jaakkola ST, Penttinen RK, Vilen ST, Jalasvuori M, Ronnholm G, Bamford JK, Bamford DH, Oksanen HM: Closely related archaeal *Haloarcula hispanica* icosahedral viruses HHIV-2 and SH1 have nonhomologous genes encoding host recognition functions. J Virol 2012, 86(9):4734-4742.

Jaatinen ST, Happonen LJ, Laurinmäki P, Butcher SJ, Bamford DH: Biochemical and structural characterisation of membrane-containing icosahedral dsDNA bacteriophages infecting thermophilic Thermus thermophilus. Virology 2008, 379: 10-19.

Jalasvuori M, Jaatinen ST, Laurinavičius S, Ahola-Iivarinen E, Kalkkinen N, Bamford DH, Bamford JK: The closest relatives of icosahedral viruses of thermophilic bacteria are among viruses and plasmids of the halophilic archaea. J Virol 2009, 83:9388-9397.

Jalasvuori M, Pawlowski A, Bamford JK: A unique group of virus-related, genome-integrating elements found solely in the bacterial family Thermaceae and the archaeal family Halobacteriaceae. J Bacteriol 2010, 192: 3231-3234

Kivela HM, Roine E, Kukkaro P, Laurinavicius S, Somerharju P, Bamford DH: Quantitative dissociation of archaeal virus SH1 reveals distinct capsid proteins and a lipid core. Virology 2006, 356(1-2):4-11.

Matsushita I, Yamashita N, Yokota A: Isolation and characterization of bacteriophage induced from a new isolate of Thermus aquaticus. Microbiol Cult Collect 1995, 11:133-138.

Matsushita I, Yanase H: The genomic structure of Thermus bacteriophage ϕIN93. J Biochem 2009, 146:775-785.

Pawlowski A, Rissanen I, Bamford JK, Krupovic M, Jalasvuori M: Gammasphaerolipovirus, a newly proposed bacteriophage genus, unifies viruses of halophilic archaea and thermophilic bacteria within the novel family Sphaerolipoviridae. Arch Virol 2014, 159(6):1541-54.

Porter K, Dyall-Smith ML: Transfection of haloarchaea by the DNAs of spindle and round haloviruses and the use of transposon mutagenesis to identify non-essential regions. Mol Microbiol 2008, 70(5):1236-1245.

Porter K, Kukkaro P, Bamford JK, Bath C, Kivelä HM, Dyall-Smith ML, Bamford DH: SH1: A novel, spherical halovirus isolated from an Australian hypersaline lake. Virology 2005, 335(1):22-33.

Porter K, Russ BE, Yang J, Dyall-Smith ML: The transcription programme of the protein-primed halovirus SH1. Microbiology 2008, 154(Pt 11):3599-3608.

Porter K, Tang S-L, Chen C-P, Chiang P-W, Hong M-J, Dyall-Smith ML: PH1, and archaeovirus of *Haloarcula hispanica* related to SH1 and HHIV-2. Archaea 2013, volume 2013, Article ID 456318.

Rissanen I, Grimes JM, Pawlowski A, Mäntynen S, Harlos K, Bamford JK, Stuart DI: Bacteriophage P23-77 capsid protein structures reveal the archetype of an ancient branch from a major virus lineage. Structure 2013, 21:718-726.

Yu M X, Slater MR, Ackermann HW: Isolation and characterization of Thermus bacteriophages. Arch Virol 2006, 151:663-679.

Zhang Z, Liu Y, Wang S, Yang D, Cheng Y, Hu J, Chen J, Mei Y, Shen P, Bamford DH, Chen X. 2012. Temperate membrane-containing halophilic archaeal virus SNJ1 has a circular dsDNA genome identical to that of plasmid pHH205. Virology 434:233-241.

Annex:

Include as much information as necessary to support the proposal, including diagrams comparing the old and new taxonomic orders. The use of Figures and Tables is strongly recommended but direct pasting of content from publications will require permission from the copyright holder together with appropriate acknowledgement as this proposal will be placed on a public web site. For phylogenetic analysis, try to provide a tree where branch length is related to genetic distance.

FAMILY: *Sphaerolipoviridae*

These viruses all share a similar particle structure, including an icosahedral protein capsid with an internal membrane. Virus particles show diameters ranging from 50 to 80 nm. The capsid geometry of one member (SH1) has been determined and found to be novel (T=28 *dextro*). They infect members of the Halobacteriaceae, and have genomes of dsDNA that are either linear or circular and range in size from around 16 to 31 kbp. Their major capsid proteins are not similar to other currently described viruses.

