Vesicle-like virion of *Haloarcula hispanica* pleomorphic virus 3 preserves high infectivity in saturated salt

Tatiana A. Demina*, Nina S. Atanasova*, Maija K. Pietilä[#], Hanna M. Oksanen[§], and Dennis H. Bamford

Institute of Biotechnology and Department of Biosciences, Viikinkaari 9, 00014 University of Helsinki, Finland.

*These authors contributed equally

[#]Current affiliation: Department of Food and Environmental Sciences, Viikinkaari 9, 00014 University of Helsinki, Finland

[§]Corresponding author: Hanna Oksanen, hanna.oksanen@helsinki.fi

Abstract

Hypersaline environments that are subject to salinity changes are particularly rich in viruses. Here we report a newly isolated archaeal halovirus, *Haloarcula hispanica* pleomorphic virus 3 (HHPV3). Its reproduction significantly retards host growth and decreases cell viability without causing lysis. HHPV3 particles require a minimum of 3 M NaCl for stability and maintain high infectivity even in saturated salt. Notably, virions are irreversibly inactivated at ~1.5 M NaCl in neutral pH, but tolerate this salinity at alkaline pH. The HHPV3 virion is a pleomorphic membrane vesicle containing two major protein species and lipids acquired nonselectively from the host membrane. The circular double-stranded DNA genome contains a conserved gene block characteristic of pleolipoviruses. We propose that HHPV3 is a member of the *Betapleolipovirus* genus (family *Pleolipoviridae*). Our findings add insights into the diversity observed among the pleolipoviruses found in hypersaline environments.

Key words: halovirus; pleomorphic virus; virus life cycle; genome sequence; *Pleolipoviridae*; archaeal virus; hypersalinity.

Introduction

Hypersaline environments are inhabited by diverse microbes including halophilic or halotolerant eukaryotes, bacteria, archaea, and viruses. Of the ~100 haloviruses isolated so far, ~90 are archaeal. These represent diverse morphotypes: tailed icosahedral (myoviruses, siphoviruses, podoviruses), tailless icosahedral with an internal membrane, spindle-shaped, and pleomorphic (Atanasova *et al.*, 2015b, Atanasova *et al.*, 2016). In nature, halophilic microbes are subjected to marked salinity changes due to rainfall and water evaporation. Inevitably NaCl concentration affects halovirus life cycles and adsorption rates (Kukkaro and Bamford, 2009, Mei *et al.*, 2015). For example, high salinity inhibits the adsorption of myovirus Hs1 to its host, *Halobacterium salinarum,* and the virus establishes persistent infection (Torsvik and Dundas, 1980). S5100, another myovirus infecting *Hbt. salinarum*, also initiates a persistent infection in saturated NaCl, but enters a lytic cycle upon reduction of the salinity (Daniels and Wais, 1990). Moreover, it has been observed that *Hbt. salinarum* virus densities in a natural brine pool vary depending on the salinity, increasing at lowered salt concentrations (Wais and Daniels, 1985). In contrast, the icosahedral, membrane-containing *Natrinema* temperate virus SNJ1 displays a lytic life cycle at high salinity (3.4 and 5 M NaCl) and lysogenic cycle at lower (3 M NaCl) salinity (Mei *et al.*, 2015).

The concentration of NaCl and other salts can also affect the infectivity of haloviruses. Infectivity of the *Haloarcula. hispanica* tailless icosahedral virus SH1 is dependent on the concentrations of NaCl and $MgCl_2$ (Porter *et al.*, 2005), and a drop in salinity causes dissociation of the particles (Kivelä *et al.*, 2006). However, 2 M Mg^{2+} alone can compensate for the absence of NaCl and maintain SH1 infectivity (Porter *et al.*, 2005). Similarly, for *Hbt. salinarum* virus ΦH, magnesium compensates for the lowered levels of NaCl or KCl (Schnabel *et al.*, 1982). *Har. hispanica* tailless icosahedral virus HHIV-2 tolerates low levels of NaCl (100 mM), but is sensitive to the absence of Ca^{2+} (Jaakkola *et al.*, 2012). On the other hand, tailless icosahedral haloarchaeal virus HCIV-1 of *Har. californiae* does not require the presence of any particular ions in high amounts (Demina *et al.*, 2016). The loss of

infectivity caused by altered salinity can, in some cases, be reversed, as shown for the haloarchaeal icosahedral tailed viruses HVTV-1, HSTV-1, and HSTV-2 (Pietilä *et al.*, 2013a, Pietilä *et al.*, 2013b).

The pleomorphic archaeal virus morphotype was discovered in 2009 with the isolation of *Halorubrum* pleomorphic virus 1 (HRPV-1), the first archaeal virus known to have a ssDNA genome (Pietilä *et al.*, 2009). Since then, 11 more pleolipoviruses have been isolated (Atanasova *et al.*, 2015a, Li *et al.*, 2014, Liu *et al.*, 2015, Pietilä *et al.*, 2012, Roine *et al.*, 2010), and HRPV-1 has been designated as the type species of the family *Pleolipoviridae* (Pietilä *et al.*, 2016). The pleolipovirus virion is a spherical lipid vesicle with two major protein species: an internal membrane-associated protein and a spike protein irregularly distributed on the surface (Pietilä *et al.*, 2010, Pietilä *et al.*, 2012). This virion organization has been observed for pleolipoviruses HRPV-1, HRPV-6, HHPV-1, HRPV-2, HGPV-1, HRPV-3, and His2, with the exception that there are two spike protein species in His2 and two membrane protein species in HGPV-1 (Pietilä *et al.*, 2010, Pietilä *et al.*, 2012). The virion lipids are acquired nonselectively from the host cell membrane (Pietilä *et al.*, 2010, Pietilä *et al.*, 2012). In contrast to the conserved virion organization, the genome types of pleolipoviruses are diverse and include linear dsDNA (His2), circular ssDNA (HRPV-1, HRPV-2, HRPV-6, HHPV-2), circular dsDNA (HHPV-1), and circular dsDNA with ssDNA regions (HRPV-3, HGPV-1, SNJ2) (Li *et al.*, 2014, Liu *et al.*, 2015, Pietilä *et al.*, 2009, Roine *et al.*, 2010, Senčilo *et al.*, 2012). Pleolipoviruses have been classified into three genera within the *Pleolipoviridae* family (Pietilä *et al.*, 2016). Alphapleolipoviruses (HHPV-1, HHPV-2, HRPV-6, HRPV-2, HRPV-1) encode a putative rolling-circle replication initiation protein, betapleolipoviruses (HRPV-3, HGPV-1, SNJ2) encode a conserved protein often predicted to have a winged-helix DNA-binding domain, and the only described gammapleolipovirus (His2) encodes a putative protein-primed DNA polymerase (Liu *et al.*, 2015, Pietilä *et al.*, 2016). Despite their different genome types and low sequence identities, all sequenced pleolipovirus genomes exhibit synteny and contain a conserved gene cluster that encodes major structural proteins and other putative proteins (Senčilo *et al.*, 2012).

With one exception, the described pleolipoviruses have a non-lytic life cycle characterized by continuous virus production accompanied by retardation of host growth. That exception is the temperate pleomorphic virus SNJ2 (Liu *et al.*, 2015). It coexists with another temperate tailless icosahedral virus, SNJ1, in *Natrinema* cells (Liu *et al.*, 2015). Successful production of SNJ2 particles depends on the presence of SNJ1, although the exact mechanism for this interaction is unknown. The SNJ2 genome has an ORF (ORF 19) characteristic of the betapleolipovirus group (Liu *et al.*, 2015), but it also encodes an integrase and a ΦH1-repressor-like protein (Liu *et al.*, 2015). SNJ2 is able to integrate into and excise from the host tRNA-Met gene (Liu *et al.*, 2015), thus possibly providing a link for the many pleolipoproviruses identified in haloarchaeal plasmids and chromosomes (Bath *et al.*, 2006, Chen *et al.*, 2014, Dyall-Smith *et al.*, 2011, Liu *et al.*, 2015, Pietilä *et al.*, 2009, Roine *et al.*, 2010, Senčilo *et al.*, 2012).

