Novel haloarchaeal viruses from Lake Retba infecting Haloferax and Halorubrum species

Carolina M Mizuno, Bina Prajapati, Soizick Lucas-staat, Télesphore
Sime-Ngando, Patrick Forterre, Dennis Bamford, David Prangishvili, Mart
Krupovic, Hanna Oksanen

1 # Novel haloarchaeal viruses from Lake Retba infecting *Haloferax* and

2 # *Halorubrum* species

3 Carolina M. Mizuno[1], Bina Prajapati[2#], Soizick Lucas-Staat[1], Telesphore Sime-Ngando[3], Patrick Forterre[1],

4 Dennis H. Bamford[2], David Prangishvili[1], Mart Krupovic[1]*, and Hanna M. Oksanen[2]*

5

6 [1] Unité Biologie Moléculaire du Gène chez les Extrêmophiles, Institut Pasteur, 25 rue du Docteur Roux, 75015

7 Paris , France

8 [2] Molecular and Integrative Biosciences Research Programme, Faculty of Biological and Environmental

9 Sciences, University of Helsinki, Finland

10 [3] CNRS UMR 6023, Université Clermont-Auvergne, Laboratoire "Microorganismes: Génome et

11 Environnement" (LMGE), F-63000 Clermont-Ferrand, France

12

13 *Corresponding authors
14 Mart Krupovic
15 Hanna M Oksanen
16

17 [#] Current affiliation: Cell and Molecular Biology Program, Institute of Biotechnology, University of Helsinki,

18 Finland

19 Running title: Haloarchaeal Viruses of Lake Retba

20

21

22

Summary

The diversity of archaeal viruses is severely undersampled compared to that of viruses infecting bacteria and eukaryotes, limiting our understanding on their evolution and environmental impacts. Here we describe the isolation and characterization of four new viruses infecting halophilic archaea from the saline Lake Retba, located close to Dakar on the coast of Senegal. Three of the viruses, HRPV10, HRPV11 and HRPV12, have enveloped pleomorphic virions and should belong to the family *Pleolipoviridae*, whereas the forth virus, HFTV1, has an icosahedral capsid and a long non-contractile tail, typical of bacterial and archaeal members of the order *Caudovirales*. Comparative genomic and phylogenomic analyses place HRPV10, HRPV11 and HRPV12 into the genus *Betapleolipovirus*, whereas HFTV1 appears to be most closely related to the unclassified *Halorubrum* virus HRTV-4. Differently from HRTV-4, HFTV1 encodes host-derived minichromosome maintenance helicase and PCNA homologs, which are likely to orchestrate its genome replication. HFTV1, the first archaeal virus isolated on a *Haloferax* strain, could also infect *Halorubrum* sp., albeit with an eight-fold lower efficiency, whereas pleolipoviruses nearly exclusively infected autochthonous *Halorubrum* strains. Mapping of the metagenomic sequences from this environment to the genomes of isolated haloarchaeal viruses showed that these known viruses are underrepresented in the available viromes.

Introduction

40

41   Hypersaline environments, where salt concentration is close to saturating, harbor a high number of virus-like

42   particles (VLP) but rather low microbial diversity (Oren, 2002; Pagaling *et al.*, 2007; Sime-Ngando *et al.*, 2011;

43   Ventosa *et al.*, 2015; Roux *et al.*, 2016). Although these environments are dominated by archaea and bacteria,

44   some eukaryotes are present, e.g. the salt-adapted unicellular green alga *Dunaliella salina*, some fungi and

45   yeast, protozoa, and the brine shrimp *Artemia* (Triantaphyllidis *et al.*, 1998; Gunde-Cimerman *et al.*, 2018).

46   Haloarchaeal virus predation is among the most important factors driving the genetic variation of different

47   haloarchaeal species. For instance, most of the differences between the closely related species are mapped

48   to the genes encoding for cell surface structures and their modification, which directly affect virus-host

49   interactions (Cuadros-Orellana *et al.*, 2007; Dyall-Smith *et al.*, 2011; Tschitschko *et al.*, 2018). As of today,

50   more than 100 viruses have been isolated from hypersaline environments of which the majority infect

51   extremely halophilic euryarchaea, all belonging to the class Halobacteria (Tang *et al.*, 2004; Pagaling *et al.*,

52   2007; Atanasova *et al.*, 2015b; Dyall-Smith *et al.*, 2019). Halophilic viruses are well adapted to high salinity

53   and some of them remain infectious even in saturated salt (Demina *et al.*, 2016a). Some viruses can even

54   survive in low salinities, a beneficial trait under changing environmental conditions (Pietilä *et al.*, 2013c).

55   Halophilic archaeal viruses fall into four different morphological groups: spindle-shaped (genus

56   *Salterprovirus*), pleomorphic (family *Pleolipoviridae*), tailless icosahedral (family *Spherolipoviridae*), and

57   tailed icosahedral (order *Caudovirales*) (Pietilä *et al.*, 2014; Pietilä *et al.*, 2016; Demina *et al.*, 2017). However,

58   filamentous and some exceptional morphotypes not reminiscent of any isolated virus, e.g. hairpin-shaped,

59   bacilliform, and chain-like VLPs, have also been visualized in hypersaline water bodies (Oren *et al.*, 1997;

60   Sime-Ngando *et al.*, 2011; Di Meglio *et al.*, 2016). Significant number of archaeal virus genes have no obvious

61   homologs in public sequence databases (Krupovic *et al.*, 2018). Thus, classification of viral sequences in

62   metagenomic datasets is still a challenge due to the high genetic diversity, despite of remarkable advances

63   in next-generation sequencing and bioinformatics. In addition, the relatively low number of described

64   halophilic archaeal virus isolates with determined complete genome sequences limits the utility of sequence

65   similarity-based analyses. For example, only few similarities to the NCBI non-redundant database at either

66   the nucleotide or amino acid level were reported following the analyses of the metaviromes from hypersaline

67   Lake Retba (Sime-Ngando *et al.*, 2011; Roux *et al.*, 2016) and halite endoliths in the Atacama Desert (Crits-

68   Christoph *et al.*, 2016).

69   Two groups of haloarchaeal viruses have been frequently isolated from geographically remote locations.

70   These include members of the family *Pleolipoviridae* and the order *Caudovirales*, respectively. Membrane-

71   containing virions of pleolipoviruses resemble extracellular membrane vesicles able to carry the virus

72   genome from host to host (Pietilä *et al.*, 2012). They represent a unique archaeal virus group with a recently

73   established taxonomic position as the first family containing viruses with either single-stranded (ss) or

74   double-stranded (ds) DNA genomes (linear or circular forms) (Pietilä *et al.*, 2016; Bamford *et al.*, 2017). The

75   non-lytic pleomorphic virus life-cycle starts with fusion between the viral and host membranes and they exit

76   the host cell most probably by budding, preserving the host membrane integrity (Svirskaite *et al.*, 2016; El

77   Omari *et al.*, 2019). Their simplistic mechanism of nucleic acid transmission resembles the function of the

78   recently described "infectious" plasmid membrane vesicles isolated from Antarctic species of haloarchaea

79   (Erdmann *et al.*, 2017), supporting the tight evolutionary relationships between viruses and non-viral mobile

80   genetic elements (Iranzo *et al.*, 2016b). Pleolipoviruses share a conserved core of four to five genes, mainly

81   encoding major structural proteins of which one is the spike protein responsible for host attachment and

82   membrane fusion (Pietilä *et al.*, 2012; Sencilo *et al.*, 2012; El Omari *et al.*, 2019).

83   The tailed icosahedral dsDNA viruses represent the most numerous archaeal virus group described today

84   (Atanasova *et al.*, 2015a). Intriguingly, all these viruses infect halophiles or methanogens of the phylum

85   Euryarchaeota (Prangishvili *et al.*, 2017). However, the identified proviruses and metagenomic studies

86   suggest wider association of archaeal tailed viruses across different orders within the Euryarchaeota, but also

87   with members of the phylum Thaumarchaeota (Krupovic *et al.*, 2010a; Krupovic *et al.*, 2011b; Danovaro *et*

88   *al.*, 2016; Philosof *et al.*, 2017; Vik *et al.*, 2017; Abby *et al.*, 2018; Ahlgren *et al.*, 2019; Lopez-Perez *et al.*,

89   2019). This group of viruses shares the same architectural principles with the icosahedral tailed dsDNA

90   bacteriophages of the order *Caudovirales* (Pietilä *et al.*, 2013b). All three different tail structures initially

91   characterized for bacteriophages, have been found among the archaeal viruses: long contractile

92   (myoviruses), long non-contractile (siphoviruses), and short non-contractile tails (podoviruses) (Atanasova *et*

93   *al.*, 2012). Genomes of archaeal caudoviruses are mosaics of genes with different evolutionary histories and

94   their gene contents and genome lengths differ considerably, and consequently, also differ their capsid sizes,

95   making this group genetically very diverse (Krupovic *et al.*, 2010a; Pietilä *et al.*, 2013c; Sencilo *et al.*, 2013;

96   Dyall-Smith *et al.*, 2019). Although their capsid structures are very conserved, their receptor binding proteins

97   have a high genetic plasticity allowing them to adapt to new hosts. Particularly, archaeal myoviruses with

98   contractile tails have very broad host ranges crossing the genus boundary (Atanasova *et al.*, 2012; Atanasova

99   *et al.*, 2015c). Furthermore, several myoviruses were shown to encode an invertible tail fiber gene module,

100   which allows these viruses to alternate between different variants of the tail fiber proteins with distinct host

101   specificities (Rossler *et al.*, 2004; Dyall-Smith *et al.*, 2018; Dyall-Smith *et al.*, 2019).