VIRUS	Genus[a]	Capsid (diameter)	Genome type	Genome size (bp)	Reference
SH1	*Alpha-*	Protein capsid of T=28 *dextro* geometry, with internal membrane layer, and large, horn-like spikes at vertices. Fragile outer layer. (D = 78 nm)	Linear dsDNA with inverted terminal repeats (309 bp) and terminal proteins.	30,898	Jäälinoja *et al*. (2008)
PH1	*Alpha-*	Round, with fragile outer layer. Probable internal membrane. (D~ 51 nm)	Linear dsDNA with inverted terminal repeats (337 bp) and terminal proteins.	28,072	Porter *et al*. (2013)
HHIV-2	*Alpha-*	Round, with probable internal membrane (D = 80 nm)	Linear dsDNA with inverted terminal repeats (309 bp) and terminal proteins.	30,578	Jaakkola *et al*. (2012)
SNJ1	*Beta-*	Round, with an internal membrane and a fragile outer layer. (D ~ 72 nm[b])	circular dsDNA	16,341	Zhang *et al*. (2012)
P23-77	*Gamma-*	Protein capsid of T=28 *dextro* geometry, with internal membrane layer, and stick-like spikes at vertices. (D = 78 nm)	Circular dsDNA	17,036	Jalasvuori *et al*. (2009)
IN93	*Gamma-*	Round, with probable internal membrane	Circular dsDNA	19,604	Matsushita and Yanase (2009)

[a]*Alpha- = Alphasphaerolipovirus*; *Beta- = Betasphaerolipovirus; Gamma- = Gammashaeorolipovirus.*
[b]Estimated from figure S3 of reference 8., Zhang et al. (2012), since no size values were stated in text of this paper.

Phylogenetic reconstruction based on the core gene products (small and large major capsid proteins [MCPs] and the packaging ATPase) of members of Sphaerolipoviridae and related proviruses found in the genomes of halophilic archaea and thermophilic bacteria (Jalasvuori et al., 2009, 2010, Porter et al., 2013) produced congruent trees, with the members of the proposed "Alpha-", "Beta-" and "Gammasphaerolipovirus" genera falling into three distinct, well-supported clades (Pawlowski et al., 2014). Based on the analysis of all three core proteins, haloarchaeal proviruses IHP, HaloMukP1 and

HaloMukP2 are related to betasphaerolipovirus SNJ1, whereas HalaPauP1 and HaloLacP1 are clearly related to alphasphaerolipoviruses. The proviruses identified in the Thermaceae genomes form a monophyletic clade with bacteriophages P23-77 and IN93.

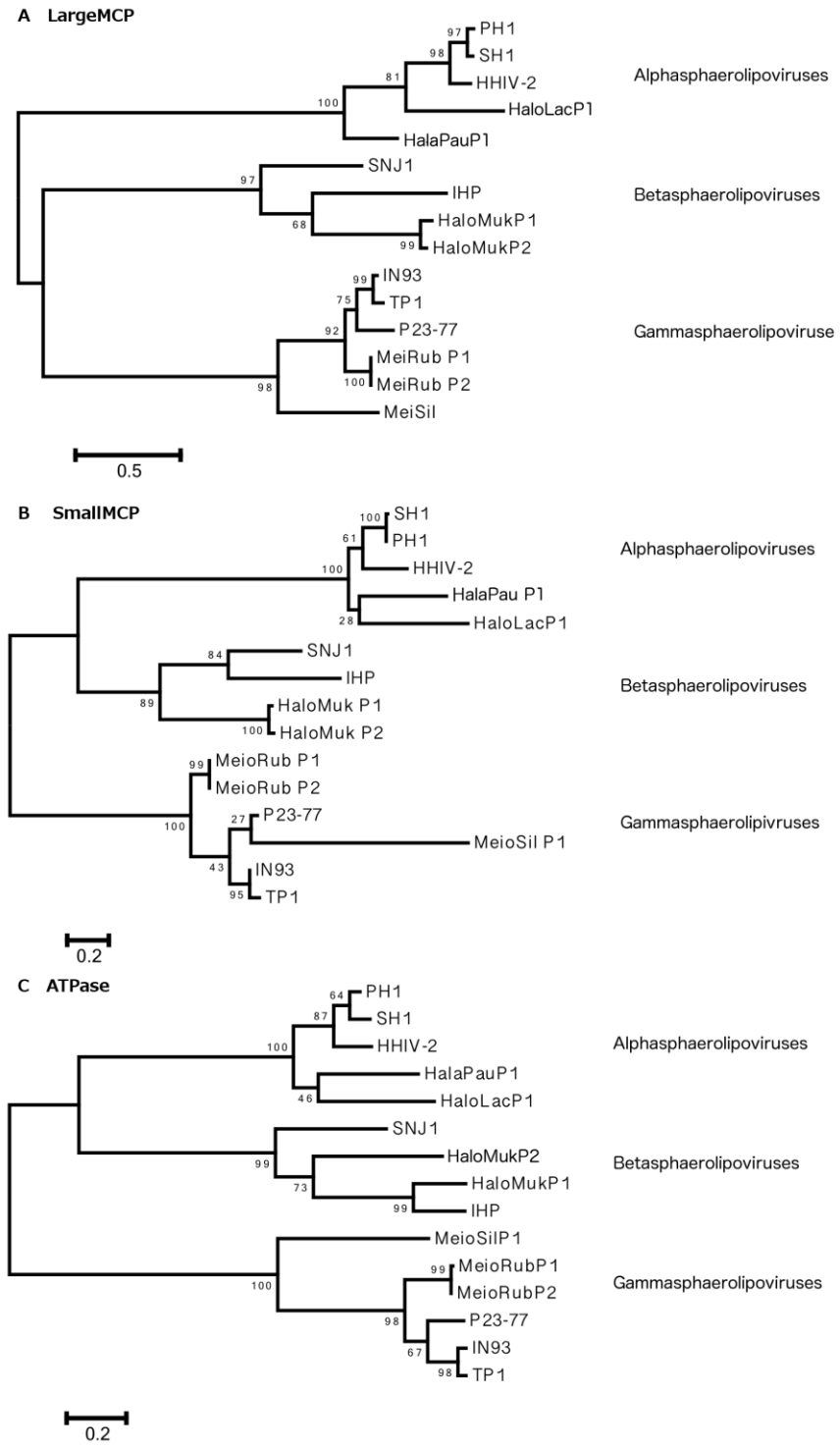


Figure 1: Molecular phylogenetic analysis of (A) large and (B) small major capsid protein and (C) ATPase sequences. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT amino acid substitution model. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The percentage of trees in which the associated taxa clustered together is shown next to the branches. All positions containing gaps and missing data were eliminated.