The overwhelming diversity of viral genomic sequences and lack of a shared common gene imposes challenges for virus classification. Given the limitation in protein fold space and virion structural constraints, it was proposed that the entire viral universe can be systemized using a structure-based approach (Bamford *et al.*, 2002, Bamford, 2003). The structures of the elements that define the viral "self", i.e., major capsid proteins, are conserved and fall into a limited number of lineages (Abrescia *et al.*, 2012, Bamford, 2003, Krupovič and Bamford, 2010). Each lineage embraces viruses with deep evolutionary relationships that now may have no sequence similarity, that can even have different genome types, and that infect hosts from all three domains of life (Bamford, 2003). Consistent with this universal concept was the discovery of a group of archaeal pleomorphic viruses that share a conserved virion organization similar to that of the bacterial plasmaviruses. However, answering the question of whether the pleomorphic viruses form a distinct ancient lineage awaits the recovery of additional novel isolates and determination of their key protein structures to high resolution.

In this study, we report the isolation and molecular characterization of a new virus, *Haloarcula hispanica* pleomorphic virus 3 (HHPV3). Based on its characteristics, including pleomorphic virion

architecture and genome organization, we propose that HHPV3 is a member of the haloarchaeal pleomorphic viruses, and specifically that it belongs to the *Betapleolipovirus* genus of the *Pleolipoviridae* family (Pietilä *et al.*, 2016). We also report eight new pleolipovirus-related putative proviruses in the chromosomes of halophilic archaea, which, combined with the previously known ones (Liu *et al.*, 2015), further highlights the abundance of pleolipoviruses.

Materials and Methods

Archaeal strains, viruses, and growth conditions

Strains and viruses used in this study are listed in Table 1 and Table S1. Strains were aerobically grown in modified growth medium (MGM) (Nuttall and Dyall-Smith, 1993) at 37 °C. Artificial, 30 % saltwater (SW) containing 240 g NaCl, 30 g $MgCl_2 \times 6H_2O$, 35 g $MgSO_4 \times 7H_20$, 7 g KCl, 5 ml of 1 M $CaCl_2 \times 2H_20$, and 80 ml of 1 M Tris-HCl pH 7.2 (per liter of water) was prepared as described in the Halohandbook (http://www.haloarchaea.com/resources/halohandbook/). One litre of medium contained 23 % SW, 5 g of peptone (Oxoid), and 1 g of Bacto yeast extract (Becton, Dickinson and Company). For the top layer and solid media, 4 g or 14 g of Bacto agar (Becton, Dickinson and Company) was added, respectively.

Virus isolation, virus stock preparation, and analysis of plaque variants

Salt crystals were collected from the bottom mud at the saltern of Samut Sakhon, Thailand (13°32'N; 100°17'E) in December, 2010. Three grams of crushed crystals were dissolved in 7 ml of 6 % SW (37 °C, 3 h), and the suspension (100 µl) was used for plaque assay with *Har. hispanica*. After 3 d incubation at 37 °C, one hazy, ring-like plaque with a diameter of ~4 mm was picked and purified by three consecutive plaque assays..

Semiconfluent plates, incubated for 3 d at 37 °C, were used for the preparation of virus agar stocks. The plate top layer was collected and incubated in MGM (3 ml per plate) at 37 °C for 1.5 h with shaking, and debris was removed by centrifugation (Sorvall F14 rotor, 6000 rpm, 4 °C, 20 min). Both, the original hazy ring-like plaque, and a new clear plaque variant were isolated from the initial agar

stock. Both plaque variants were purified three consecutive times followed by agar stock preparation. The plaques were screened by PCR using virus-specific primers (1F and 1R, Table S2), which had been designed according to the obtained virus genome sequence (see below). For the preparation of PCR templates, a picked plaque was incubated in ultrapure water (4 °C, overnight), SDS was added to the final concentration of 1 % (w/v), the mixture was boiled for 5 min followed by dilution in ultrapure water (1:10). The obtained PCR products were sequenced as described below (Sanger sequencing).

All further experiments were performed using the virus stocks of the clear plaque variant, designated as HHPV3, except for the host range test (see below) which assayed both plaque variants. Stability of the agar stock titer at 4 °C was monitored weekly for three months by plaque assay. HRPV-1, HGPV-1, and HRPV-3 agar stocks were prepared as described previously (Pietilä *et al.*, 2009, Pietilä *et al.*, 2012).

Virus host range

HHPV3 infectivity was tested on 47 haloarchaeal strains (Table S1) by a spot-on-lawn assay. Drops (10 µl) of undiluted and 1:100 diluted HHPV3 agar stock were applied on the lawn (top layer) of the test strains inoculated in early stationary growth phase. The plates were incubated at 37 °C for 3–5 d and screened for growth inhibition. Positive results were verified by plaque assay.

Adsorption assay

Cells from 100 ml of *Har. hispanica* liquid culture (OD_{550}= 1.0) were collected by centrifugation (Sorvall F14 rotor, 7000 rpm, 22 °C, 15 min) and suspended into 10 ml of fresh MGM at 37 °C. The culture was infected with HHPV3 using a multiplicity of infection (MOI) of 0.14 and grown aerobically at 37 °C. Samples were collected until 6 h p.i., diluted in ice-cold MGM (1:100), centrifuged (Eppendorf table centrifuge, 13 000 rpm, 1 min, 22 °C), and the titer was then determined by plaque assay. The adsorption rate constant (k) was calculated using the formula $k = \frac{2.3}{B \times t} \times$

$log\ \frac{p0}{p}$, where p0 and p represent free virus concentrations at the zero time point and after time period t, respectively, while B represents the concentration of the cells (Adams, 1959).

Virus life cycle

In order to optimize virus production, host cultures at different cell densities (4.9×10^8, 6.0×10^8, 9.5×10^8, 1.6×10^9, and 1.8×10^9 CFU/ml) were tested using different MOI values (5, 15, 30, and 40). In the optimized protocol, a liquid culture of *Har. hispanica* was infected with HHPV3 using a MOI of 15 when cell culture turbidity was OD_{550}=1.0 (~1.5×10^9 CFU/ml). At 3 h p.i., the culture was washed twice by collecting (Sorvall SA600 rotor, 8000 rpm, 20 °C, 15 min) and then resuspending the cells in the original volume of fresh MGM. To determine the number of infective centers (IC) and viable cells (VC), cells were collected at 0–3 h p.i. for IC or 0-30 h p.i. for VC (Eppendorf table centrifuge, 13 000 rpm, 22 °C, 1 min) and resuspended in MGM. IC was determined by plaque assay and VC was assayed by plating the samples for measurement of cell numbers. Free progeny virus concentration in the culture supernatant was determined at 4 to 30 h p.i. by plaque assay. Simultaneously, turbidity (OD_{550}) of the infected and uninfected (control) cultures was monitored until 30 h p.i.

Stability assays

Stability of HHPV3 infectivity under different conditions was determined by plaque assay after 3 h and 24 h incubation at 4 °C unless otherwise noted.

Sensitivity to the nonionic detergents Triton X-100 and Nonidet P-40 was tested by incubating virus agar stock with the detergent (1:1000 and 1:10000) at 22 °C for 15 min.

To study the effect of pH on virus infectivity, 1 M stock solutions of the buffers 1,4-piperazinediethanesulfonic acid (PIPES) (for pH 6.1) and Tris-HCl (for pH 7.5, 8.0, and 9.0) were prepared and diluted into 70 mM in 26 % SW devoid of Tris-HCl. Virus stock was diluted 1:1000 in each of the test solutions.

To monitor temperature stability, aliquots (0.5 ml) of virus stock solution were incubated at different temperatures ranging from 4 to 90 °C for 30 min.

Requirements for ions normally present in 26 % SW were determined by diluting virus agar stock 1:1000 in 26 % SW pH 7.5 that lacked one, two, or three ionic compounds. For $CaCl_2$ sensitivity test, virus agar stock was diluted 1:10000 in 10 mM ethylene glycol-bis(2-aminoethylether)-*N,N,N',N'*-tetraacetic acid (EGTA) in 26 % SW pH 7.5 devoid of $CaCl_2$. In addition, the effect of $CaCl_2$ on virus infectivity was tested by diluting the agar stock 1:10000 in 0 mM, 2 mM, 4 mM, 6 mM, 8 mM, or 10 mM $CaCl_2$ in 26 % SW pH 7.5.

For NaCl sensitivity test, virus agar stock was diluted 1:1000 in 26 % SW pH 7.5 with variable molar concentrations of NaCl (0-5.5 M). The 5.5 M NaCl solutions (at 4 °C) contained salt crystals (saturated condition). In order to test whether viruses incubated at low NaCl concentrations can recover their infectivity after being transferred back to the optimal NaCl concentration (3.56 M), HHPV3 agar stock was diluted 1:1000 in 26 % SW pH 7.5 containing 0, 1, 1.5, or 2 M NaCl. After 3 h incubation at 22 °C, 5 M NaCl was added to the test solutions to obtain the optimal 3.56 M NaCl. The number of infective particles was then determined after 1, 2, or 60 min incubation at 22 °C. The effect of pH on virus infectivity at different NaCl concentrations was tested by incubating virus stock (1:1000) in 1.5 M or 3.56 M NaCl-containing 26 % SW of pH 7.5 (70 mM Tris-HCl), pH 9.0 (70 mM Tris-HCl) or pH 6.1 (70 mM PIPES).