102   Both culture-independent and culture-dependent approaches indicate that haloviruses represent a globally

103   distributed reservoir of orphan genes encoding novel functions (Aalto *et al.*, 2012; Atanasova *et al.*, 2012;

104   Roux *et al.*, 2016). In addition, many halophilic archaea carry proviruses in their chromosomes (Krupovic *et*

105   *al.*, 2010a; Dyall-Smith *et al.*, 2011; Makarova *et al.*, 2014; Liu *et al.*, 2015; Maier *et al.*, 2015; Demina *et al.*,

106   2016a; Atanasova *et al.*, 2018b). The co-evolution of viruses and host cells in the presence of high

107   recombination frequency in halophilic microbes have resulted in a globally distributed complex network of

108   viruses, proviruses, membrane vesicle, transposons, and plasmids sharing the common genetic pool and

109   displaying dynamic interplay across time and space (Zhang *et al.*, 2012; Forterre *et al.*, 2014; Atanasova *et*

110   *al.*, 2015c; Liu *et al.*, 2015; Iranzo *et al.*, 2016b; Atanasova *et al.*, 2018b; Dyall-Smith and Pfeiffer, 2018; Wang

111   *et al.*, 2018a; Wang *et al.*, 2018b). Here we report on the isolation of four new haloarchaeal viruses from

112   saline Lake Retba. Morphological and genomic characterization of these viruses allowed their tentative

113   taxonomic assignments. The siphovirus HFTV1, to the best of our knowledge, is the first virus isolated on a

114   *Haloferax* strain, and it should belong in the order *Caudovirales* with other archaeal and bacterial tailed

115   dsDNA viruses. Bacterial and archaeal caudoviruses together with eukaryotic herpesviruses form the HK97–

116   like virus lineage (Abrescia *et al.*, 2012). The pleomorphic archaeal viruses might belong to a tentative new

117   virus lineage comprising membrane vesicle-like archaeal viruses of the family *Pleolipoviridae*. Collectively,

118   our results further expand the knowledge on the genomic diversity and host range of haloarchaeal viruses

119   and provide insights into their genome evolution.

Results and Discussion

Isolation of novel archaeal viruses of *Halorubrum* and *Haloferax*

The viruses designated HRPV10, HRPV11, HRPV12 and HFTV1 were isolated together with their host strains

from saline Lake Retba (14°50'14'' N, 17°14'55'' W), close to Dakar, the capital of Senegal, in May, 2011 (Table

1). Sample LR1 collected from the center of the Lake Retba contained grey water and a grey sediment mixed

with salt, whereas sample LR2 (purple water with white sediment) was collected close to the site where salt

is collected for trade. The salinity of the LR1 and LR2 samples was 290-300 g/L and 250 g/L, respectively,

whereas the temperature (27°C) and pH (7.8) were the same at both sampling sites.

The haloarchaeal isolates were obtained by directly plating of the samples on MGM plates (see Methods).

The obtained colonies were colony-purified on solid media. The pure cultures of the isolated halophilic

archaeal strains (19 in total) were identified as members of the class *Halobacteria* by partial 16S rRNA gene

sequence analysis (Fig. 1, Table 2). The isolates belong to three of the six families of the class Halobacteria:

Halorubraceae (11 isolates), Haloferacaceae (7 isolates) and Halobacteriaceae (1 isolate). All isolates from

Halorubraceae were identified as *Halorubrum* spp., eight of which form a clade with *Halorubrum*

lacusprofundi.

LR2-19 clusters with *Hrr. sodomense*. LR1-22 clusters with an uncharacterized species of *Halorubrum* and

LR2-20 did not cluster with any other strain. Among the isolates within Haloferacaceae, all seven were

classified as *Haloferax* spp., clustering with *Hfx. volcanii*. LR2-15, the only representative of Halobacteriaceae,

clustered with *Halomicroarcula limicola* and was classified as *Halomicroarcula* sp. LR2-15. The viruses were

isolated on the endogenous Lake Retba strains (Table 2) using the same Lake Retba samples (see Methods

for details). The host strains of the viruses HRPV10, HRPV11, and HRPV12 are *Halorubrum* spp. LR2-17, LR2-

12, and LR1-23, respectively (Table 1), whereas HFTV1 infects *Haloferax* sp. LR2-5, making HFTV1 the first

known virus isolated on *Haloferax* strain. The defective proviruses identified in *Hfx. mediterranei* (Li *et al.*,

143 2013) and a variant of Halorubrum virus HF1 capable of infecting *Hfx. volcanii* [(Nuttall and Dyall-Smith, 1993)

144 not available to our knowledge, personal communication], are the only reports on *Haloferax* viruses.

145

146 HRPV10, HRPV11, and HRPV12 produce hazy plaques that are 3-10 mm in diameter, whereas the HFTV1

147 plaques are clear (Table 1; Supplementary Fig. 1). The virus isolate plaque morphologies were different from

148 each other, and HRPV10 and HRPV11, in particular, produce very hazy plaques difficult to document as

149 figures, but they are visible in optimal lightning conditions (Supplementary Fig. 1A and B). The plaque

150 morphologies of the HRPV10, HRPV11, and HRPV12 resemble the plaques of the members in the family

151 *Pleolipoviridae* suggesting that also the plaques of the new pleomorphic virus isolates have non-lytic life cycle

152 and the plaque is a consequence of the host cell growth retardation due to the virus infection (Pietilä *et al.*,

153 2009; Svirskaite *et al.*, 2016). The plaques were purified by three consecutive time to obtain pure virus

154 cultures (see Methods). The virus stocks gave typical titers of 10^{11}-10^{12} pfu/ml (Table 1), suggesting that they

155 might be promising model systems for studies on haloarchaeal virus functions and virus-host interactions.

156 Infectivity of the viruses remained unchanged at 4 °C for a period of four weeks (data not shown).

157 For virus purification, virions were collected from the virus stocks by using two-step polyethylene glycol-NaCl

158 precipitation and purified to near homogeneity by rate zonal centrifugation in sucrose followed by

159 equilibrium centrifugation in CsCl. In the case of HRPV10, HRPV11 and HRPV12, this approach yielded highly

160 pure virion preparations based on the specific infectivities (2-5×10^{13} pfu/mg of protein; Table 3), negative

161 staining and transmission electron microscopy (TEM) analysis of the purified particles (Fig. 2A-C), and SDS-

162 PAGE gel analysis (Fig. 3A-C). Specific infectivities and protein patterns of the purified HRPV10, HRPV11, and

163 HRPV12 viruses were comparable with data reported for pleomorphic viruses purified by using the

164 comparable precipitation and preparative ultracentrifugation techniques (e.g. viruses HRPV-1, HRPV-2,

165 HRPV-3, HRPV-6, HHPV-1, His2, and HHPV4) yielding highly pure virus material (specific infectivities 2-5×10^{13}

166 pfu/mg of protein) (Pietilä *et al.*, 2012; Atanasova *et al.*, 2018b). HFTV1 virus particles were purified in high

167 numbers based on TEM (Fig. 3D) and protein quantities (Table 3) but the purified particles had specific

168   infectivity of ~2×10^9 pfu/mg of protein (Table 3), which is 3-4 magnitudes lower than e.g. the specific

169   infectivities of the purified virus samples of haloarchaeal tailed virus HSTV-1 (~9×10^{12} pfu/mg of protein) and

170   icosahedral membrane-containing virus HCIV-1 (~1×10^{12} pfu/mg of protein), of which have been analyzed

171   structurally (Pietilä *et al.*, 2013b; Demina *et al.*, 2016b; Santos-Perez *et al.*, 2019). The negative staining and

172   TEM of the purified HFTV1 particles revealed that some of particles had lost their genome explaining partly

173   the loss of infectivity (Fig. 2D).

174

175   **Membrane vesicle-like virions of HRPV10, HRPV11 and HRPV12**

176   The purified HRPV10, HRPV11 and HRPV12 virions were tailless round particles with a diameter of ~55 nm

177   (Fig. 2A-C). The virion morphologies resembled one another and those of viruses in the family *Pleolipoviridae*

178   (Pietilä *et al.*, 2012). All three virion types equilibrated in CsCl density gradients (mean density of 1.30-1.35

179   g/ml) suggested that they contain lipids as one of their structural components. HRPV11 and HRPV12 were

180   sensitive to chloroform, a widely used organic solvent, whereas HRPV10 was resistant (Table 1). Infectivity

181   of all three viruses in the presence of non-ionic detergents Nonidet P-40 or Triton X-100 decreased by 7-11

182   orders of magnitude (Table 1).

183   The lipid compositions of the viruses and their host strains were verified by thin-layer chromatography and

184   ammonium molybdate staining. The major polar lipids of *Haloarcula hispanica* – phosphatidylglycerol (PG),

185   phosphatidylglycerophosphate methyl ester (PGP-Me), phosphatidylglycerosulfate (PGS), and triglycosyl

186   glycerodiether (TGD) – have been previously identified (Bamford *et al.*, 2005) and were used as a control (Fig.

187   3A-C). The patterns of lipid species of *Halorubrum* sp. LR2-17, LR2-12, and LR1-23 were identical to each

188   other, containing probably PG, PG-Me and PGS (Fig. 3A-C). In all three viruses, lipids were found to be a

189   structural component of the virions (Fig. 3A-C). The virus lipid profiles were identical with each other and

190   resembled the lipid profiles of their respective hosts suggesting that viruses use non-selective lipid uptake as

191   also shown previously for other pleolipoviruses (Pietilä *et al.*, 2010; Pietilä *et al.*, 2012). Virions of all three

192   viruses contained two major structural protein species, which were ~60 kDa and ~7 kDa in mass when

193   resolved in polyacrylamide gel (Fig. 3D-F). The patterns were different from each other and the major

194   structural proteins were identified based on the gene homology to those of other pleolipoviruses (see below).