GENUS*: Alphasphaerolipovirus*

Alignments of the nucleotide sequences of the type species genome with the other two proposed species of this genus show high levels of identity (figures are given as % identity).

	PH1 genome (KC252997)	SH1 genome (NC007217)	HHIV-2 genome (JN968479)
PH1	-	72	54
SH1		-	59
HHIV-2			-

There is a high level of gene synteny between the three viruses. The figure below shows an alignment of HHIV-2, SH1 and PH1, where homologous genes are colour coded and many are named (e.g. VP1, VP2, etc.). It clearly shows the gene arrangement has been strongly conserved between the three viruses. (Figure from Porter et al. 2013).

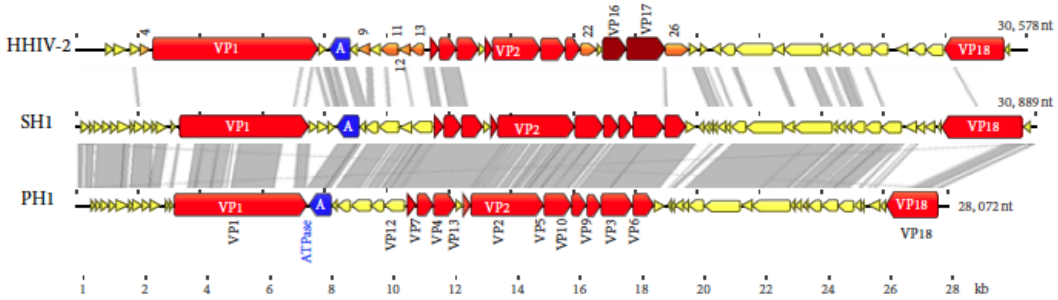


Figure 2. Genome alignment of the three members of the genus Alphasphaerolipovirus.

Alignment of the **major capsid protein** (**VP7**) sequences of alphasphaerolipoviruses SH1, PH1 and HHIV-2 show high levels of amino acid identity (figures given as % identity).

	PH1 VP7	SH1 VP7	HHIV-2 VP7
PH1	-	98.4	71.9
SH1	98.4	-	72.4
HHIV-2	71.9	72.4	-

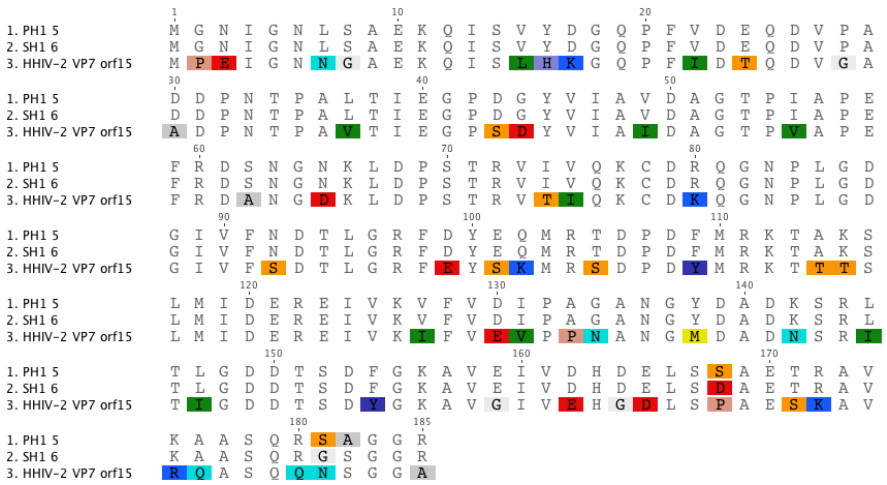


Figure 3. Alignment of major capsid protein sequences of alphasphaerolipoviruses PH1, SH1 and HHIV-2. Positions with varying amino acids are coloured.

Similarly high levels of protein sequence identity are shown by other capsid proteins of alphasphaerolipoviruses SH1, PH1 and HHIV-2, for example, capsid protein **VP4.**