HRPV-1, HRPV-3, and HGPV-1 sensitivity to NaCl was tested by diluting (1:1000) virus stocks in the optimal virus buffers (Pietilä *et al.*, 2009, Pietilä *et al.*, 2012) with various NaCl concentrations (0-5.5 M), and infectivity was determined by plaque assay after 3 h and 24 h of incubation.

Virus purification

HHPV3 particles were precipitated from agar stocks using 10 % (w/v) PEG6000 (incubation at 4 °C for 1 h), collected by centrifugation (Sorvall SA600 rotor, 6000 rpm, 4 °C, 20 min), and dissolved in 26 % SW pH 7.5 using 1:10 of the original agar stock volume (overnight on ice). To obtain 1× purified

virus, the PEG-precipitated virus particles were subjected to equilibrium centrifugation in CsCl solution in 26 % SW pH 7.5 containing 2.5 M NaCl, mean density 1.32 g/ml (Sorvall TH641 rotor, 25 000 rpm, 20 °C, 17 h). The light scattering virus zone was collected, diluted 1:2 in 18 % SW pH 7.5, and concentrated by differential centrifugation (Beckman Ti50 rotor, 30 000 rpm, 20 °C, 2.5 h). The 1× particles were resuspended in 26 % SW on ice overnight.

To obtain 2× purified virus for lipid analysis, PEG-precipitated viruses (0.4 ml) were loaded onto linear 10-40 % sucrose gradients in 26 % SW pH 7.5 and centrifuged (Sorvall TH641 rotor, 25 000 rpm, 15 °C, 17 h). The light scattering virus zone was collected, diluted 1:2 in 26 % SW pH 7.5, and subjected to equilibrium centrifugation in CsCl solution prepared as described above (Sorvall rotor TH641, 25 000 rpm, 20 °C, 17 h). The collected 2× virus zone was diluted 1:2 in 18 % SW pH 7.5, and then concentrated by differential centrifugation (Beckman Ti50 rotor, 30 000 rpm, 20 °C, 2.5 h). The 2× virus particles were resuspended in 26 % SW overnight on ice. The 2× virus particles of HRPV-1 were purified as described previously (Pietilä *et al.*, 2009).

Negative staining and transmission electron microscopy

For transmission electron microscopy (TEM), 5-µl samples of the 1× or 2× purified virus particles were adsorbed on copper pioloform coated grids at 22 °C for 1 min. HHPV3 particles were negatively stained with 1 % (w/v) ammonium molybdate (pH 7.0), 1% (w/v) phosphotungstic acid (pH 7.0), or 3 % (w/v) uranylacetate (pH 4.5) with 60 s, 60 s, and 30 s incubation times, respectively, at 22 °C. The negatively stained viral particles were visualized by Jeol 1400 transmission electron microscope (Electron Microscopy Unit, Institute of Biotechnology, University of Helsinki) operating at 80 kV acceleration voltage. In order to test HHPV3 sensitivity to these negative stains, 1× purified viruses were incubated in the undiluted stain solution (1:1) for the appropriate staining times (see above) at 22 °C, after which infectivity was determined by plaque assay.

Analysis of virion structural components

Virion components were analyzed using tricine-sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; 16 % acrylamide in the separation gel) (Olkkonen and Bamford, 1989). Gels were stained with Coomassie Brilliant Blue R 250 (Serva) for proteins and Sudan Black B (Sigma Aldrich) for lipids. For the visualization of glycoproteins, 1× purified HHPV3 particles were analyzed in 16 % SDS-PAGE gels stained with Pro-Q Emerald (Invitrogen) (for glycoproteins) and SYPRO-Ruby (for all proteins) according to the manufacturer's instructions with the exception that Pro-Q Emerald staining was performed twice.

Viral lipids were isolated from 2× purified HHPV3 particles according to the protocol by Folch (1957) as modified by Kates (1972) (Folch *et al.*, 1957, Kates *et al.*, 1972). The extracted lipids were analyzed on pre-activated thin layer chromatography (TLC) silica plates and visualized by ammonium molybdate staining (Arnold *et al.*, 2000). The plate was quickly dipped into a solution containing 10 % (v/v) H_2SO_4 and 5 % (w/v) ammonium molybdate, after which the excess liquid was dried and the plate was incubated at 140 °C for 15 min.

Mass spectrometry analysis of viral proteins

Viral proteins were separated by SDS-PAGE and then silver-stained (O'Connell and Stults, 1997). Protein bands were cut from the gel and digested "in-gel". Cysteine bonds were reduced by incubation with 0.045 M dithiothreitol (Sigma-Aldrich, USA) for 20 min at 37 °C and alkylated with 0.1 M iodoacetamide (Fluka, Sigma-Aldrich, USA) (22 °C). Following trypsin digestion (0.75 µg of trypsin [Promega] at 37 °C overnight), peptides were purified using C18 microspin columns (Harvard Apparatus) according to the manufacturer's protocol. The dried peptides were reconstituted in 30 µl of buffer A (0.1 % trifluoroacetic acid in 1% acetonitrile). Liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) analysis was carried out on an EASY-nLC II, (Thermo Fisher Scientific, Germany) connected to a Velos Pro-Orbitrap Elite hybrid mass spectrometer (Thermo Fisher Scientific, Germany) with nano-electron spray ion source (Thermo Fisher Scientific,

Germany). The LC-MS/MS samples were separated using a two-column setup consisting of a 2 cm C18-A1 trap column (Thermo Fisher Scientific, Germany), followed by a 10 cm C18-A2 analytical column (Thermo Fisher Scientific, Germany). The samples were separated using a 70 min linear gradient of 5 to 35 % buffer B (0.1 % trifluoroacetic acid in 98 % acetonitrile), followed by 5 min gradient of 35 to 80 % buffer B, 1 min gradient of 80 to 100 % buffer B, and 9 min column wash with 100 % buffer B at a constant flow rate of 0.3 µl/min. A 6 µl-sample was injected for each LC-MS/MS analysis run. Full MS scan was acquired with a resolution of 60 000 at normal mass range in the Orbitrap analyzer method set to fragment the 20 most intense precursor ions with collision-induced dissociation (energy 35). Data was acquired using TLQ Tune software and the acquired MS2 scans were searched against our in-house constructed protein database based on HHPV3 and *Har. hispanica* protein sequences using the Sequest search algorithms in Thermo Proteome Discoverer. Precursor mass tolerance and fragment mass tolerance were set to less than 15 ppm and 0.8 Da. Carbamidomethylation (+57.021464 Da) of cysteine residues was used as static modification, and oxidation (+15.994491 Da) of methionine was used as dynamic modification. Only full-tryptic peptides with a maximum of one missed cleavage were considered.

Virus genome isolation and sequencing

For nucleic acid extraction, 1× purified HHPV3 particles were diluted in ultrapure water and treated with proteinase K (0.5 µg/ml, Thermo Scientific) and 2 % (w/v) SDS at 37 °C for 45 min. Phenol-ether extracted nucleic acid was precipitated using NaCl or ammonium acetate and ethanol. Purified nucleic acid was treated with RQ1 DNase (Promega), RNase A (Fermentas), Exonuclease III (Fermentas), Mung Bean nuclease (New England Biolabs), Bal31 exonuclease (New England Biolabs), and five restriction enzymes (*BamH*I [Roche], *Hind*III [Roche], *Apa*l [Promega], *Pst*I [Roche], *Msc*l [New England Biolabs]) according to the manufacturer's instructions and then analyzed in agarose gels stained with ethidium bromide.

A mixture of (i) untreated DNA, (ii) DNA treated with *Sulfolobus* polymerase IV (New England Biolabs) and T4 ligase (Thermoscientific), and (iii) DNA treated with Klenow polymerase (Fermentas) and T4 ligase (Thermoscientific) was used for virus genome sequencing at the DNA Sequencing and Genomics Laboratory, Institute of Biotechnology, University of Helsinki. The purified DNA was fragmented by sonication (Bioruptor NGS, Diagenode), and a standard Illumina TruSeq library was constructed. The obtained library was analysed on an MiSeq instrument using MISeq reagent kit v3 (600 cycles). The obtained paired-end sequences were trimmed using Cutadapt (Martin, 2011) and assembled with SPAdes (Bankevich *et al.*, 2012).