195

Three new pleomorphic viruses are members of the genus *Betapleolipovirus*

197   The nucleic acids extracted from the purified virions of HRPV10, HRPV11, and HRPV12 were sensitive to RQ1

198   DNase and Exonuclease III, resulting in complete degradation or extensive fragmentation, respectively. Mung

199   bean nuclease, which is specific to ssDNA (Fig. 4A), digested the HRPV10, HRPV11, and HRPV12 genomes into

200   discrete fragments (Fig. 4B-D), suggesting that the genomes are dsDNA molecules with nicks or single-

201   stranded regions. Sequencing of the genomes and assembly of the reads using a *de novo* assembly algorithm

202   (with default parameters) implemented in the CLC Genomics Workbench (QIAGEN Bioinformatics) resulted

203   in single contigs for each genome. Each contig contained direct terminal repeats of 18-20 bp. The sequencing

204   reads overlapping both termini were as abundant as those from other genomic positions, indicating that the

205   genomes are circular. Consistently, assembly of the sequencing reads with the SPAdes algorithm (Bankevich

206   *et al.*, 2012) resulted in contigs identical to those assembled with CLC Genomics Workbench, albeit with

207   alternative start positions and terminal repeats, confirming the circular nature of the genomes

208   (Supplementary Fig. 2).

209   To assess the relationship between the 3 new pleolipoviruses and the previously described members of the

210   family *Pleolipoviridae,* we calculated intergenomic distances between pairs of viruses by pairwise

211   comparisons of their nucleotide sequences and constructed the corresponding phylogenomic tree using

212   VICTOR(Fig. 5A), a Genome BLAST Distance Phylogeny method (Meier-Kolthoff and Goker, 2017). The

213   clustering of the pleolipoviruses in the resultant tree is consistent with previous classification based on the

214   presence of genes for the putative replication-initiation or DNA polymerase proteins and relatedness of the

215   VP3-like proteins (Pietilä *et al.*, 2016). HRPV10, HRPV11 and HRPV12 form a well-supported clade with

216   members of the genus *Betapleolipovirus*, namely, *Halorubrum* virus HRPV-3 and *Halogeometricum* virus

217 HGPV-1. The clade also includes *Haloarcula* virus HHPV3 and *Natrinema* virus SNJ2, two tentative members

218 of the *Betapleolipovirus* genus (Liu *et al.*, 2015; Bamford *et al.*, 2017).

219 The genomes of HRPV10, HRPV11, and HRPV12 were 9296, 9368, and 9944 bp in length, respectively, and

220 their GC% contents were 55.2-55.7%. Genomes were predicted to contain 13-16 ORFs oriented in both

221 transcriptional directions forming at least two putative operons (Fig. 5B; Supplementary Table 1). Genomes

222 of HRPV10, HRPV11 and HRPV12 are very similar to each over (92-95% nucleotide identity over the whole

223 length), but different from other characterized pleolipoviruses. The most closely related virus HGPV1 shares

224 68% identity over just 14% of the genome as determined by BLASTN. Consistently, comparison of the

225 HRPV10, HRPV11 and HRPV12 proteins against the proteomes of all other known pleolipoviruses has

226 revealed betapleolipoviruses as the closest relatives, with the largest number of sequence matches to

227 betapleolipoviruses HGPV-1 (gene *2*, ORF5, ORF6, ORF7, ORF9, ORF13) and HRPV3 (ORF12 and ORF9),

228 whereas ORF4 was most similar to the homolog encoded by alphapleolipovirus HHPV-2 (Supplementary

229 Table 2). The presence of the signature gene encoding the putative replication protein (ORF11 in HRPV10)

230 unequivocally relates HRPV10, HRPV11 and HRPV12 to betapleolipoviruses. However, single-gene

231 phylogenies reconstructed for the core proteins, namely, spike protein (El Omari *et al.*, 2019) (Supplementary

232 Fig. 3A) and the putative NTPase (Supplementary Fig. 3B), were not entirely consistent with this assignment

233 (Fig. 5A), most likely reflecting occasional recombination between pleolipoviruses belonging to different

234 genera, consistent with previous observations (Wang *et al.*, 2018a). Comparison of HRPV ORFs to the non-

235 redundant protein sequence database reveal that the most similar sequences are found in the genomes of

236 *Halorubrum coriense*, *Halorubrum terrestre* and *Halorubrum* sp. T3, indicating the presence of related

237 proviruses within these organisms (Supplementary Table 1). Proviruses related to pleolipoviruses have been

238 described in haloarchaeal strains (Liu *et al.*, 2015; Demina *et al.*, 2016a; Atanasova *et al.*, 2018a; Wang *et al.*,

239 2018a).

240 The close genetic similarity between HRPV10, HRPV11 and HRPV12 allows tracing the evolutionary events

241 which took place in a relatively recent past. In particular, HRPV11 and HRPV12 share two small genes

11

242 (HRPV11-ORF9 and HRPV12-ORF9; HRPV11-ORF14 and HRPV12-ORF15), encoding putative DNA-binding

243 proteins carrying zinc-binding domains, which are absent in HRPV10, whereas ORF13 of HRPV12 is not found

244 in the two other viruses (Supplementary Table 1). Notably, the closest homolog of the latter gene is encoded

245 by an uncultivated tailed haloarchaeal virus eHP-27 (51% identity; E=3e-57) (Garcia-Heredia *et al.*, 2012),

246 followed by homologs from diverse haloarchaea. Given that HRPV12 ORF13, which encodes a putative

247 AdoMet-dependent methyltransferase (Supplementary Table 1), is not present in any other pleolipovirus

248 (Fig. 5B), in all likelihood, it has been introduced into the HRPV12 genome horizontally from an unrelated

249 haloarchaeal virus, following the divergence of HRPV12 from a common ancestor with HRPV10 and HRPV11.

250 By contrast, the homolog of HRPV11 ORF9 has been apparently lost from the HRPV10 genome due to an

251 inactivating point mutation, resulting in a long intergenic region between ORFs 8 and 9. Furthermore, analysis

252 of the nucleotide similarity pattern along the HRPV10, HRPV11 and HRPV12 genomes uncovered a

253 hypervariable region within ORF4, which encodes for a putative receptor-binding spike protein, one of the

254 two major virion proteins suggested to be involved in host recognition and virus entry (Pietilä *et al.*, 2010).

255 Notably, ORF4 homologs in HRPV11 and HRPV12 do not display appreciable similarity within the central

256 region (Fig. 5B), pinpointing a highly variable protein domain, which is most likely to be critical for host

257 recognition and binding; a similar conservation pattern is also observed in alphapleolipoviruses (e.g.,

258 compare HRPV-2 and HRPV-6 in Fig. 5B).

259 Comparative genomics analysis has shown that besides the five core genes conserved in all pleolipoviruses

260 (except for His2, which contains four core genes), HRPV10, HRPV11 and HRPV12 encode several putative

261 proteins specific to members of the genus *Betapleolipovirus*. These include homologs of HRPV10 ORF8 and

262 ORF11, which are conserved in all currently known betapleolipoviruses, as well as ORF10 and ORF13,

263 conserved in a subset of betapleolipoviruses, but not in viruses from the two other genera (Fig. 5B). Previous

264 sequence analyses did not provide insights into the putative functions of the four conserved proteins. Indeed,

265 HRPV10 ORF11-like proteins, which were suggested to represent replication initiation proteins of

266 betapleolipoviruses (Krupovic *et al.*, 2018), remain recalcitrant to functional annotation based on sequence

267 similarity searches. However, profile-profile comparisons initiated with the sequence of HRPV10 ORF8

268   revealed homology to various PD-(D/E)XK family nuclease, including type II restriction endonucleases

269   (Supplementary Table 1). Notably, the protein is not restricted to betapleolipoviruses, but is also conserved

270   in several other groups of unrelated haloarchaeal viruses, including members of the *Caudovirales* (HHTV-1)

271   and *Sphaerolipoviridae* (SH1, PH1, HCIV-1, HHIV-2). HRPV10 ORF10 and ORF13 encode putative DNA-binding

272   proteins with winged helix-turn-helix and ribbon-helix-helix domains, respectively (Supplementary Table 1),

273   and may be involved in transcriptional regulation of the viral and/or host genes.

274

275   The first virus isolated on *Haloferax* has an icosahedral head, non-contractile tail and circularly

276   permuted dsDNA genome

277   Micrographs of the purified HFTV1 virions revealed icosahedral particles with a long non-contractile tail

278   typical of the siphovirus morphotype (Fig. 2D). The diameter of the head was ~50 nm and the tail length was

279   ~60 nm. The major protein species of HFTV1 virions were approximately 50, 40, 22 and 16 kDa in mass (Fig.

280   3G). The infectivity of HFTV1 in the presence of chloroform, Nonidet P-40, or Triton X-100 remained

281   unchanged (Table 1), suggesting that the virion does not contain a membrane moiety and consists only of

282   proteins and nucleic acid.

283   The nucleic acid extracted from purified HFTV1 virions was sensitive to RQ1 DNase treatment, but resistant

284   to Exonuclease III and Mung bean nuclease, indicating that the genome is a dsDNA molecule (Supplementary

285   Fig. 4). Genome sequencing and read assembly were performed as described above for pleolipoviruses and

286   yielded a 38,059 bp-long circular contig (GC% ~54%), which appears to represent a complete viral genome.