	PH1 VP4	SH1 VP4	HHIV-2 VP4
PH1 VP4		94.0%	81.5%
SH1 VP4	94.0%		82.3%
HHIV-2 VP4	81.5%	82.3%	

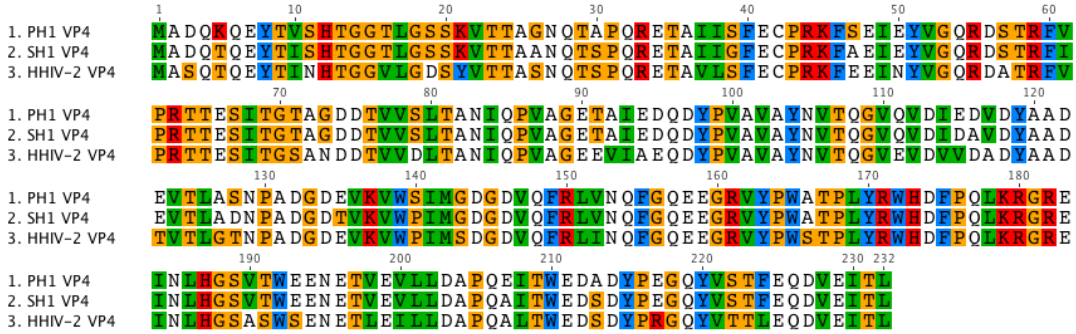


Figure 4. Amino acid sequence alignment of the VP4 proteins of alphasphaerolipoviruses SH1, PH1 and HHIV-2. Colour coding shows similar amino acids.

Virus morphology and structure

These viruses have similar morphology, being round, 50-80 nm in diameter, and having a layered capsid structure with an internal membrane (SH1: Jäälinoja HT *et al.* (2008), HHIV-2: Jaakkola *et al*. (2012), PH1: Porter et al. (2013)). The best described virus is SH1, where cryoelectronmicroscopy has provided a clear view of the particle structure and the nature of the capsid layers, including the lipid layer (see Jäälinoja HT et al. (2008)). The capsid membrane lipids of SH1 have been shown to be (selectively) acquired from the host cell (Jäälinoja HT *et al.* (2008), Kivela HM *et al.* (2006)). The protein capsid of SH1 is arranged in a T=28 *dextro* lattice. The major capsid proteins are VP4 and VP7, while protein proteins VP3 and VP6 form spikes at the 5-fold vertices (Jäälinoja HT et al. (2008) and references within).

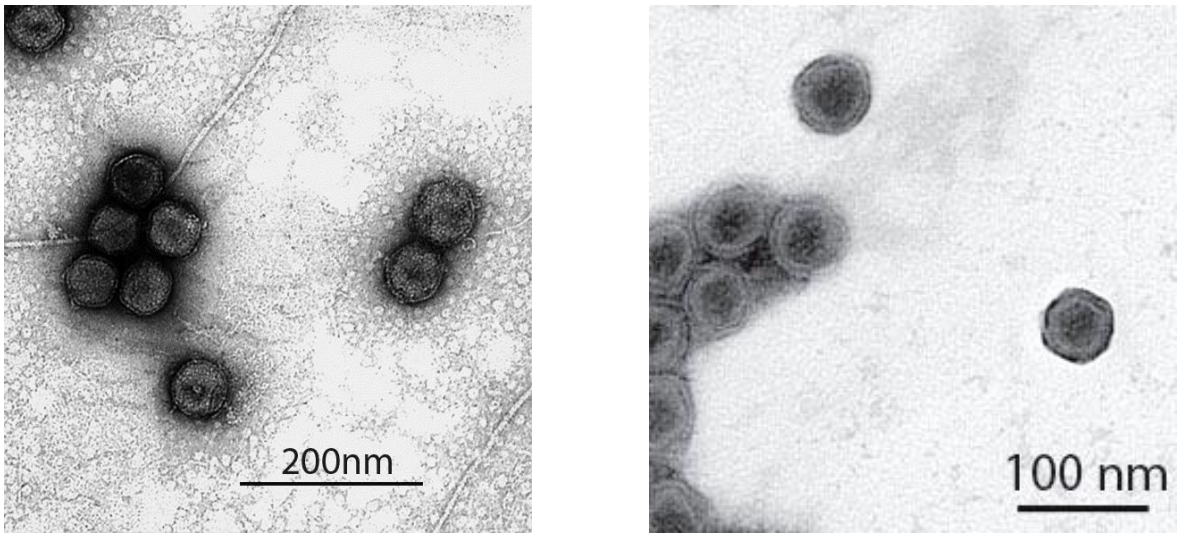


Figure 5. Negative stain EM (uranyl acetate) of SH1 virus (left) and PH1 (right). Particles are round, with a fragile outer-layer. (host cell flagella are also visible in SH1 preparation)

GENUS: *Betasphaerolipoviruses*

Apart from the virion structure described in Zhang et al. (2012), which clearly shows similarity to alphasphaerolipoviruses (SH1, PH1 and HHIV-1), additional support for the proposal of SNJ1 as a genus within the Sphaerolipoviridae is provided by the data described below.

1. Two genes of SNJ1 specify proteins that show similarity to corresponding proteins of alphasphaerolipoviruses. These are SNJ1 genes specifying ORF17 protein (**YP_001687802**) and the packaging ATPase (**YP_001687808**).

a. SNJ1 ORF17 Protein YP_001687802: A BLASTP search of Genbank using the SNJ1 ORF17 protein sequence (YP_001687802) and with the filter for viruses only (taxid:10239) gives only two significant matches: SH1 orf 11 (YP_271868.1) and HHIV-2 protein 4 (YP_005352790.1). The expect values were 10^{-6} and 10^{-9}, respectively. This shows a specific relationship to alphasphaerolipoviruses.

From the alignment data below, it can be seen that the level of similarity of the ORF17 SNJ1 protein sequence with alphasphaerolipovirus homologs is around 22 - 30%, while within the genus *Alphasphaerolipovirus* the level of similarity is much higher, around 60% or higher.