The circular conformation of the molecule was verified by Sanger sequencing of the PCR product using primers 2F and 2R (see Results, Table S2). The PCR product was purified by MultiScreen PCR 96 (Cat No. LSKMPCR50; Millipore) and the sequencing was performed using BigDye® Terminator v3.1 Cycle Sequencing Kit (Part No. 4336921; Applied Biosystems) according to the manufacturer's instructions. Sequencing reactions were performed with ABI3130XL Genetic Analyzer containing 16-capillaries (Applied Biosystems) and cleaned using CleanSEQ kit (Agencourt).

Genome sequence accession number

The HHPV3 genome sequence is available in the NCBI database with the accession number KX344510.

Virus genome annotation

The HHPV3 genome sequence was analyzed using Geneious (version 6.1.6) (Kearse *et al.*, 2012). The replication origin was predicted using Ori-Finder 2 (Luo *et al.*, 2014). ORFs were predicted by Glimmer (Delcher *et al.*, 2007) and GeneMarkS (Besemer *et al.*, 2001), and final coordinates were chosen in collinearity with previous annotations of related viral and putative proviral sequences. ORF GC percentages were calculated using "Genomics %G~C Content Calculator" by Science Buddies (http://www.sciencebuddies.org/science-fair-projects/project_ideas/Genom_GC_Calculator.shtml). The isoelectric points and molecular masses of (putative) proteins were calculated using the EXPASY

"Compute pI/MW" tool. Transmembrane helices were predicted with TMHMM v 2.0 server (Krogh *et al.*, 2001), and coiled coil regions were identified using the COILS tool (Lupas *et al.*, 1991). Signal peptides were predicted with Signal P (version 4.1) (Petersen *et al.*, 2011). BLASTX and BLASTP (dated 5 April 2016) were used to search for sequences homologous to (putative) HHPV3 proteins (Altschul *et al.*, 1990). Conserved domains were searched against NCBI's conserved domain database (Marchler-Bauer *et al.*, 2015) and using InterProScan at EMBL-EBI (Jones *et al.*, 2014) (search dated 5 April 2016). Amino acid or nucleotide identities were calculated at EMBL-EBI using EMBOSS Needle (McWilliam *et al.*, 2013). Phylogenetic analysis of the protein sequences of HHPV3, pleolipoviruses, and related putative provirus regions was conducted using Molecular Evolutionary Genetics Analysis (MEGA) software (version 5.05) (Tamura *et al.*, 2011). The sequences were aligned by MUSCLE (Edgar, 2004), and the maximum likelihood trees were built using the Jones-Taylor-Thornton (JTT) amino acid substitution model with 1000 bootstrap values.

Results

Isolation of halovirus HHPV3 from salt crystals

Originally, a single plaque of hazy ring-like morphology was obtained on the lawn of *Har. hispanica*, but during subsequent plating of the hazy plaque also clear plaques appeared. Both plaque types were plague-purified and the virus stocks of the two variants gave typical titers of ~1×10^{11} and ~1×10^{9} PFU/ml for the clear and hazy ones, respectively. The plaque morphology of the clear variant was always clear, whereas the hazy plaque stocks occasionally produced both plaque types (Figure S1). The PCR products amplified using the plaques of the two variants as templates were 100 % identical (~680 bp of gene *3* encoding a putative spike protein, see below). In addition, the two plaque variants had the same very narrow host range. From the tested 47 haloarchaeal strains (Table S1), only *Haloarcula hispanica* served as a host. Most likely clear and hazy plaques represent the same virus species. The clear plaque variant due to its stability and higher titer, was selected, designated HHPV3, and used in this study.

High virus production and host growth retardation by HHPV3 infection

About 85 % of HHPV3 particles adsorbed to the *Har. hispanica* cells in 4 h (Figure 1 A). The adsorption rate constant calculated for the first 30 min p.i. was 2.4×10^{-12} ml/min. The highest virus production was obtained by infecting early stationary growth phase cells (OD_{550}=1.0) at an MOI of 15. Already at 30 min p.i. the number of infective centers (IC) was equal to the number of viable cells (VC). By 3 h p.i. the number of IC was ~10 times that of VC, indicating that the majority of cells were infected (Figure S2). Virus release was detected starting ~ 6 h p.i., and the titer of free progeny viruses in the culture supernatant increased to a the maximum of $\sim 8 \times 10^{10}$ PFU/ml (Figure 1 B). By ~12 h p.i. the turbidity (OD_{550}) of the infected culture reached ~1.3 and remained at this level, while growth of the uninfected culture continued until the OD_{550} was ~1.7 (Figure 1 B). Also at 12 h p.i. we observed a significant decrease in the number of VC in the infected culture. At that time the optical densities of the uninfected and infected cultures were equal (OD_{550} ~1.3) (Figure 1 B), but already their VC counts differed by about two orders of magnitude (Figure 1 C).

Effects of pH and ion concentration on HHPV3 infectivity

HHPV3 tolerated a pH range of 4.5–9 (in 26% SW containing 3.5 M NaCl) (Figure S3 A) and temperatures up to 40 °C without significant loss of infectivity (Figure S3 B). Infectivity decreased in the absence of $CaCl_2$ (three orders of magnitude decrease in 24 h) and NaCl (one order of magnitude decrease in 24 h), while the absence of magnesium or potassium had no major impact (Figure S3 C). When all calcium ions were removed from the buffer with a chelating agent (EGTA), HHPV3 infectivity dropped by four orders of magnitude (Figure S3 D). The presence of 2 mM $CaCl_2$ was sufficient to maintain virus infectivity (Figure S3 D). HHPV3 infectivity was stable in 26% SW containing from 3 to 5.5 M NaCl, the latter being a saturated solution containing salt crystals (Figure 2 A). A peculiar, pH-dependent, irreversible loss of infectivity was observed in 26% SW containing 1.5 M NaCl. The decrease, evident at pH 7.5, was more marked at pH 6.1 but not apparent at pH 9.0 (Figure 2 A). When the effect of pH was tested in 26 % SW containing 3.5 M NaCl buffered to pH

6.1, 7.5, or 9.0, no drastic differences in HHPV3 infectivity were observed (5.3×10^{10}, 8.2×10^{10}, and 2.3×10^{10} PFU/ml, respectively, after 24 h).

In comparison, betapleolipoviruses HGPV-1 and HRPV-3 were highly infective in their optimal buffers over a range of NaCl concentrations from 0 M to 5.5 M (Figure 2 C and D). However, infectivity of the pleolipovirus type species HRPV-1 was greatest at 2–3 M NaCl and decreased at both higher and lower concentrations (Figure 2 B). Following 24 h incubation at 0 M and 5.5 M (saturated) NaCl, HRPV-1 infectivity decreased five and two orders of magnitude, respectively (Figure 2 B).

Purification of HHPV3

The optimized 1× purification protocol (equilibrium centrifugation in CsCl) recovered approximately 15 % of the infective HHPV3 particles (Table S3). The mean equilibrium density of HHPV3 virions was approximately 1.26 g/ml in CsCl (Figure 3 A), and the specific infectivity of the concentrated 1× virus particles was ~8.4×10^{11} PFU/A_{260}. For the lipid analysis, the 2× purification protocol including tandem equilibrium centrifugations in sucrose and CsCl was employed. Yield of the 2× viruses was lower, around 1 %, with a specific infectivity of ~4.4×10^{11} PFU/A_{260}. The light scattering 2× virus zone had a mean equilibrium density of 1.28 g/ml in CsCl.

The 2× purified HHPV3 virions were spherical but slightly pleomorphic particles with an average diameter of ~50 nm (Figure 3 B) resembling the morphology of pleolipoviruses (Pietilä *et al.*, 2016). However, whereas some pleolipoviruses exhibit drastic structural variability depending on the negative stain used (Pietilä *et al.*, 2012), the morphology of HHPV3 was similar for the three negative stains tested: uranyl acetate (Figure 3 B), phosphotungstic acid, and ammonium molybdate (data not shown). The best image quality was achieved using uranyl acetate (Figure 3 B). All three stains reduced HHPV3 (2× virus) infectivity by approximately one order of magnitude.