287   A total of 70 ORFs were predicted in the HFTV1 genome using Prodigal (Hyatt *et al.*, 2010), of which 28 (40%)

288   did not have any clear homologs in the public databases (Supplementary Table 3, Supplementary Fig. 5). Half

289   (35) of the gene products had sequence similarity (35-84% identity; Supplementary Table 3) to haloviruses:

290   15 to halophilic archaeal siphovirus HRTV-4 originating from a salt water sample from Margherita di Savoia,

291   Italy (Sencilo *et al.*, 2013), and 20 to uncultivated environmental haloviruses identified in the solar saltern of

292   Santa Pola, Spain (Garcia-Heredia *et al.*, 2012). The remaining 10% of the genes had closest homologs

293 encoded in cellular organisms. Namely, the most significant similarities were shared with archaea from the

294 order Halobacteriales (*Natrialba, Natronobacterium, Haloarcula, Halococcus* and *Haloterrigena*) and one to

295 *Cellulophaga baltica*, a marine bacterium from the order Flavobacteriales (Supplementary Fig. 5). All ORFs

296 but one are arranged in the same transcriptional direction (Fig. 6).

297 To determine the packaging mechanism employed by HFTV1, we analyzed the bias in distribution of the

298 1,657,094 sequencing reads along the HFTV1 genome using PhageTerm, a tool that relies on the detection

299 of biases in the number of sequencing reads observable at natural DNA termini compared with the rest of

300 the viral genome (Garneau *et al.*, 2017). The analysis revealed a pattern of sequencing read coverage

301 consistent with a circularly permuted, terminally redundant genome and headful packaging mechanism

302 initiated from a *pac* site, similar to that of bacteriophage P1 (Supplementary Fig. 6). Consistently, phylogeny

303 of the large subunit of the terminase (Supplementary Fig. 3B), an enzyme responsible for genome packaging

304 in bacterial and archaeal members of the order *Caudovirales*, revealed a relatively close relationship of HFTV1

305 to *Halorubrum* virus HRTV-4, a siphovirus for which the genome was also found to be circularly permuted

306 (Sencilo *et al.*, 2013).

307 The phylogenomic analysis using VICTOR (Meier-Kolthoff and Goker, 2017) confirmed the relationship of

308 HFTV1 with *Halorubrum* virus HRTV-4, and also revealed relationship to four uncultivated viruses, eHP-1,

309 eHP-15, eHP-19 and eHP-34 (Garcia-Heredia *et al.*, 2012), for which the hosts have not been previously

310 predicted (Supplementary Fig. 7A). The genomes of the latter viruses are generally collinear with those of

311 HFTV1 and HRTV-4. The highest sequence similarity between the genomes is observed within the genes

312 encoding for putative virion morphogenesis proteins, such as the major capsid protein, the large subunit of

313 the terminase and tail proteins (Fig. 6). Given this genomic conservation, we predict that the uncultivated

314 HFTV1-like viruses eHP-1, eHP-15, eHP-19 and eHP-34 infect halophilic archaea.

315 The genome of HFTV1 encodes several proteins putatively involved in DNA metabolism, namely a replicative

316 minichromosome maintenance (MCM) helicase (gp58), DNA polymerase sliding clamp protein PCNA (gp64),

317 DNA methyltransferase (gp61) and Rad52-like recombinase (gp50) (Fig. 6, Supplementary Table 3). The MCM

318   is the principal helicase responsible for unwinding of the dsDNA duplex during chromosomal replication in

319   archaea and eukaryotes (Bell and Botchan, 2013). MCM homologs have been previously identified in archaeal

320   viruses and plasmids with moderately-sized genomes (20-50 kb) (Krupovic *et al.*, 2018) and phylogenetic

321   analyses have suggested that mobile genetic elements have horizontally acquired the *mcm* genes from

322   cellular organisms on multiple independent occasions (Krupovic *et al.*, 2010b). The PCNA sliding clamp is

323   another key replication protein in archaea and eukaryotes and is known as a "molecular tool-belt" due to its

324   interaction with multiple other proteins involved in DNA replication and repair, including replicative DNA

325   polymerase, DNA ligase, replication factor C, Flap Endonuclease 1 (FEN1) and RNase H (Pan *et al.*, 2011).

326   Similar to MCM helicases, PCNA homologs have been previously identified in some haloarchaeal virus

327   genomes (Raymann *et al.*, 2014), whereas certain other archaeal viruses have been shown to specifically

328   recruit the host PCNA for the replication of their genomes (Gardner *et al.*, 2014). Thus, the virus-encoded

329   MCM and PCNA homologs are likely to orchestrate the replication of the HFTV1 genome.

330   Despite the synteny within the morphogenetic gene modules of HFTV1, HRTV-4, eHP-1, eHP-15, eHP-19, and

331   eHP-34, the genome replication modules of these viruses appear to be very different. Namely, among the

332   five viruses, only HFTV1 encodes both MCM helicase and PCNA. Notably, there is only one other known

333   archaeal halophilic virus, podovirus HSTV-1, which harbors genes for both proteins in its genome (Pietilä *et*

334   *al.*, 2013b; Raymann *et al.*, 2014). By contrast, the MCM helicase is encoded only by eHP-34, whereas HRTV-4,

335   the closest relative of HFTV1, as well as eHP-1, eHP-15 and eHP-19 do not encode either of the two replication

336   proteins. These observations reaffirm that virion formation and genome replication are uncoupled processes

337   and evolve independently (Krupovic and Bamford, 2010), as is also evident in the case of pleolipoviruses,

338   where viruses from the three genera encode non-homologous genome replication proteins (Krupovic *et al.*,

339   2018). Consequently, viral genomes are often mosaics of genes with different evolutionary histories (Juhala

340   *et al.*, 2000; Pope *et al.*, 2015; Iranzo *et al.*, 2016a; Yutin *et al.*, 2018).

341   Genetic mosaicism in tailed bacteriophage genomes is thought to be generated by illegitimate recombination

342   (Krupovic *et al.*, 2011a) or relaxed homologous recombination (De Paepe *et al.*, 2014). The former occurs at

343 essentially random positions within the genome, with nonviable recombinants being purged by natural

344 selection (Pedulla *et al.*, 2003). The latter process involves promiscuous phage-encoded recombinases, such

345 as phage λ recombinase Redβ, which catalyze homologous recombination by annealing short and diverged

346 sequences (De Paepe *et al.*, 2014). Among the phage recombinases, the Rad52-like family is by far the largest

347 and most diversified (Lopes *et al.*, 2010). Interestingly, HFTV1 encodes a divergent member (gp50) of the

348 Rad52-like family of recombinases (Rad52, PDB profile 5JRB_A, HHpred probability of 94%), which might

349 facilitate genome remodeling in the replication modules of HFTV1-like viruses. Homologs of HFTV1 gp50 are

350 also encoded by HRTV-4, eHP-1, eHP-15, and eHP-34 as well as by several other uncultivated halophilic

351 viruses. Notably, the closest homologs of the HFTV1 PCNA are encoded by cellular organisms, suggesting that

352 the corresponding gene has been acquired by HFTV1 from halophilic archaea, rather than inherited from a

353 common ancestor shared with other viruses. Similarly, the closest homologs of the orthologous HFTV1 and

354 eHP-34 MCM helicases are encoded by halophilic archaea, whereas homologs from other viral groups are

355 more divergent. This suggests that the *pcna* and *mcm* genes have been acquired directly from the hosts in

356 different groups of archaeal viruses on several occasions, independently of each other.

357

358 **Narrow host range of haloarchaeal viruses from Lake Retba**

359 To determine the host range of viruses isolated from Lake Retba, we first tested their infectivity towards the

360 19 autochthonous haloarchaeal strains (Table 2). Despite the overall close genomic similarity, the

361 pleomorphic viruses HRPV10, HRPV11, and HRPV12 were found to have distinct host ranges. HRPV11

362 displayed the broadest host range, being able to infect four different *Halorubrum* strains isolated from Lake

363 Retba, whereas HRPV10 and HRPV12 each could infect only two different strains. *Halorubrum* sp. LR2-12 was

364 susceptible to all three pleomorphic viruses, albeit with highly different efficiencies of plating (EOP) (Table

365 2). Notably, *Haloferax* virus HFTV1 was found to infect hosts across the genus boundary. In addition to its

366 own isolation host, LR2-5, belonging to the genus *Haloferax* (Table 2), HFTV1 was able to infect *Halorubrum*

367 sp. LR1-23, albeit with an eight orders of magnitude lower efficiency (Table 2). Similarly, a previous cross-

368   infectivity study has shown that haloarchaeal siphoviruses generally display genus-restricted host ranges,

369   although some isolates were found to infect hosts belonging to two or three genera (Atanasova *et al.*, 2015c).

370   Next, we set out to explore the infectivity of the four viruses against *Haloferax* and *Halorubrum* strains

371   isolated from geographically remote locations (Israel, Italy, Slovenia, Spain, Thailand and Antarctica). Namely,

372   we tested 41 distinct *Halorubrum* strains originating from 10 different sampling sites and seven *Haloferax*

373   strains from five distinct locations (Supplementary Table 4). Among the 51 strains tested, only *Halorubrum*

374   sp. E200-4 isolated from Eilat, Israel was sensitive to pleomorphic virus HRPV11, albeit with a considerably

375   lower ($\sim 6 \times 10^{-3}$) EOP. This observation is consistent with the previous finding that most pleolipoviruses are

376   highly specific to their isolation hosts (Atanasova *et al.*, 2012; Atanasova *et al.*, 2015c), but also indicates that

377   occasional cross-infections that transcend site and time of isolation are possible. Similar patterns of infection,

378   whereby viruses preferentially infect hosts from the same site rather than hosts isolated from similar but

379   geographically remote sites, are also typical of bacterial virus-host systems from different ecological niches

380   (Vos *et al.*, 2009; Koskella *et al.*, 2011), including hypersaline environments (Villamor *et al.*, 2018). Thus, a

381   pronounced biogeographical pattern emerges in haloarchaeal virus-host interactions, possibly due to

382   increased diversification of the species composition of communities as a function of increasing geographic

383   and environmental distance (Weitz *et al.*, 2013). The specificity of viruses to autochthonous strains seemingly

384   contrasts the conclusions drawn from comparative (meta)genomic analysis of halophilic viral communities

385   which indicated that hypersaline viral communities should be considered as a genetic continuum across

386   continents (Roux *et al.*, 2016). Collectively, the results of the large-scale comparative genomics and local

387   infectivity studies suggest that the gene complements responsible for virion formation and adaptation to

388   environmental conditions are shared by haloarchaeal viruses across the globe, whereas the incessant

389   evolutionary arms race drives local adaptation of viruses and their hosts at a finer scale.