Alignment of SNJ1 ORF17 protein (YP_001687802) with *alphasphaerolipovirus* homologs. Pairwise similarity values are given in the table below, while the alignment (blacked-out letters are similar amino acids – except for N- and C- terminal extensions) is given beneath.

	PH1 ORF09	SH1 ORF11	HHIV-2 pr...	SNJ1 prote...
PH1 ORF09		86.5%	59.8%	31.0%
SH1 ORF11	86.5%		58.0%	29.1%
HHIV-2 protein 4	59.8%	58.0%		30.6%
SNJ1 protein 4	31.0%	29.1%	30.6%	

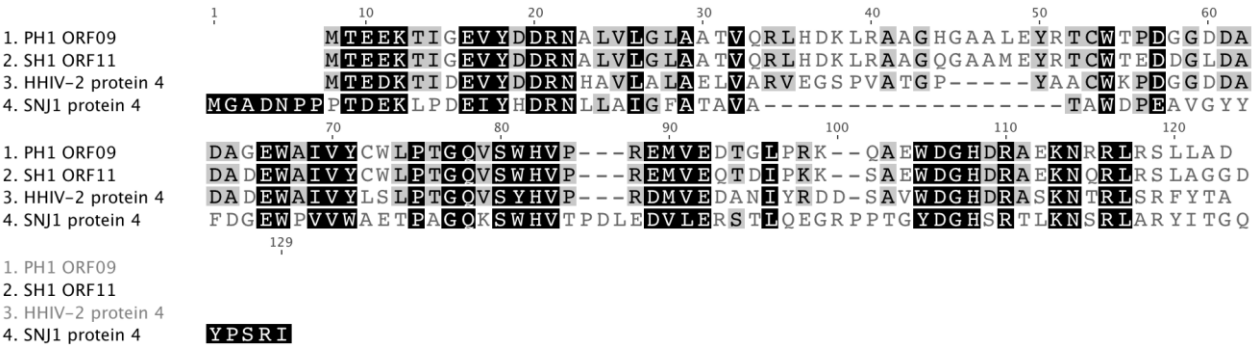


Figure 6. Alignment of SNJ1 ORF17 protein (YP_001687802) with alphasphaerolipovirus homologs.

b. Packaging ATPase protein (YP_001687808)
A BLASTP search at GenBank using the SNJ1 ATPase protein sequence (with the organism filter for virus viruses, taxid:10239) only gave two significant matches (April 3, 2013) i.e., to the packaging ATPase proteins of SH1 (orf17, YP_271874.1) and HHIV-2 (YP_005352793; HaHiIcV2_gp07). The expect values were 10^{-8} & 10^{-9}, respectively. This again demonstrates a specific relationship of this SNJ1 protein to alphasphaerolipoviruses.

The table below shows the pairwise similarity values for the four ATPase proteins. The CLUSTALW alignment is also given below, with similar amino acids indicated by black shading (except for the c-terminal extension of the SNJ1 protein)

	HHIV-2 At...	SNJ1 Atpase	SH1 Atpase	PH1 Atpase
HHIV-2 Atpase		21.2%	78.4%	81.7%
SNJ1 Atpase	21.2%		22.0%	22.0%
SH1 Atpase	78.4%	22.0%		91.7%
PH1 Atpase	81.7%	22.0%	91.7%	

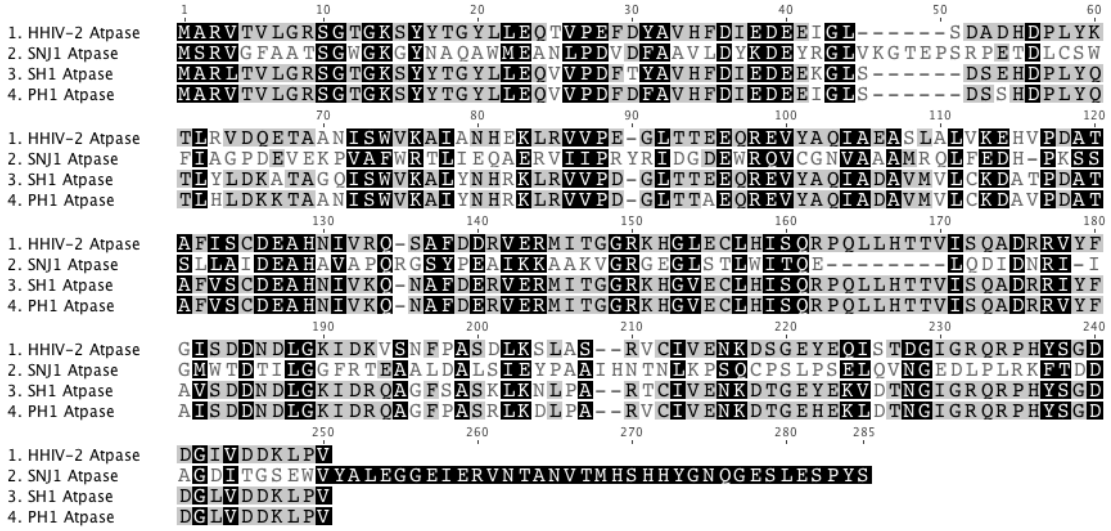


Figure 6. Alignment of SNJ1 putative packaging ATPase (YP_001687802) with alphasphaerolipovirus homologs.