HHPV3 virion protein and lipid composition

HHPV3 virion contains two major protein species VP1 and VP3 ("VP" for virion protein) encoded by genes *1* and *3* (Figure 4, Table S4, see genome annotation below). Staining with Pro-Q Emerald dye showed that none of the HHPV3 structural proteins were glycosylated (Figure 4 B and C). We observed an infectivity decrease of eight orders of magnitude after HHPV3 was treated with Triton X-100 and a decrease of four orders of magnitude with the nonionic detergent Nonidet P-40. These observations, which was consistent with the lipid signal obtained for the purified particles with Sudan Black B staining (Figure 4 A) suggested that the virions contain lipids. The presence of lipids in HHPV3 virions was confirmed by lipid extraction and TLC (Figure 5). The major lipid species present are phosphatidylglycerol, phosphatidylglycerophosphate methyl ester, phosphatidylglycerosulfate, and triglycosyl glycerodiether. Comparison of the HHPV3 lipid profile with that of *Har. hispanica* (Bamford *et al.*, 2005) revealed that all lipid species were present in approximately the same proportions in both (Figure 5), providing evidence that virion lipids were acquired from the host nonselectively.

HHPV3 genome

The HHPV3 genome was degraded by RQ1 DNase, ExoIII nuclease and several restriction endonucleases, but not with RNase A (Figures S4A and S4B), indicating that it is a DNA molecule. Both Mung Bean and Bal31 nucleases (Figure S4C and D) digested the DNA into multiple fragments. These patterns might indicate that the dsDNA molecule contains nicks or single-stranded regions as has been shown e.g. for pleolipovirus HRPV-3 (Senčilo *et al.*, 2012). Illumina sequencing of the genome yielded a contig with identical ends (127-nt region, nt coordinates 2603-2729) suggesting a circular nature of the molecule. The circular conformation was verified by PCR and subsequent Sanger sequencing of the region including the contig ends. The size of the genome is 11,648 kbp. In addition, the circular conformation was confirmed by genomic DNA restriction endonuclease digestion patterns. *BamH*I, *Hind*III, *Apa*I, and *Pst*I (each having one restriction site in the genome) resulted in one fragment, while *Mse*I and *Nru*I (each with two restriction sites) yielded two fragments

(Figure S4 B). The conserved motif GCCCA, which was previously shown to precede the single-stranded nicks/regions in betapleolipoviruses HRPV-3 (Senčilo *et al.*, 2012) and SNJ2 (Liu *et al.*, 2015), was also found in the HHPV3 sequence: 10 motifs on the positive strand and 12 on the negative strand (Figure 6 A).

Assignment of HHPV3 to the genus *Betapleolipovirus*

The HHPV3 genome was predicted to have 17 ORFs (numbered from 1 to 17; Table S6) arranged in two clusters with opposite transcriptional directions (Figure 6 A). Two ORFs were confirmed to be protein-encoding genes: gene *1* encoding VP1 and gene *3* encoding VP3 (see protein analyses above). The location of transition sites in Z-curves (Figure 6 B) and predicted origin recognition box (ORB) sequences (Table S5) suggest that the replication origin may be located between ORFs 8 and 9, and the replication terminus between ORF 17 and gene *1* (Figure 6 A). The predicted ORFs varied in size, encoding proteins of 47-613 kDa with a pI of 4-5, except for putative proteins 7 (pI 9.8) and 13 (pI 6.2). Transmembrane helices were predicted in the proteins VP1, VP3, and the putative proteins 2, 4, and 5, all encoded by adjacent genes and ORFs (Table S6). Coiled coil regions were predicted in six (putative) proteins including VP3 (Table S6). Signal peptides of 28 and 19 amino acids were predicted in VP3 and putative protein 4, respectively. Conserved domains were found in VP3 and putative proteins 5, 6, 13, and 15 (Table S7). Putative protein 6 contained a conserved domain characteristic of P-loop NTPases and this protein is homologous to the putative NTPases of the known pleolipoviruses (Figure 6 C). Putative protein 13 is predicted to have a role in transcription regulation based on its conserved domain (Table S7) and similarities to haloarchaeal RNA polymerase sigma subunits. Putative protein 15 had a Dam-replacing domain (Table S7) and showed similarities to restriction endonuclease-like proteins of haloarchaea. Based on sequence comparison with HRPV-1, we propose that HHPV3 gene *1* encodes a protein associated with the virion membrane (VP1), gene *3* encodes a spike protein (VP3) anchored in the membrane, and putative protein 6 is an NTPase.

Genes *1* and *3*, as well as ORFs 4, 5, 6, 7, and 11 all displayed similarity to the conserved cluster of genes seen in viruses classified in the family *Pleolipoviridae* (Pietilä *et al.*, 2016). The sequence similarities are low (Figure 6 C and Table S6), but the genomes are collinear. The overall nucleotide identity between the genomes of HHPV3 and previously characterized pleolipoviruses is ~43%, and the highest identity is to HRPV-2 (~48%) and HGPV-1 (~47%). Together, the observed gene synteny, the absence of a rolling-circle replication initiation protein (RCR Rep) gene, and the presence of ORFs 7 and 11 suggest that HHPV3 belongs to the genus *Betapleolipovirus*.

Presence of putative HHPV3-related proviruses in several haloarchaea

We found HHPV3-related sequence elements in many haloarchaeal chromosomes, some of which had been previously reported as putative betapleolipovirus proviruses (Liu *et al.*, 2015, Pietilä *et al.*, 2016). The highest similarity was found between the HHPV3 genome and a region in chromosome 1 of *Har. hispanica* ATCC 33960. The amino acid sequence of three HHPV3 putative proteins (7, 9, 10) was 100 % identical to the corresponding proteins in *Har. hispanica* (Figure 6 D), while others were less similar (e.g., VP3 and putative protein 11). Another region with high similarity to HHPV3 was identified in *Haloarcula marismortui* ATCC 43049 (region 1, Figure 6 D). Interestingly, HHPV3 putative protein 11 is only 27 % identical to the corresponding protein in *Har. hispanica*, but is 99 % identical to that of *Har. marismortui*. When HHPV3 is compared with the putative proviruses in *Har. marismortui* (region 1) and *Har. hispanica* (Figure 6 D), the VP3 homologs display much lower amino acid identity (56 %) than do the other (putative) proteins in the putative provirus cluster (91-98 %), which suggests that VP3 is evolving more rapidly. Sequence elements related to HHPV3 ORFs were also identified in the genomes of *Haloarcula argentinensis* DSM 12282 (two regions), *Haloarcula vallismortis* ATCC 29715, *Haloarcula japonica* DSM 6131, *Haloarcula rubripromontorii* SL3, *Haloarcula amylolytica* JCM 13557, *Halobacterium* sp. CBA1132, and *Halosimplex carlsbadense* 2-9-1 (Figure S5). In addition, HHPV3-related ORFs were found in previously reported putative pleoliproviruses residing in the genomes of *Halomicrobium mukohataei*

DSM 12286 (Figure S5), *Haloastagnicola larsenii* XH-48, *Halorhabdus utahensis* DSM 12940, *Haloferax volcanii* DS2, *Halopiger xanaduensis* SH-6, *Natrialba magadii* ATCC 43099, and *Halovivax ruber* XH-70 (Liu *et al.*, 2015). In all cases, these proviral regions in archaea were found between tRNA gene(s) and an integrase gene.

Phylogenetic analyses showed that within the dataset used here, the sequences most closely related to HHPV3 are those of the putative proviruses identified in *Har. hispanica* (region 1), *Har. marismortui* (regions 1 and 2), *Har. argentinensis* (regions 1 and 2), *Har. japonica, Har. amylolytica, Har. vallismortis, Har. rubripromontorii* SL3, and more distantly *Halomicrobium mukohataei* (region 1) and *Halosimplex carlsbadense* (Figure S6, colored in pink). All of these putative proviruses have a homolog of HHPV3 ORF2, an ORF that is so far absent from other characterized pleolipoviruses.

Discussion

Classification of HHPV3

The *Pleolipoviridae* family comprises halophilic archaeal viruses with different genome types that are all characterized by pleomorphic virions (Pietilä *et al.*, 2016). Indeed, the remarkable variability in their DNA genome types sets this virus group apart from all other described viral families and calls for exploration of the diversity of additional environmental isolates. Here we report the isolation and characterization of a new haloarchaeal pleolipovirus, HHPV3, that infects *Haloarcula hispanica*. As is typical for the pleolipoviruses, HHPV3's narrow host range is limited to its isolation host (Atanasova *et al.*, 2012, Atanasova *et al.*, 2015a, Bath *et al.*, 2006, Pietilä *et al.*, 2009, Roine *et al.*, 2010). All of the characterized pleolipoviruses have non-lytic life cycles, but the degree of host growth retardation varies. For instance, HHPV3 infection results in substantial retardation of the *Har. hispanica* growth and decreased cell viability, while HRPV-1 infection has almost no effect on the growth of *Halorubrum* sp. PV6 (Pietilä *et al.*, 2009). Further, His2 infection significantly retards *Har. hispanica* growth and decreases the number of viable cells by two to three orders of magnitude, but

causes no detectable physiological changes such as leakage of ATP or reduction in oxygen consumption in the infected cells (Pietilä *et al.*, 2012, Svirskaitė *et al.*, 2016). It has been proposed that pleolipoviruses enter the host cell by fusion of the membrane of the vesicle-like virion with the host cytoplasmic membrane, and that the progeny virions exit by budding (Pietilä *et al.*, 2016).