390   Haloarchaeal myovirus isolates appear to display a broader host range (Atanasova *et al.*, 2012; Atanasova *et*

391   *al.*, 2015c) than siphoviruses, such as HFTV1. This tendency appears to be general, because bacterial

392   myoviruses also display broader host range than siphoviruses and podoviruses (Wichels *et al.*, 1998). The

393   broader host range of archaeal myoviruses might be linked to the larger genomes and, accordingly,

394   functionally more diverse gene content (Krupovic *et al.*, 2018) including e.g. many auxiliary genes involved in

395   DNA and RNA metabolism (Sencilo *et al.*, 2013). For instance, HVTV-1 encodes an almost complete replisome

396   (Pietilä *et al.*, 2013c; Kazlauskas *et al.*, 2016), whereas HGTV-1 encodes an RNA ligase and lysyl-tRNA

397   synthetase and has 36 tRNA genes for all universal genetic code amino acids (Sencilo *et al.*, 2013).

398   Presumably, this extended gene baggage renders myoviruses more promiscuous and partly independent of

399   the corresponding cellular machineries compared to viruses with smaller genomes.

400

401   **Scarcity of haloarchaeal virus isolates in the environment**

402   The relative abundance of viruses in any particular sample or environment can be estimated by mapping the

403   sequence reads from a metavirome to the reference genomes and expressed as **R**eads recruited **P**er **K**b of

404   genome per **G**b of metagenome (RPKG). We used this approach to compare the relative abundance of the

405   four viruses described in this study to that of the previously reported cultivated and uncultivated

406   haloarchaeal viruses. To this end, we analyzed saltern viromes sequenced from Lake Retba (Roux *et al.*, 2016)

407   and South Bay Salt Works (Rodriguez-Brito *et al.*, 2010). Notably, the samples for the preparation of the Lake

408   Retba virome (Roux *et al.*, 2016) were collected during the same sampling trip as those used to isolate viruses

409   reported herein. However, none of the cultivated haloarchaeal viruses, including those described here, were

410   sufficiently similar to the sequences present in the available viromes. By contrast, uncultured viruses

411   predicted to infect *Haloquadratum walsbyi* recruited around 15,000 RPKG and formed a distinct clade in the

412   phylogenomic tree (Supplementary Fig. 7). Apart from these, other uncultured viruses with no identified host

413   and one virus predicted to infect nanohaloarchaea have recruitment of around 10 RPKG. The fact that all

414   currently cultured viruses recruit only negligible number of reads, even when the virome originates from the

415   same site as virus isolates is likely to reflect the still scarce and biased sampling of the (halo)archaeal virome.

416   At least in the case of the Lake Retba viruses and the corresponding virome, the two have been isolated at

417   the same time and thus temporal variation in virus diversity cannot explain this result. Given the low

418   abundance of *Halorubrum* spp. in salterns from warm environments (Garcia-Heredia *et al.*, 2012), the

419   currently used culture-based approaches appear to be biased towards isolation of viruses that represent a

420   rather minor fraction of the natural haloarchaeal virome. We note, however, that *Halorubrum* species

421   represent one of the dominant components of the haloarchaeal communities in the cold hypersaline

422   environments, such as Deep Lake in Antarctica (DeMaere *et al.*, 2013); thus, *Halorubrum* viruses might

423   specifically dominate the cold-adapted haloarchaeal viromes. To obtain further insights into the actual

424   diversity of haloarchaeal viruses and to initiate studies on the biology of ecologically relevant virus-host

425   systems, future work should focus on improving the cultivation protocols for the dominant inhabitants of the

426   hypersaline environments, such as *Haloquadratum* spp. (Oh *et al.*, 2010; Dyall-Smith *et al.*, 2011).

427   Nevertheless, further characterization of the "cultivable minority" component of the haloarchaeal virome,

428   as described in this study, provides important insights into the general mechanisms of haloarchaeal virus

429   evolution and might lead to the establishment of virus-host systems in genetically tractable haloarchaeal

430   hosts, such as *Haloferax*, for in-depth studies on virus-host interactions.

Experimental Procedures

Sampling and growth conditions

Samples were collected from Lake Retba, Senegal in May, 2011 (14°50'14'' N, 17°14'55'' W). The Lake Retba sample 1 (LR1) was collected in the center of the lake, where salt was precipitated at the bottom. LR1 sample contained grey water with grey sediment mixed with salt. The Lake Retba sample 2 (LR2) consists of purple water with white salt sediment. The temperature, pH, and salinity of the water at the sampling site was measured at the time of the sampling.

Isolation of microorganisms and viruses was carried out during the summer and autumn 2011. For isolation, the liquid phase and the sediment (including precipitated salt) were separated by decanting. Water was transferred to clean bottles. The sediments were dissolved by adding of 6% SW buffer (see below) until salts dissolved at the room temperature (magnetic stirring). Liquid phase and the dissolved sediment were treated as one sample.

Strains and viruses were aerobically grown in modified growth medium (MGM) (Nuttall and Dyall-Smith, 1993) at 37 °C. For plaque assay, different dilutions of virus sample were mixed with host culture (300 µl) and melted top layer agar (3 ml) and plated on MGM plates. For plaque assay, the hosts were grown for 2-3 over nights to obtain stationary phase culture. For making of MGM, 30% saltwater (SW) containing 240 g NaCl, 30 g $MgCl_2 \times 6H_2O$, 35 g $MgSO_4 \times 7H_2O$, 7 g KCl, 5 ml of 1 M $CaCl_2 \times 2H_2O$, and 80 ml of 1 M Tris-HCl pH 7.2 (per liter of water) was prepared as described in the Halohandbook (Dyall-Smith, 2009). One litre of MGM medium contained 5 g of peptone (Oxoid), and 1 g of Bacto yeast extract (Becton, Dickinson and Company). Top layer, solid, and liquid medium contained 18% SW, 20% SW, and 23% SW, respectively. For the top layer and solid media, 4 g or 14 g of Bacto agar (Becton, Dickinson and Company) was added, respectively.

452

Isolation of microorganisms, 16S rRNA gene sequencing and phylogenetic tree

454   To isolate strains from the samples, aliquots of samples (100 µl) were directly plated on MGM plates and

455   grown at 37 °C in a covered box. A selection of colonies with different morphologies and colors were picked

456   and colony purified by streaking single colonies on solid media by three consecutive times. The archaeal

457   strains used in the study are listed in Table 2. The strains were identified based on their partial 16S rRNA gene

458   sequences, which were determined as described previously (Sime-Ngando *et al.*, 2011). The 16S rRNA genes

459   were amplified by PCR. The primers were either universal for both the bacteria and archaea, or specific for

460   the archaea (Eder *et al.*, 1999). The sequences of the universal prokaryotic forward primers were 5'-

461   AGAGTTTGATCCTGGCTCAG-3' (F27) and 5'-TCCGTGCCAGCAGCCGCCG -3' (F530), and those of the universal

462   prokaryotic reverse primers were 5′-ACGGHTACCTTGTTACGACTT-3' (1512uR) and 5'-

463   CGTATTACCGCGGCTGCTGG-3' (R518). The archaea-specific primers were 5'-TCYGGTTGATCCTGCC-3' (8aF)

464   and 5'-AGGAGGTGATCCAGCC-3' (AR1456). The reaction mixture (50 µl of total volume) contained 1X Taq™

465   buffer (Promega, Madison, WI, USA), 1.5 mM $MgCl_2$, dNTPs at a concentration of 0.2 mM each, 1 U of Taq

466   polymerase, each primer at a concentration of 125 pmol, and 5 ng of template DNA. The amplification was

467   ended by an extension step for 10 min at 72°C. Negative and positive controls were included. Five µl of PCR

468   products were loaded onto 0.8 % agarose gel in TAE 1X (Tris-acetic acid-EDTA buffer) and visualized under

469   UV light after ethidium bromide staining. PCR products obtained were cloned using TOPO TA cloning kit

470   (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. After blue-white selection, positive

471   clones were grown at 37°C overnight on 96-well tissue culture plates in the presence of kanamycin.

472   The clones were picked and suspended in TE followed by boiling at 96°C, and used as a template DNA for PCR

473   amplification using M13 primers targeting the cloning vector (5'-GTAAAACGACGGCCAG-3' and 5'-

474   CAGGAAACAGCTATGAC-3'). The selected clones were grown as previously to extract plasmid DNA using

475   Nucleospin Plasmid preparation Kit (Macherey-Nagel, EURL, France) and sent for Sanger sequencing using

476   M13 primers on both strands. For a first classification, we used the SILVA r128 rRNA classifier (Pruesse *et al.*,

477   2012). For the phylogenetic tree, sequences were aligned using MUSCLE (Edgar, 2004) and maximum

478   likelihood trees were constructed using the program FastTree2 (Price *et al.*, 2010). Bootstrapping (100

479   replicates) was performed using the Seqboot program in the PHYLIP package (Felsenstein, 1993).

480   The sequences are deposited in the NCBI data bank under the accession numbers MG462733-MG462751

481   (Table 2).