2. Major Capsid Proteins

Purified virions of SNJ1 display two major capsid proteins, PB2 and PB6 (Zhang et al., 2012). These proteins are of similar size to the two major capsid proteins (VP4 and VP7) of alphasphaerolipoviruses.

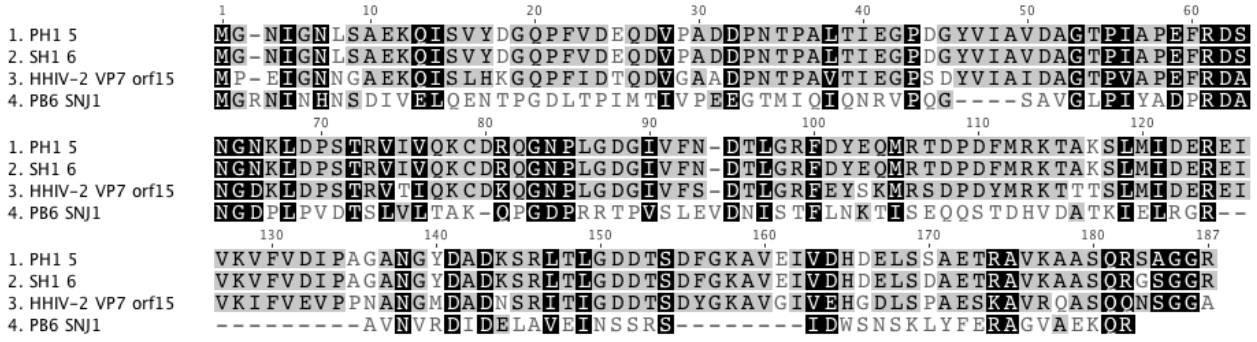
The major capsid protein PB6 (gene 26, YP_001687811) of SNJ1 can be aligned to corresponding VP7 (major capsid) proteins of alphasphaerolipoviruses (see part a, below), and while the overall similarity is low (see table below), the alignment below shows numerous conserved residues throughout the entire alignment, indicating a specific relationship between them.

The second most prominent capsid protein of SNJ1 is of similar length to VP4 of alphasphaerolipoviruses. The alignment presented in part b, below, shows there is much weaker similarity than is shown between the major capsid proteins in part a, but distinctive, conserved motifs throughout the alignment suggest these proteins are also (distantly) related.

Table: Similarity values (%) between the aligned major capsid (VP7-like) proteins of alpha- and beta-sphaerolipoviruses are shown in the table below.

	PH1 5	SH1 6	HHIV-2 V...	PB6 SNJ1
PH1 5		98.4%	71.9%	17.6%
SH1 6	98.4%		72.4%	17.6%
HHIV-2 VP7 orf15	71.9%	72.4%		14.8%
PB6 SNJ1	17.6%	17.6%	14.8%	

A.



B.

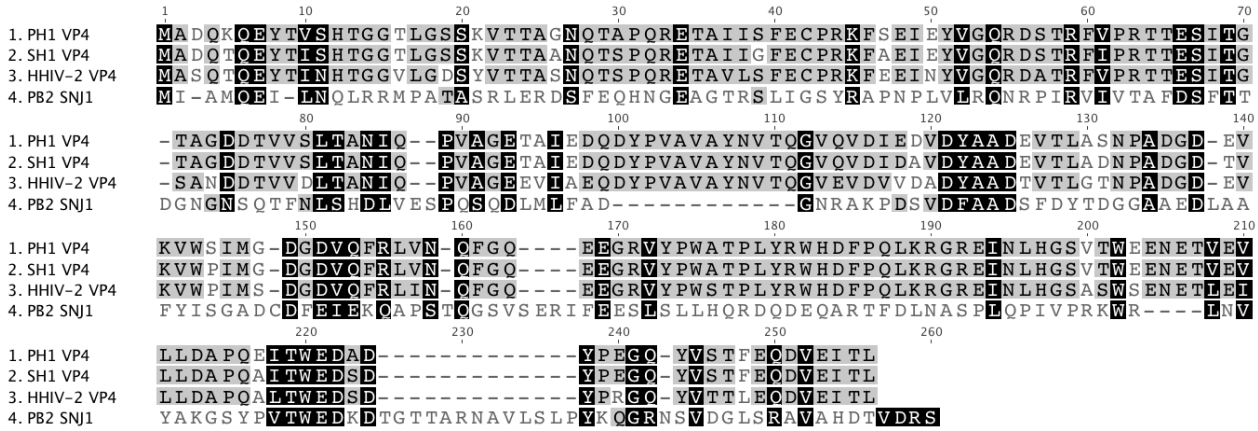


Figure 7. Multiple alignments of the major capsid protein sequences of SNJ1 with the alphasphaerolipovirus homologs. A. Major capsid protein VP7 of alphasphaerolipoviruses aligned with PB6 of SNJ1. Similar amino acids are blocked in black. The SNJ1 protein is slightly smaller (158 aa) than the alphasphaerolipovirus major capsid proteins (185 aa). B. Alignment of capsid protein PB2 of SNJ1 with VP4 capsid proteins of alphasphaerolipoviruses PH1, SH1 and HHIV-2. Similar amino acids are blocked in black (except for the c-terminal extension).