Here we suggest that the HHPV3 virion exemplifies the relatively simple virion organization observed for pleolipoviruses. The vesicular virion contains only two major protein species: the spike protein VP1 and a membrane-associated protein VP3 (Figure 4). The spike proteins of some pleolipoviruses are modified; for instance, HRPV-1 VP4 is glycosylated (Pietilä *et al.*, 2010), while His2 VP28 and HGPV-1 VP4 both contain lipid moieties (Pietilä *et al.*, 2012). None of HHPV3's structural proteins are modified with glycans (Figure 4 B and C).

Based on genome sequence comparisons, we proposed that HHPV3 be classified in the *Betapleolipovirus* genus of the *Pleolipoviridae* family (Pietilä *et al.*, 2016). The gene synteny, as well as the absence of a rolling-circle replication initiation protein (RCR Rep) gene and the presence of ORFs 7 and 11, suggests that HHPV3 belongs to the genus *Betapleolipovirus*. In addition, the single-stranded interruptions in circular, double-stranded DNA genomes of HRPV-3, HGPV-1, and SNJ2 (Liu *et al.*, 2015, Senčilo *et al.*, 2012) might to be present also in HHPV3 (this study). This feature might be a feature specific to betapleolipoviruses.

The results of this study add further evidence for the abundance of pleolipovirus-related proviruses present in archaea as plasmids or as elements integrated into chromosomal tRNA gene(s) (Bath *et al.*, 2006, Chen *et al.*, 2014, Dyall-Smith *et al.*, 2011, Liu *et al.*, 2015, Pietilä *et al.*, 2009, Roine *et al.*, 2010, Senčilo *et al.*, 2012). Taken together, these observations indicate that pleolipoviruses are wide-spread in saline environments, probably owing to their simple, vesicle-like virion architecture that supports survival in the extremely saline niches of the planet.

Betapleolipovirus adaptation to various salinities

Overall the pleolipoviruses display a range of tolerances to various salinities, an ability obviously crucial for survival in hypersaline environments. The infectivity of particular pleolipoviruses is dependent on the total salinity of the environment (Pietilä *et al.*, 2012). The type species, HRPV-1, has an optimum at 2-3 M NaCl and loses infectivity with both decreased and increased salinities (Figure 2 B). However, HHPV3 as well as approximately half of the previously described pleolipoviruses (HRPV-2, HGPV-1, HRPV-8, and HAPV-2) (Atanasova *et al.*, 2012, Atanasova *et al.*, 2015a) were isolated from salt crystals, which attests to their ability to tolerate saturated salt. HHPV3 requires extreme concentrations of NaCl (at least 3.5 M NaCl) for the maintenance of high infectivity over a broad range of pH. In contrast, the betapleolipoviruses HRPV-3 and HGPV-1 had no significant loss of titer over a remarkable range of NaCl concentrations – from no NaCl to NaCl saturation (5.5 M) (Figure 2 C and D), demonstrating high adaptation to environments with changing salinity.

HHPV3 also demonstrated a peculiar, irreversible, pH-dependent infectivity drop at 1.5 M NaCl. It was previously shown that NaCl may affect the isoelectric points of bacteriophages (Langlet *et al.*, 2008). Possibly alteration of the surface properties of HHPV3 particles by NaCl affects their isoelectric points, which in turn could cause particle aggregation at a particular pH.

Halophilic archaea cope with osmotic stress by regulation of ion flux, synthesis and accumulation of osmolytes, changing protein expression, etc. (Roberts, 2004). Virions, being inert particles, cannot actively regulate their interactions with the surrounding environment. Instead they rely on other capabilities, such as virion properties. Pleolipoviruses use for their virion structure a pleomorphic membrane vesicle, randomly decorated with spike proteins, that functions as a flexible intercellular carrier of their DNA in a challenging hypersaline environment. The release of membrane vesicles is an important physiological process utilized by fungi, bacteria, and archaea, i.e., cells from all three domains of life (Deatherage and Cookson, 2012). Hyperthermophilic archaea produce membrane

vesicles for the transport of metabolites, toxins, and even DNA (Marguet *et al.*, 2013). Moreover, it has been shown that *Thermococcus nautilus* produces membrane vesicles containing plasmid pTN3, which is in fact the genome of a defective virus (Gaudin *et al.*, 2014). Apparently, membrane vesicles – the vehicle for pleolipovirus genome transfer and survival in conditions of changing salinity – is a versatile and multifunctional structure widely used in nature.

The haloarchaeal pleolipoviruses classification challenge

Pleolipoviruses isolated from hypersaline sources from around the globe challenge a fundamental traditional view of virus taxonomy by sharing the same virion organization, while employing dramatically different genome types. Their classification is, however, compatible with a viral taxonomy based on conserved virus structures rather than fast-evolving virus genomic sequences (Abrescia *et al.*, 2012, Bamford, 2003, Krupovič and Bamford, 2010). So far, four structure-based lineages of the icosahedral viruses have been established for the eukaryotic, bacterial, and archaeal viruses (Oksanen *et al.*, 2012). However, to establish a structure-based phylogeny for the enveloped viruses is a more demanding task, as it is quite challenging to resolve structures in virions that lack symmetry (Abrescia *et al.*, 2012). In addition, it is presently unclear which protein structure should be defined as the viral self for pleomorphic viruses. The inner membrane protein seems to be most plausible, because the spike protein is more likely to be involved in virion-host interactions and thus its structure is expected to be less conserved. At this point, more data is needed to determine whether pleolipoviruses form a distinct, ancient virus lineage. If they in fact do, we would expect to find viruses of this lineage associated also with bacteria and eukaryotes. The only probable structural resemblance known so far is to the bacterial mycoplasmavirus L172 (Dybvig *et al.*, 1985), which has pleomorphic morphology and a protein profile similar to that of the known pleolipoviruses.

Acknowledgements

We thank Helin Veskiväli for excellent technical assistance. The authors acknowledge the support of employees and the use of experimental resources of Instruct, a Landmark ESFRI project. We thank

Academy of Finland (funding grant 272853) and University of Helsinki for the support to EU ESFRI Instruct Centre for Virus Production (ICVIR) used in this study. This study was supported by Academy Professor (Academy of Finland) funding grants 283072 and 255342 (D.H.B.), as well as Ella and Georg Ehrnrooth Foundation grant (T.A.D.) and Alfred Kordelin Foundation grant (T.A.D.).

References

Abrescia, N.G., Bamford, D.H., Grimes, J.M., Stuart, D.I., 2012. Structure unifies the viral universe. Annu.Rev.Biochem. 81, 795-822.

Adams, M.H., 1959. Bacteriophages. Interscience Publishers, New York.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J.Mol.Biol. 215, 3, 403-410.

Arnold, H.P., Zillig, W., Ziese, U., Holz, I., Crosby, M., Utterback, T., Weidmann, J.F., Kristjanson, J.K., Klenk, H.P., Nelson, K.E., Fraser, C.M., 2000. A novel lipothrixvirus, SIFV, of the extremely thermophilic crenarchaeon *Sulfolobus*. Virology. 267, 2, 252-266.

Atanasova, N.S., Bamford, D.H., Oksanen, H.M., 2016. Virus-host interplay in high salt environments. Environ.Microbiol.Rep.

Atanasova, N.S., Demina, T.A., Buivydas, A., Bamford, D.H., Oksanen, H.M., 2015a. Archaeal viruses multiply: temporal screening in a solar saltern. Viruses. 7, 4, 1902-1926.

Atanasova, N.S., Oksanen, H.M., Bamford, D.H., 2015b. Haloviruses of archaea, bacteria, and eukaryotes. Curr.Opin.Microbiol. 25, 40-48.

Atanasova, N.S., Roine, E., Oren, A., Bamford, D.H., Oksanen, H.M., 2012. Global network of specific virus-host interactions in hypersaline environments. Environ.Microbiol. 14, 2, 426-440.

Bamford, D.H., 2003. Do viruses form lineages across different domains of life?. Res.Microbiol. 154, 4, 231-236.