482

483   **Isolation of viruses**

484   Viruses were isolated either by direct plating or enrichment culture techniques. The pure cultures of LR1

485   host strains (Table 2) were used to isolate viruses from LR1 samples. Hosts of LR2 (Table 2) were used to

486   isolate viruses from LR2 samples. The LR1 and LR2 samples were first centrifuged at 13 000 rpm (Table-top

487   Eppendorf centrifuge) for 10 min at room temperature and the supernatant we used for virus isolation. To

488   remove microorganisms from the LR1 sediment sample, the sample was filtered (pore size 0.2 µm). For direct

489   plating, 100 µl of samples was mixed with dense host culture (300 µl) and melted top layer agar (3 ml) and

490   poured on a plate, which was incubated at 37 °C until a dense lawn of archaea was observed (typically two

491   to three days). For enrichments, 500 µl of samples was mixed with host culture (1 ml grown for 2-3 days) and

492   incubated in a shaker (200 rpm) 1-2 overnights. The enrichment samples (100 µl and 500 µl) were plated as

493   above. The obtained plaques were plaque purified. The plaque purification was carried out by growing the

494   viruses on their host strain to obtain separate plaques by using plaque assay. Plaques were picked by sterile

495   toothpick or Pasteur pipette, and a single plaque was resuspended in 0.5 ml of MGM liquid medium. The

496   plaque purification was repeated by using the single plaque as the starting material. The single plaque

497   purification was carried out by three consecutive times for each virus.

498   For all plaque assays and preparation of the virus stocks, appropriate virus dilution (100 µl) mixed with host

499   culture (300 µl) and melted top layer agar (3 ml) was plated. The plates were incubated for 2-3 days. The

500   virus stocks were prepared from semiconfluent plates. Top layer media from the semiconfluent plates were

501   collected by a sterile glass triangle into a sterile Erlenmeyer bottle and 2 ml of liquid medium was added per

502   each collected plate. The suspension was incubated for 1.5 hours at 37 °C. Cell debris and agar were removed

503   by centrifugation (Thermo Scientific F12 rotor, 8000 rpm, 20 min, 5 °C). The supernatant was put into a clean

504   bottle and it is referred as a virus stock. One semiconfluent plate produces approximately 3-3.5 ml of virus

505   stock. Stability of viruses (virus stocks stored at 4 °C) was monitored for four weeks by plaque assay. To test

506   the sensitivity of the viruses to organic solvents and detergents, viruses (virus stock in MGM) were incubated

507   in 20% (v/v) chloroform, 0.1% (v/v) Nonidet P-40, or 0.1% Triton X-100 for 15 min at 22°C. MGM was used as

508   a control. The infectivity of the viruses was determined by plaque assay and the experiments have been

509   repeated at least for two times.

510

511   **Virus purification and particle analysis**

512   The virus stocks (typically made of 200 plates producing 600-700 ml of virus stock or 400 plates producing

513   1200-1400 ml of virus stock) were treated with DNase I (70 µg/ml; 30 min at 37 °C; Sigma-Aldrich) prior the

514   purification. Viruses were precipitated from the virus stocks by using two-step polyethylene glycol (PEG)-

515   NaCl precipitation (Yamamoto *et al.*, 1970). First, the impurities were precipitated by using 4% (w/v) PEG

516   6000 (no NaCl added due to the high salinity of the virus stock). PEG was dissolved by magnetic stirring for

517   30 min at 4°C. After centrifugation (Thermo Scientific F12 rotor, 8000 rpm, 40 min, 5 °C), PEG was added to

518   the supernatant to obtain a final concentration of 11% (w/v). After dissolution of PEG and centrifugation (see

519   above), the obtained virus precipitate was dissolved in 18% SW buffer followed by removal of the aggregates

520   and undissolved components (Thermo Scientific F20 rotor, 7000 rpm, 10 min, 5 °C). Viruses were first purified

521   by rate zonal ultracentrifugation in sucrose by using linear 5-20% sucrose gradients (18% SW buffer; Sorvall

522   rotor AH629, 24 000 rpm, 15 °C). The running times were 2.5 h (HRPV10), 1 h 45 min (HRPV11), and 3 h

523   (HRPV12 and HFTV1). After rate zonal centrifugation, viruses were purified by equilibrium centrifugation in

524   CsCl gradients (mean ρ=1.30-1.35 g/ml in 18% SW; Sorvall rotor AH629, 20 000 rpm, 19 h, 20 °C), and

525   concentrated by differential centrifugation (Sorvall rotor T647.5, 32 000 rpm, 3-5 h, 15 °C). Virus purifications

526   were repeated at least three times for each virus. Protein concentrations were determined by Bradford assay

527   using bovine serum albumin as a standard (Bradford, 1976). The proteins were analyzed by using modified

528 tricine-sodium dodecyl sulfate polyacrylamide gel electrophoresis (14% acrylamide in the separation gel;

529 (Schägger and von Jagow, 1987)). Gels were stained with Coomassie Brilliant Blue R 250 (Serva).

530 Viral lipids were isolated by chloroform-methanol extraction from the purified HRPV10, HRPV11, and HRPV12

531 virus particles and from the early-stationary-phase cells of *Haloarcula hispanica* (Juez *et al.*, 1986) and

532 *Halorubrum* sp. LR2-17, LR2-12, LR1-23 strains as previously described (Folch *et al.*, 1957; Kates *et al.*, 1972).

533 Extracted lipids were dissolved in chloroform-methanol (9:1) and analyzed on pre-activated thin layer

534 chromatography (TLC) silica plates, which were developed with chloroform–methanol–90% acetic acid

535 (65:4:35 [vol/vol/vol]). Lipids were visualized by ammonium molybdate staining (Arnold *et al.*, 2000). The

536 plate was quickly dipped into a solution containing 10% (v/v) H_2SO_4 and 5% (w/v) ammonium molybdate,

537 after which the excess liquid was dried, and the plate was incubated at 140 °C for around 15 min.

538 For transmission electron microscopy, 5 µl samples of the purified virus particles were adsorbed on copper

539 pioloform coated grids (Electron Microscopy Unit, HiLIFE Institute of Biotechnology, University of Helsinki).

540 The particles were negatively stained either with 3% (w/v) uranylacetate (pH 4.5) or 1% (w/v)

541 phosphotungstic acid (pH 7.0) , and visualized by JEOL 1400 transmission electron microscope (Electron

542 Microscopy Unit, HiLIFE Institute of Biotechnology, University of Helsinki) operating at 80 kV acceleration

543 voltage.

544

545 **Virus genome analysis, sequencing and annotation**

546 Nucleic acid was purified from the pure virus particles. The particles in 18% SW were diluted 1:4 in 20 mM

547 Tris-HCl, pH 7.2 and treated with 1% (w/v) sodium dodecyl sulphate and 100 µg/ml proteinase K (Thermo

548 Scientific) in the presence of 1 mM ethylenediaminetetraacetic acid (EDTA) for an hour at 37 °C. Nucleic acid

549 was extracted by phenol-ether extraction and followed by precipitation with NaCl and ethanol. Purified

550 nucleic acids were treated with RQ1 DNase (Promega), Exonuclease III (Fermentas), Mung bean nuclease

551 (MBN 0.025, 0.5, or 5.0 U/µg DNA; Promega) according to manufacturers' instructions. For MBN experiments,

552   phage φX174 ssDNA genome and its dsDNA replicative form RFII (New England Biolabs) were used as

553   controls.

554   Libraries were prepared using TruSeq PCRfree library preparation. Samples were sequenced by Illumina

555   MiSeq 600 cycles (Illumina Inc., San Diego, CA) with 2x300 bp read length. The sequencing reads were

556   trimmed based on the quality scores (limit 0.05) from a base-caller algorithm available in the sequencing

557   files. The trimming was performed using the modified-Mott trimming algorithm implemented in the CLC

558   Genomics Workbench v7 (QIAGEN Bioinformatics) and the trimmed reads were subsequently assembled into

559   contigs using the same software with default parameters. Protein-coding genes were predicted using

560   Prodigal (Hyatt *et al.*, 2010), and tRNA genes using tRNAscan-SE (Lowe and Eddy, 1997). Additional

561   annotation of genes was done by comparing against the NCBI NR, COG (Tatusov *et al.*, 2003), and TIGRfam

562   (Haft *et al.*, 2001) databases, and also manually annotated using HHPRED server (Zimmermann *et al.*, 2018).

563   The sequences are deposited in the NCBI GenBank data bank under the accession numbers MG550110 -

564   MG550113.

565   All pairwise comparisons of the nucleotide sequences were conducted using the Genome-BLAST Distance

566   Phylogeny (GBDP) method (Meier-Kolthoff *et al.*, 2013) under settings recommended for prokaryotic viruses

567   (Meier-Kolthoff   and   Goker,   2017).   All   reference   genomes   were   downloaded   from

568   https://www.ncbi.nlm.nih.gov/genome/browse/

569   Genome phylogenies were constructed using VICTOR (Meier-Kolthoff and Goker, 2017), a Genome BLAST

570   Distance Phylogeny (GBDP) method which calculates intergenomic distances between pairs of viruses based

571   on pairwise comparison of nucleotide sequences. The resulting intergenomic distances (including 100

572   replicates each) were used to infer a balanced minimum evolution tree with branch support via FASTME

573   including SPR postprocessing (Lefort *et al.*, 2015) for the formula D0. The trees were rooted at the outgroup

574   and visualized with FigTree (Rambaut, 2006). For both single gene phylogenies, the sequences were aligned

575   using MUSCLE (Edgar, 2004). Maximum likelihood trees were constructed using the program FastTree2 (Price

576   *et al.*, 2010). Bootstrapping was performed using the Seqboot program in the PHYLIP package (Felsenstein,

577   1993). Comparisons among related viral genomes and reference genomes were performed using tBLASTx or

578   BLASTN (Edgar, 2010).

579

580   **Virus-host interaction studies**

581   Infectivity of HRPV10, HRPV11, HRPV12, and HFTV1 viruses was tested on 19 Lake Retba strains (Table 2) and

582   48 culture collection strains representing genus *Halorubrum* or *Haloferax* (Supplementary Table 3) by spot-

583   on-lawn assay. Undiluted and diluted (10^{-2}) virus stocks (10 µl) were applied on the top layer agar inoculated

584   with the test strain. The virus host strain and MGM medium were used as positive and negative controls. All

585   positive results (growth inhibitions) were verified by plaque assay.