GENUS: *Gammasphaerolipovirus*

Virion morphology

P23-77 virus particles are spherical, tailless and have an average diameter of 78 nm. Approximately 15 nm long stick-like spikes emerge from the five-fold vertices (Jaatinen et al., 2008, Figure 8). An inner lipid membrane is located between the 6 nm thick capsid and the circular dsDNA genome. Capsid and membrane are connected by proteins at the five-fold vertices. The lipids are selectively acquired from the host cell during virus assembly (Jalasvuori et al., 2009). The P23-77 capsid consists of 270 hexameric and 12 pentameric capsomers, arranged in a T=28, dextro lattice. The only other characterized virus with such unusual capsid architecture is haloarchaeal virus SH1 (Jäälinoja et al., 2008), type species of the newly proposed genus “Alphasphaerolipovirus” within the family “Sphaerolipoviridae”.

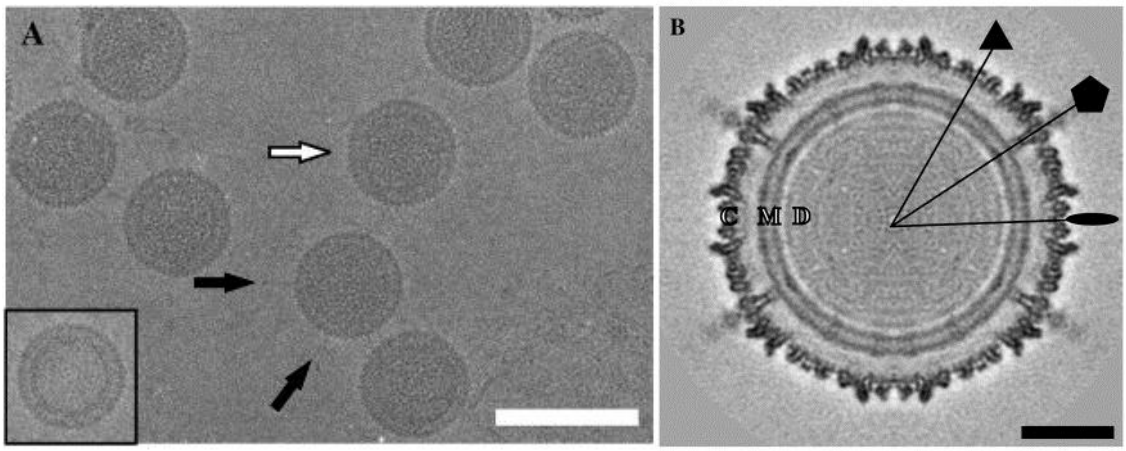


Figure 8: (A) Electron micrograph of the P23-77 virion. Particles are spherical, tailless and mostly filled with DNA (white arrow). Very rarely, empty particles are observed (inset). Thin spikes extend from the surface of some virions (black arrow). Scale bar, 100 nm. (B) Three dimensional image reconstruction of the P23-77 virion. Symmetry axes are designated with a black ellipse (2-fold), triangle (3-fold) and pentagon (5 fold). Capsid shell (C), membrane (M) and DNA (D) are indicated. Scale bar, 20 nm. Figure reproduced from (Jaatinen et al., 2008) with permission from Elsevier.

Genome analysis

The circular dsDNA genomes of P23-77 (Jalasvuori et al., 2009) and IN93 (Matsushita and Yanase, 2009) differ in size and gene content (Figure 9). The genome of IN93 is 2568 nt larger than the genome of P23-77 and contains six additional ORFs. The main difference is the presence of an integration cassette encoding a LexA-like repressor, an endonuclease and an integrase required for the lysogenic cycle (ORFs 36-39) in the genome of IN93. The genes of the integration cassette are the only ones located on the opposite strand with respect to the rest of the genes. P23-77 lacks the integration cassette, which is reflected by its smaller genome size and the strictly lytic life style. The nucleotide similarity across the aligned genomes is 56%. There is a high level of gene synteny between the two genomes. 32 of the 37 predicted ORFs in the P23-77 genome have homologs in IN93, yet 78% of the gene products lack similarity to any other protein sequences in public databases. The viral core proteins – the putative genome packaging ATPase (ORF13) and the small (ORF16) and large (ORF17) major capsid proteins (MCPs) – are among the most conserved proteins in P23-77 and IN93 with sequence identities of 79%, 74% and 79%, respectively. The three core genes are also conserved in the archaeal members of the "Sphaerolipoviridae".