Bamford, D.H., Burnett, R.M., Stuart, D.I., 2002. Evolution of viral structure. Theor.Popul.Biol. 61, 4, 461-470.

Bamford, D.H., Ravantti, J.J., Rönnholm, G., Laurinavičius, S., Kukkaro, P., Dyall-Smith, M., Somerharju, P., Kalkkinen, N., Bamford, J.K., 2005. Constituents of SH1, a novel lipid-containing virus infecting the halophilic euryarchaeon *Haloarcula hispanica*. J.Virol. 79, 14, 9097-9107.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J.Comput.Biol. 19, 5, 455-477.

Bath, C., Cukalac, T., Porter, K., Dyall-Smith, M.L., 2006. His1 and His2 are distantly related, spindle-shaped haloviruses belonging to the novel virus group, *Salterprovirus*. Virology. 350, 1, 228-239.

Besemer, J., Lomsadze, A., Borodovsky, M., 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. 29, 12, 2607-2618.

Chen, S., Wang, C., Xu, J.P., Yang, Z.L., 2014. Molecular characterization of pHRDV1, a new virus-like mobile genetic element closely related to pleomorphic viruses in haloarchaea. Extremophiles. 18, 2, 195-206.

Daniels, L.L., Wais, A.C., 1990. Ecophysiology of Bacteriophage S5100 Infecting *Halobacterium cutirubrum*. Appl.Environ.Microbiol. 56, 11, 3605-3608.

Deatherage, B.L., Cookson, B.T., 2012. Membrane vesicle release in bacteria, eukaryotes, and archaea: a conserved yet underappreciated aspect of microbial life. Infect.Immun. 80, 6, 1948-1957.

Delcher, A.L., Bratke, K.A., Powers, E.C., Salzberg, S.L., 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics. 23, 6, 673-679.

Demina, T.A., Pietilä, M.K., Svirskaite, J., Ravantti, J.J., Atanasova, N.S., Bamford, D.H., Oksanen, H.M., 2016. Archaeal virus HCIV-1 highlights conserved elements in icosahedral membrane-containing DNA viruses from extreme environments. mBio accepted.

Dyall-Smith, M.L., Pfeiffer, F., Klee, K., Palm, P., Gross, K., Schuster, S.C., Rampp, M., Oesterhelt, D., 2011. *Haloquadratum walsbyi*: limited diversity in a global pond. PLoS One. 6, 6, e20968.

Dybvig, K., Nowak, J.A., Sladek, T.L., Maniloff, J., 1985. Identification of an enveloped phage, mycoplasma virus L172, that contains a 14-kilobase single-stranded DNA genome. J.Virol. 53, 2, 384-390.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 5, 1792-1797.

Folch, J., Lees, M., Sloane Stanley, G.H., 1957. A simple method for the isolation and purification of total lipides from animal tissues. J.Biol.Chem. 226, 1, 497-509.

Gaudin, M., Krupovic, M., Marguet, E., Gauliard, E., Cvirkaite-Krupovic, V., Le Cam, E., Oberto, J., Forterre, P., 2014. Extracellular membrane vesicles harbouring viral genomes. Environ.Microbiol. 16, 4, 1167-1175.

Gonzalez, C., Gutierrez, C., Ramirez, C., 1978. Halobacterium vallismortis sp. nov. An amylolytic and carbohydrate-metabolizing, extremely halophilic bacterium. Can.J.Microbiol. 24, 6, 710-715.

Jaakkola, S.T., Penttinen, R.K., Vilen, S.T., Jalasvuori, M., Rönnholm, G., Bamford, J.K.H., Bamford, D.H., Oksanen, H.M., 2012. Closely related archaeal *Haloarcula hispanica* icosahedral viruses HHIV-2 and SH1 have nonhomologous genes encoding host recognition functions. J.Virol. 86, 9, 4734-4742.

Javor, B., Requadt, C., Stoeckenius, W., 1982. Box-shaped halophilic bacteria. J.Bacteriol. 151, 3, 1532-1542.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R., Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics. 30, 9, 1236-1240.

Juez, G., Rodriguez-Valera, F., Ventosa, A., Kushner, D.J., 1986. *Haloarcula hispanica* spec. nov. and *Haloferax gibbonsii* spec, nov., two new species of extremely halophilic archaebacteria. Syst.Appl.Microbiol. 8, 1–2, 75-79.

Kates, M., Work, T.S., Work, E., 1972. Techniques of lipidology: isolation, analysis and identification of lipids. North-Holland, Amsterdam.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28, 12, 1647-1649.

Kivelä, H.M., Roine, E., Kukkaro, P., Laurinavičius, S., Somerharju, P., Bamford, D.H., 2006. Quantitative dissociation of archaeal virus SH1 reveals distinct capsid proteins and a lipid core. Virology. 356, 1-2, 4-11.

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J.Mol.Biol. 305, 3, 567-580.

Krupovič, M., Bamford, D.H., 2010. Order to the viral universe. J.Virol. 84, 24, 12476-12479.

Kukkaro, P., Bamford, D.H., 2009. Virus-host interactions in environments with a wide range of ionic strengths. Environ.Microbiol.Rep. 1, 1, 71-77.

Langlet, J., Gaboriaud, F., Duval, J.F., Gantzer, C., 2008. Aggregation and surface properties of F-specific RNA phages: implication for membrane filtration processes. Water Res. 42, 10, 2769-2777.

Li, M., Wang, R., Zhao, D., Xiang, H., 2014. Adaptation of the *Haloarcula hispanica* CRISPR-Cas system to a purified virus strictly requires a priming process. Nucleic Acids Res. 42, 4, 2483-2492.

Liu, Y., Wang, J., Liu, Y., Wang, Y., Zhang, Z., Oksanen, H.M., Bamford, D.H., Chen, X., 2015. Identification and characterization of SNJ2, the first temperate pleolipovirus integrating into the genome of the SNJ1-lysogenic archaeal strain. Mol.Microbiol. 98, 6, 1002-1020.

Luo, H., Zhang, C.T., Gao, F., 2014. Ori-Finder 2, an integrated tool to predict replication origins in the archaeal genomes. Front.Microbiol. 5, 482.

Lupas, A., Van Dyke, M., Stock, J., 1991. Predicting coiled coils from protein sequences. Science. 252, 5009, 1162-1164.

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Bryant, S.H., 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res. 43, Database issue, D222-6.

Marguet, E., Gaudin, M., Gauliard, E., Fourquaux, I., le Blond du Plouy, S., Matsui, I., Forterre, P., 2013. Membrane vesicles, nanopods and/or nanotubes produced by hyperthermophilic archaea of the genus *Thermococcus*. Biochem.Soc.Trans. 41, 1, 436-442.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. 17, 1.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P., Lopez, R., 2013. Analysis Tool Web Services from the EMBL-EBI. Nucleic Acids Res. 41, Web Server issue, W597-600.

Mei, Y., He, C., Huang, Y., Liu, Y., Zhang, Z., Chen, X., Shen, P., 2015. Salinity regulation of the interaction of halovirus SNJ1 with its host and alteration of the halovirus replication strategy to adapt to the variable ecosystem. PLoS One. 10, 4, e0123874.

Mylvaganam, S., Dennis, P.P., 1992. Sequence heterogeneity between the two genes encoding 16S rRNA from the halophilic archaebacterium *Haloarcula marismortui*. Genetics. 130, 3, 399-410.

Nuttall, S.D., Dyall-Smith, M.L., 1993. HF1 and HF2: novel bacteriophages of halophilic archaea. Virology. 197, 2, 678-684.

O'Connell, K.L., Stults, J.T., 1997. Identification of mouse liver proteins on two-dimensional electrophoresis gels by matrix-assisted laser desorption/ionization mass spectrometry of in situ enzymatic digests. Electrophoresis. 18, 3-4, 349-359.

Oksanen, H.M., Pietilä, M.K., Senčilo,A., Atanasova,N.S., Roine, E., Bamford, D.H., 2012. Virus universe: Can it be constructed from a limited number of viral architectures, in: Witzany, G. (Ed.), Viruses: Essential Agents of Life. Dordrecht: Springer Science+Business Media, pp. 83-105.

Olkkonen, V.M., Bamford, D.H., 1989. Quantitation of the adsorption and penetration stages of bacteriophage φ6 infection. Virology. 171, 1, 229-238.

Oren, A., 1983. *Halobacterium sodomense* sp. nov., a Dead Sea halobacterium with an extremely high magnesium requirement. Int.J.Syst.Bacteriol. 33, 2, 381-386.