586

587   **Analysis of metaviromes**

588   Viromes were downloaded from Metavir 2 (Roux *et al.*, 2014). Only sequence matches longer than 50 bp

589   withe-value less than 1e-5 and more than 95% identity were considered. The recruitment of each genome

590   from the virome was calculated by dividing the number of hits by the length of the contig (in kb) and by the

591   size of the database (in Gb). This normalized measure is abbreviated as RPKG (**R**eads recruited **P**er **K**b of

592   genome per **G**b of metagenome).

593   In order to test the performance of the currently available tools for identifying archaeal viruses in

594   metagenomic dataset, we ran the VirSorter analysis (Roux *et al.*, 2015) against the RefSeq virus database. Of

595   the four genomes analyzed, only HFTV1 was considered by VirSorter to be of viral origin under the category

596   2 ("quite sure"), with three detected "phage hallmark genes". None of the pleolipoviruses was recognized as

597   a virus, pointing to a need for improvement of the database of virus hallmark genes.

598

Acknowledgments

The authors have no conflict of interest to declare.

References

Aalto, A.P., Bitto, D., Ravantti, J.J., Bamford, D.H., Huiskonen, J.T., and Oksanen, H.M. (2012) Snapshot of virus evolution in hypersaline environments from the characterization of a membrane-containing Salisaeta icosahedral phage 1. *Proc Natl Acad Sci U S A* **109**: 7079-7084.

Abby, S.S., Melcher, M., Kerou, M., Krupovic, M., Stieglmeier, M., Rossel, C. *et al.* (2018) Candidatus Nitrosocaldus cavascurensis, an ammonia oxidizing, extremely thermophilic archaeon with a highly mobile genome. *Front Microbiol* **9**: 28.

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

Ahlgren, N.A., Fuchsman, C.A., Rocap, G., and Fuhrman, J.A. (2019) Discovery of several novel, widespread, and ecologically distinct marine Thaumarchaeota viruses that encode amoC nitrification genes. *ISME J*: [Epub ahead of print].

Arnold, H.P., Zillig, W., Ziese, U., Holz, I., Crosby, M., Utterback, T. *et al.* (2000) A novel lipothrixvirus, SIFV, of the extremely thermophilic crenarchaeon Sulfolobus. *Virology* **267**: 252-266.

Atanasova, N.S., Bamford, D.H., and Oksanen, H.M. (2015a) Haloarchaeal virus morphotypes. *Biochimie* **118**: 333-343.

Atanasova, N.S., Oksanen, H.M., and 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., and Oksanen, H.M. (2012) Global network of specific virus-host interactions in hypersaline environments. *Environ Microbiol* **14**: 426-440.

633 Atanasova, N.S., Demina, T.A., Buivydas, A., Bamford, D.H., and Oksanen, H.M. (2015c) Archaeal viruses
634 multiply: temporal screening in a solar saltern. *Viruses* **7**: 1902-1926.
635 Atanasova, N.S., Demina, T.A., Krishnam Rajan Shanthi, S.N.V., Oksanen, H.M., and Bamford, D.H. (2018a)
636 Extremely halophilic pleomorphic archaeal virus HRPV9 extends the diversity of pleolipoviruses with
637 integrases. *Res Microbiol* **169**: 500-504.
638 Atanasova, N.S., Heiniö, C.H., Demina, T.A., Bamford, D.H., and Oksanen, H.M. (2018b) The unexplored
639 diversity of pleolipoviruses: the surprising case of two viruses with identical major structural modules. *Genes*
640 *(Basel)* **9**: 131.
641 Bamford, D.H., Pietilä, M.K., Roine, E., Atanasova, N.S., Dienstbier, A., Oksanen, H.M., and ICTV Report
642 Consortium (2017) ICTV Virus Taxonomy Profile: Pleolipoviridae. *J Gen Virol* **98**: 2916-2917.
643 Bamford, D.H., Ravantti, J.J., Rönnholm, G., Laurinavicius, S., Kukkaro, P., Dyall-Smith, M. *et al.* (2005)
644 Constituents of SH1, a novel lipid-containing virus infecting the halophilic euryarchaeon Haloarcula hispanica.
645 *J Virol* **79**: 9097-9107.
646 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S. *et al.* (2012) SPAdes: a new
647 genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**: 455-477.
648 Bell, S.D., and Botchan, M.R. (2013) The minichromosome maintenance replicative helicase. *Cold Spring Harb*
649 *Perspect Biol* **5**: a012807.
650 Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein
651 utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.
652 Crits-Christoph, A., Gelsinger, D.R., Ma, B., Wierzchos, J., Ravel, J., Davila, A. *et al.* (2016) Functional
653 interactions of archaea, bacteria and viruses in a hypersaline endolithic community. *Environ Microbiol* **18**:
654 2064-2077.
655 Cuadros-Orellana, S., Martin-Cuadrado, A.B., Legault, B., D'Auria, G., Zhaxybayeva, O., Papke, R.T., and
656 Rodriguez-Valera, F. (2007) Genomic plasticity in prokaryotes: the case of the square haloarchaeon. *ISME J*
657 **1**: 235-245.
658 Danovaro, R., Dell'Anno, A., Corinaldesi, C., Rastelli, E., Cavicchioli, R., Krupovic, M. *et al.* (2016) Virus-
659 mediated archaeal hecatomb in the deep seafloor. *Sci Adv* **2**: e1600492.
660 De Paepe, M., Hutinet, G., Son, O., Amarir-Bouhram, J., Schbath, S., and Petit, M.A. (2014) Temperate phages
661 acquire DNA from defective prophages by relaxed homologous recombination: the role of Rad52-like
662 recombinases. *PLoS Genet* **10**: e1004181.
663 DeMaere, M.Z., Williams, T.J., Allen, M.A., Brown, M.V., Gibson, J.A., Rich, J. *et al.* (2013) High level of
664 intergenera gene exchange shapes the evolution of haloarchaea in an isolated Antarctic lake. *Proc Natl Acad*
665 *Sci U S A* **110**: 16939-16944.
666 Demina, T.A., Atanasova, N.S., Pietilä, M.K., Oksanen, H.M., and Bamford, D.H. (2016a) Vesicle-like virion of
667 Haloarcula hispanica pleomorphic virus 3 preserves high infectivity in saturated salt. *Virology* **499**: 40-51.
668 Demina, T.A., Pietilä, M.K., Svirskaite, J., Ravantti, J.J., Atanasova, N.S., Bamford, D.H., and Oksanen, H.M.
669 (2016b) Archaeal Haloarcula californiae icosahedral virus 1 highlights conserved elements in icosahedral
670 membrane-containing DNA viruses from extreme environments. *mBio* **7**.
671 Demina, T.A., Pietilä, M.K., Svirskaite, J., Ravantti, J.J., Atanasova, N.S., Bamford, D.H., and Oksanen, H.M.
672 (2017) HCIV-1 and other tailless icosahedral internal membrane-containing viruses of the family
673 Sphaerolipoviridae. *Viruses* **9**.
674 Di Meglio, L., Santos, F., Gomariz, M., Almansa, C., Lopez, C., Anton, J., and Nercessian, D. (2016) Seasonal
675 dynamics of extremely halophilic microbial communities in three Argentinian salterns. *FEMS Microbiol Ecol*
676 **92**.
677 Dyall-Smith, M. (2009). Halohandbook. URL http://www.haloarchaea.com/resources/halohandbook/
678 Dyall-Smith, M., and Pfeiffer, F. (2018) The PL6-family plasmids of Haloquadratum are virus-related. *Front*
679 *Microbiol* **9**: 1070.
680 Dyall-Smith, M., Pfeifer, F., Witte, A., Oesterhelt, D., and Pfeiffer, F. (2018) Complete genome sequence of
681 the model halovirus phiH1 (phiH1). *Genes (Basel)* **9**.
682 Dyall-Smith, M., Palm, P., Wanner, G., Witte, A., Oesterhelt, D., and Pfeiffer, F. (2019) Halobacterium
683 salinarum virus ChaoS9, a Novel Halovirus Related to PhiH1 and PhiCh1. *Genes (Basel)* **10**.