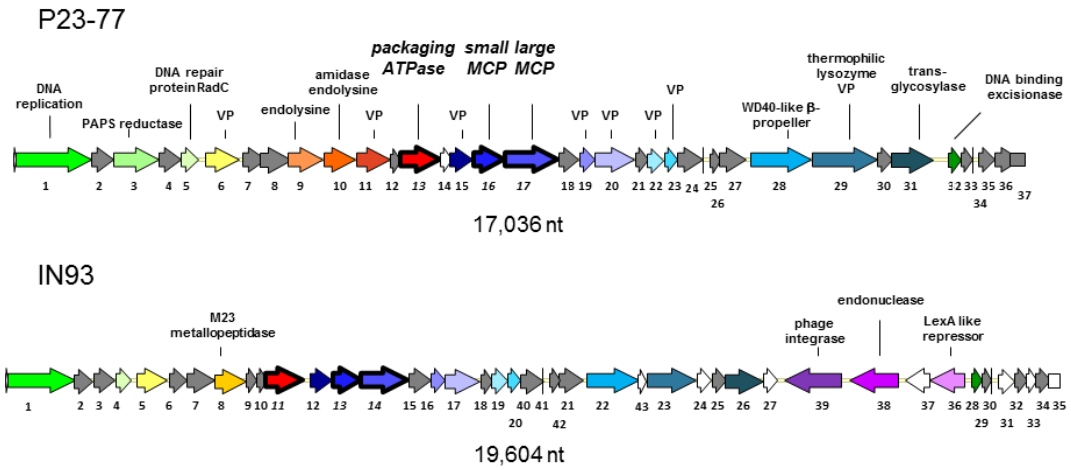


Figure 9: Genomes of proposed gammasphaerolipoviruses P23-77 (GQ403789) and IN93 (AB063393). Genomes are linearized for clearer presentation. ORFs are represented by arrows. ORF numbers are according to gene bank entry. We have assigned four new ORFs (ORF40-43) to the genome of IN93. Genes are shown in color when their gene products had been identified as structural component of the virus (VP= virion protein) or a function is assigned according to experimental data or hits in BLAST search, respectively. Genes of unknown function are marked grey (shared by all members) or white (found in only one member). Genes encoding viral core proteins ATPase and major capsid proteins (MCP) are marked in italics and bold framed arrows.

Structure of major capsid proteins and capsid organization

Recently, the small (VP16) and the large (VP17) MCPs of P23-77 were crystallized and their structures determined (Rissanen et al., 2013). The core fold of both proteins is a nearly identical eight-stranded beta-barrel, which is not found in other dsDNA viruses of Bacteria (Figure 3, A-C). The capsid surface of P23-77 is covered with small turret-like protrusions (Jaatinen et al., 2008). The high resolution structures of the P23-77 capsid proteins, fitted into the electron cryo-microscopy reconstruction (cryo-EM) of the P23-77 virion (Jaatinen et al., 2008), showed that turrets are formed by the upper domain of VP17, while VP16 and the lower domain of VP17 form the base of the capsomers (Rissanen et al., 2013, Figure 10D).

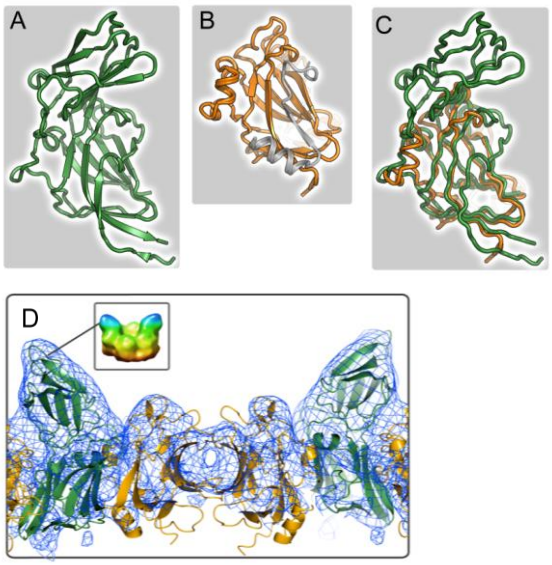


Figure 10: Structures of (A) P23-77 major capsid protein (MCP) VP17 (green, PDB ID code 3ZMN) and (B) P23-77 MCP VP16 (orange, PDB ID code 3ZMO) show the eight stranded single beta-barrel core fold. In addition, VP17 has an upper domain. (C) VP16 superimposed on the lower domain of VP17. (D) P23-77 capsid protein structures fitted into the P23-77 virion cryo-EM reconstruction (EMDB ID code: emdb_1525). The upper domains of VP17 form turrets protruding from the capsomer base (inset).

P23-77 has two distinct types of pseudohexameric capsomers (Figure 11). Both have two turrets built by the upper domain of VP17 and arranged either on the same side or on the opposite corners of the capsomer. Two types of turreted capsomers are also found in alphaspaerolipovirus SH1, but they have either two or three turret protrusions (Jäälinoja et al., 2008). The two coat proteins of SH1 are likely to participate in building the capsomers in the same way as in P23-77: small and large MCPs form the hexagonal base of the capsomer with turrets produced by an upper domain of the large MCP. Indeed, the X-ray structures of the P23-77 MCPs are superimposable within the cryo-EM density maps of the SH1 capsomers, producing a reasonable fit (Rissanen et al., 2013). Collectively, the same capsid geometry (T=28), structural similarity between the corresponding MCPs as well as similar capsid stabilization principles utilized by P23-77-like and SH1-like viruses suggest that the two viral groups have evolved from a common ancestor.

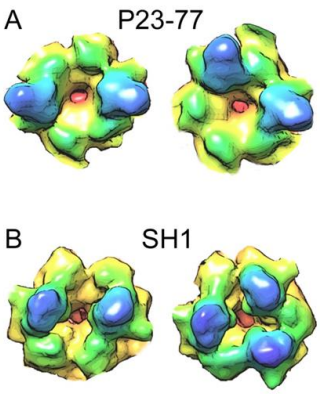


Figure 11: Comparison of the two capsomer types of P23-77 (A) and SH1 (B).