Oren, A., Ginzburg, M., Ginzburg, B.Z., Hochstein, L.I., Volcani, B.E., 1990. *Haloarcula marismortui* (volcani) sp. nov., nom. rev., an extremely halophilic bacterium from the Dead Sea. Int.J.Syst.Bacteriol. 40, 2, 209-210.

Oren, A., Ventosa, A., Gutierrez, M.C., Kamekura, M., 1999. *Haloarcula quadrata* sp. nov., a square, motile archaeon isolated from a brine pool in Sinai (Egypt). Int.J.Syst.Bacteriol. 49 Pt 3, 1149-1155.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat.Methods. 8, 10, 785-786.

Pietilä, M.K., Roine, E., Paulin, L., Kalkkinen, N., Bamford, D.H., 2009. An ssDNA virus infecting archaea: a new lineage of viruses with a membrane envelope. Mol.Microbiol. 72, 2, 307-319.

Pietilä, M.K., Atanasova, N.S., Manole, V., Liljeroos, L., Butcher, S.J., Oksanen, H.M., Bamford, D.H., 2012. Virion architecture unifies globally distributed pleolipoviruses infecting halophilic archaea. J.Virol. 86, 9, 5067-5079.

Pietilä, M.K., Laurinavičius, S., Sund, J., Roine, E., Bamford, D.H., 2010. The single-stranded DNA genome of novel archaeal virus *Halorubrum* pleomorphic virus 1 is enclosed in the envelope decorated with glycoprotein spikes. J.Virol. 84, 2, 788-798.

Pietilä, M.K., Laurinmäki, P., Russell, D.A., Ko, C.C., Jacobs-Sera, D., Butcher, S.J., Bamford, D.H., Hendrix, R.W., 2013a. Insights into head-tailed viruses infecting extremely halophilic archaea. J.Virol. 87, 6, 3248-3260.

Pietilä, M.K., Laurinmäki, P., Russell, D.A., Ko, C.C., Jacobs-Sera, D., Hendrix, R.W., Bamford, D.H., Butcher, S.J., 2013b. Structure of the archaeal head-tailed virus HSTV-1 completes the HK97 fold story. Proc.Natl.Acad.Sci.U.S.A. 110, 26, 10604-10609.

Pietilä, M.K., Roine, E., Senčilo, A., Bamford, D.H., Oksanen, H.M., 2016. *Pleolipoviridae*, a newly proposed family comprising archaeal pleomorphic viruses with single-stranded or double-stranded DNA genomes. Arch.Virol. 161, 1, 249-56.

Porter, K., Kukkaro, P., Bamford, J.K.H., Bath, C., Kivelä, H.M., Dyall-Smith, M.L., Bamford, D.H., 2005. SH1: A novel, spherical halovirus isolated from an Australian hypersaline lake. Virology. 335, 1, 22-33.

Roberts, M.F., 2004. Osmoadaptation and osmoregulation in archaea: update 2004. Front.Biosci. 9, 1999-2019.

Roine, E., Kukkaro, P., Paulin, L., Laurinavicius, S., Domanska, A., Somerharju, P., Bamford, D.H., 2010. New, closely related haloarchaeal viral elements with different nucleic acid types. J.Virol. 84, 7, 3682-3689.

Schnabel, H., Zillig, W., Pfaffle, M., Schnabel, R., Michel, H., Delius, H., 1982. *Halobacterium halobium* phage φH. EMBO J. 1, 1, 87-92.

Senčilo, A., Paulin, L., Kellner, S., Helm, M., Roine, E., 2012. Related haloarchaeal pleomorphic viruses contain different genome types. Nucleic Acids Res. 40, 12, 5523-5534.

Svirskaitė, J., Oksanen, H.M., Daugelavičius, R., Bamford, D.H., 2016. Monitoring physiological changes in haloarchaeal cell during virus release. Viruses. 8, 3, 10.3390/v8030059.

Takashina, T., Hamamoto, T., Otozai, K., Grant, W.D., Horikoshi, K., 1990. *Haloarcula japonica* sp. nov., a new triangular halophilic archaebacterium. Syst.Appl.Microbiol. 13, 2, 177-181.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol.Biol.Evol. 28, 10, 2731-2739.

Torreblanca, M., Rodriguez-Valera, F., Juez, G., Ventosa, A., Kamekura, M., Kates, M., 1986. Classification of non-alkaliphilic halobacteria based on numerical taxonomy and polar lipid composition, and description of *Haloarcula* gen. nov. and *Haloferax* gen. nov. Syst.Appl.Microbiol. 8, 1, 89-99.

Torsvik, T., Dundas, I.D., 1980. Persisting phage infection in *Halobacterium salinarium* str. 1. J.Gen.Virol. 47, 1, 29-36.

Wais, A.C., Daniels, L.L., 1985. Populations of bacteriophage infecting *Halobacterium* in a transient brine pool. FEMS Microbiol.Ecol. 1, 6, 323-326.

Figure legends

Figure 1. HHPV3 life cycle. (A) Adsorption of HHPV3 to *Har. hispanica* cells at 37 °C. Error bars represent standard deviation (n = 3). (B, C) HHPV3 infection cycle. (B) Growth curves of uninfected (open circles) and infected (closed circles) *Har. hispanica* cultures. Bars represent free progeny viruses. (C) Viable cells in uninfected (open squares) and infected (closed squares) cultures. X axis is the same in (B) and (C). MOI was 15, and cultures were grown aerobically at 37 °C. At 3 h p.i. (arrow in B), the cultures were washed to remove unadsorbed virus particles.

Figure 2. Effect of NaCl on infectivity of (A) HHPV3, (B) HRPV-1, (C) HRPV-3, and (D) HGPV-1. Virus stocks were diluted in (A) 26% SW pH 7.5, (B) optimal HRPV-1 buffer (Pietilä *et al.*, 2009), (C and D) optimal buffers for HRPV-3 and HGPV-1 (Pietilä *et al.*, 2012) at different NaCl molarities (5.5 M is saturated). Infectivity was determined by plaque assay after 3 h (open circles) and 24 h (closed circles) incubation at 4 °C. For HHPV3 (A) infectivity was also determined after 24 h incubation in 26% SW pH 6.1 (triangles) and pH 9.0 (squares) containing 1.5 M NaCl.

Figure 3. Purification of HHPV3. (A) Purification profile of the 1× purified HHPV3 after CsCl equilibrium centrifugation. Fraction No 1 is the surface of the tube, and fraction No 5 contains the virus light-scattering zone. Absorbance (A_{260}, black dots); density (g/ml, grey dots); number of infective particles (PFU/ml, bars). (B) TEM of 2× purified HHPV3 particles stained with 3% (w/v) uranyl acetate.

Figure 4. HHPV3 structural proteins and lipids (1× purified particles) analyzed in 16% polyacrylamide-tricine-SDS gel stained with (A) Coomassie blue (for proteins) and Sudan Black B (for lipids, arrow). (B) Pro-Q Emerald 300 (for glycoproteins only), and (C) SYPRO-Ruby (for proteins). Left to right: (M, kDa) CandyCane molecular weight standards containing a mixture of glycosylated (visible in B) and nonglycosylated proteins; HRPV-1 proteins; HHPV3 proteins. HHPV3 major structural proteins (VP1 and VP3, bold) and HRPV-1 proteins (VP3 and VP4) are indicated.

Figure 5. Thin-layer chromatogram of lipids extracted from *Haloarcula hispanica* cells and from 2× purified HHPV3 particles. The major lipid species of *Har. hispanica* (Bamford *et al.*, 2005) are indicated: PG, phosphatidylglycerol; PGP-Me, phosphatidylglycerophosphate methyl ester; PGS, phosphatidylglycerosulfate; TGD, triglycosyl glycerodiether.

Figure 6. (A) HHPV3 genome. Genes and ORFs are shown as arrows. Genes/ORFs belonging to the conserved betapleolipoviral block are colored. Approximate locations of predicted replication

origin and terminus are marked with a solid line and asterisk and a dotted line, respectively. (B) Z-curves (AT, GC, RY, and MK disparity curves) of HHPV3 genome sequence used for prediction or replication origin and terminus. (C) Comparison of HHPV3 genome and the genomes of betapleolipoviruses. ORFs/genes are numbered and the homologs are shown with the same color. Pairwise amino acid identities (%) are shown between the sequences. (D) Comparison of HHPV3 genome and the related proviral regions in the chromosomes of *Haloarcula hispanica* ATCC 33960 and *Haloarcula marismortui* ATCC 434049 (region 1, see Figure S5 for region 2). Homologs and pairwise amino acid identities (%) are indicated as for (C).