684 Dyall-Smith, M.L., Pfeiffer, F., Klee, K., Palm, P., Gross, K., Schuster, S.C. *et al.* (2011) *Haloquadratum walsbyi*:
685 limited diversity in a global pond. *PLoS One* **6**: e20968.
686 Eder, W., Ludwig, W., and Huber, R. (1999) Novel 16S rRNA gene sequences retrieved from highly saline brine
687 sediments of kebrit deep, red Sea. *Arch Microbiol* **172**: 213-218.
688 Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic*
689 *Acids Res* **32**: 1792-1797.
690 Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460-
691 2461.
692 El Omari, K., Li, S., Kotecha, A., Walter, T.S., Bignon, E.A., Harlos, K. *et al.* (2019) The structure of a prokaryotic
693 viral envelope protein expands the landscape of membrane fusion proteins. *Nat Commun* **10**: 846.
694 Erdmann, S., Tschitschko, B., Zhong, L., Raftery, M.J., and Cavicchioli, R. (2017) A plasmid from an Antarctic
695 haloarchaeon uses specialized membrane vesicles to disseminate and infect plasmid-free cells. *Nat Microbiol*
696 **2**: 1446-1455.
697 Felsenstein, J. (1993) *PHYLIP (phylogeny inference package), version 3.5 c*: Joseph Felsenstein.
698 Folch, J., Lees, M., and Sloane Stanley, G.H. (1957) A simple method for the isolation and purification of total
699 lipides from animal tissues. *J Biol Chem* **226**: 497-509.
700 Forterre, P., Krupovic, M., Raymann, K., and Soler, N. (2014) Plasmids from Euryarchaeota. *Microbiol Spectr*
701 **2**: PLAS-0027-2014.
702 Garcia-Heredia, I., Martin-Cuadrado, A.B., Mojica, F.J., Santos, F., Mira, A., Anton, J., and Rodriguez-Valera,
703 F. (2012) Reconstructing viral genomes from the environment using fosmid clones: the case of haloviruses.
704 *PLoS One* **7**: e33802.
705 Gardner, A.F., Bell, S.D., White, M.F., Prangishvili, D., and Krupovic, M. (2014) Protein-protein interactions
706 leading to recruitment of the host DNA sliding clamp by the hyperthermophilic Sulfolobus islandicus rod-
707 shaped virus 2. *J Virol* **88**: 7105-7108.
708 Garneau, J.R., Depardieu, F., Fortier, L.C., Bikard, D., and Monot, M. (2017) PhageTerm: a tool for fast and
709 accurate determination of phage termini and packaging mechanism using next-generation sequencing data.
710 *Sci Rep* **7**: 8292.
711 Gunde-Cimerman, N., Plemenitas, A., and Oren, A. (2018) Strategies of adaptation of microorganisms of the
712 three domains of life to high salt concentrations. *FEMS Microbiol Rev* **42**: 353-375.
713 Haft, D.H., Loftus, B.J., Richardson, D.L., Yang, F., Eisen, J.A., Paulsen, I.T., and White, O. (2001) TIGRFAMs: a
714 protein family resource for the functional identification of proteins. *Nucleic Acids Res* **29**: 41-43.
715 Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: prokaryotic
716 gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119.
717 Iranzo, J., Krupovic, M., and Koonin, E.V. (2016a) The double-stranded DNA virosphere as a modular
718 hierarchical network of gene sharing. *mBio* **7**: e00978-00916.
719 Iranzo, J., Koonin, E.V., Prangishvili, D., and Krupovic, M. (2016b) Bipartite network analysis of the archaeal
720 virosphere: evolutionary connections between viruses and capsidless mobile elements. *J Virol* **90**: 11043-
721 11055.
722 Juez, G., Rodriguez-Valera, F., Ventosa, A., and Kushner, D.J. (1986) Haloarcula hispanica spec. nov. and
723 Haloferax gibbonsii spec, nov., two new species of extremely halophilic archaebacteria. *Syst Appl Microbiol*
724 **8**: 75-79.
725 Juhala, R.J., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F., and Hendrix, R.W. (2000) Genomic sequences of
726 bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. *J Mol Biol*
727 **299**: 27-51.
728 Kates, M., Work, T.S., and Work, E. (1972) *Techniques of Lipidology: Isolation, Analysis and Identification of*
729 *Lipids*. North-Holland, Amsterdam.
730 Kazlauskas, D., Krupovic, M., and Venclovas, C. (2016) The logic of DNA replication in double-stranded DNA
731 viruses: insights from global analysis of viral genomes. *Nucleic Acids Res* **44**: 4551-4564.
732 Koskella, B., Thompson, J.N., Preston, G.M., and Buckling, A. (2011) Local biotic environment shapes the
733 spatial scale of bacteriophage adaptation to bacteria. *Am Nat* **177**: 440-451.
734 Krupovic, M., and Bamford, D.H. (2010) Order to the viral universe. *J Virol* **84**: 12476-12479.

735 Krupovic, M., Forterre, P., and Bamford, D.H. (2010a) Comparative analysis of the mosaic genomes of tailed
736 archaeal viruses and proviruses suggests common themes for virion architecture and assembly with tailed
737 viruses of bacteria. *J Mol Biol* **397**: 144-160.
738 Krupovic, M., Gribaldo, S., Bamford, D.H., and Forterre, P. (2010b) The evolutionary history of archaeal MCM
739 helicases: a case study of vertical evolution combined with hitchhiking of mobile genetic elements. *Mol Biol*
740 *Evol* **27**: 2716-2732.
741 Krupovic, M., Prangishvili, D., Hendrix, R.W., and Bamford, D.H. (2011a) Genomics of bacterial and archaeal
742 viruses: dynamics within the prokaryotic virosphere. *Microbiol Mol Biol Rev* **75**: 610-635.
743 Krupovic, M., Spang, A., Gribaldo, S., Forterre, P., and Schleper, C. (2011b) A thaumarchaeal provirus testifies
744 for an ancient association of tailed viruses with archaea. *Biochem Soc Trans* **39**: 82-88.
745 Krupovic, M., Cvirkaite-Krupovic, V., Iranzo, J., Prangishvili, D., and Koonin, E.V. (2018) Viruses of archaea:
746 structural, functional, environmental and evolutionary genomics. *Virus Res* **244**: 181-193.
747 Lefort, V., Desper, R., and Gascuel, O. (2015) FastME 2.0: A Comprehensive, Accurate, and Fast Distance-
748 Based Phylogeny Inference Program. *Mol Biol Evol* **32**: 2798-2800.
749 Li, M., Liu, H., Han, J., Liu, J., Wang, R., Zhao, D. *et al.* (2013) Characterization of CRISPR RNA biogenesis and
750 Cas6 cleavage-mediated inhibition of a provirus in the haloarchaeon Haloferax mediterranei. *J Bacteriol* **195**:
751 867-875.
752 Liu, Y., Wang, J., Liu, Y., Wang, Y., Zhang, Z., Oksanen, H.M. *et al.* (2015) Identification and characterization
753 of SNJ2, the first temperate pleolipovirus integrating into the genome of the SNJ1-lysogenic archaeal strain.
754 *Mol Microbiol* **98**: 1002-1020.
755 Lopes, A., Amarir-Bouhram, J., Faure, G., Petit, M.A., and Guerois, R. (2010) Detection of novel recombinases
756 in bacteriophage genomes unveils Rad52, Rad51 and Gp2.5 remote homologs. *Nucleic Acids Res* **38**: 3952-
757 3962.
758 Lopez-Perez, M., Haro-Moreno, J.M., de la Torre, J.R., and Rodriguez-Valera, F. (2019) Novel Caudovirales
759 associated with Marine Group I Thaumarchaeota assembled from metagenomes. *Environ Microbiol*: [Epub
760 ahead of print].
761 Lowe, T.M., and Eddy, S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in
762 genomic sequence. *Nucleic Acids Res* **25**: 955-964.
763 Maier, L.K., Dyall-Smith, M., and Marchfelder, A. (2015) The adaptive immune system of *Haloferax volcanii*.
764 *Life (Basel)* **5**: 521-537.
765 Makarova, K.S., Wolf, Y.I., Forterre, P., Prangishvili, D., Krupovic, M., and Koonin, E.V. (2014) Dark matter in
766 archaeal genomes: a rich source of novel mobile elements, defense systems and secretory complexes.
767 *Extremophiles* **18**: 877-893.
768 Meier-Kolthoff, J.P., and Goker, M. (2017) VICTOR: genome-based phylogeny and classification of prokaryotic
769 viruses. *Bioinformatics* **33**: 3396-3404.
770 Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.P., and Goker, M. (2013) Genome sequence-based species
771 delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**: 60.
772 Mizuno, C.M., Rodriguez-Valera, F., Kimes, N.E., and Ghai, R. (2013) Expanding the marine virosphere using
773 metagenomics. *PLoS Genet* **9**: e1003987.
774 Nuttall, S.D., and Dyall-Smith, M.L. (1993) HF1 and HF2: novel bacteriophages of halophilic archaea. *Virology*
775 **197**: 678-684.
776 Oh, D., Porter, K., Russ, B., Burns, D., and Dyall-Smith, M. (2010) Diversity of Haloquadratum and other
777 haloarchaea in three, geographically distant, Australian saltern crystallizer ponds. *Extremophiles* **14**: 161-169.
778 Oren, A. (2002) Molecular ecology of extremely halophilic Archaea and Bacteria. *FEMS Microbiol Ecol* **39**: 1-
779 7.
780 Oren, A., Bratbak, G., and Heldal, M. (1997) Occurrence of virus-like particles in the Dead Sea. *Extremophiles*
781 **1**: 143-149.
782 Pagaling, E., Haigh, R.D., Grant, W.D., Cowan, D.A., Jones, B.E., Ma, Y. *et al.* (2007) Sequence analysis of an
783 Archaeal virus isolated from a hypersaline lake in Inner Mongolia, China. *BMC Genomics* **8**: 410.
784 Pan, M., Kelman, L.M., and Kelman, Z. (2011) The archaeal PCNA proteins. *Biochem Soc Trans* **39**: 20-24.
785 Pedulla, M.L., Ford, M.E., Houtz, J.M., Karthikeyan, T., Wadsworth, C., Lewis, J.A. *et al.* (2003) Origins of highly
786 mosaic mycobacteriophage genomes. *Cell* **113**: 171-182.

787 Philosof, A., Yutin, N., Flores-Uribe, J., Sharon, I., Koonin, E.V., and Beja, O. (2017) Novel abundant oceanic
788 viruses of uncultured Marine Group II Euryarchaeota. *Curr Biol* **27**: 1362-1368.
789 Pietilä, M.K., Atanasova, N.S., Oksanen, H.M., and Bamford, D.H. (2013a) Modified coat protein forms the
790 flexible spindle-shaped virion of haloarchaeal virus His1. *Environ Microbiol* **15**: 1674-1686.
791 Pietilä, M.K., Roine, E., Paulin, L., Kalkkinen, N., and Bamford, D.H. (2009) An ssDNA virus infecting archaea:
792 a new lineage of viruses with a membrane envelope. *Mol Microbiol* **72**: 307-319.
793 Pietilä, M.K., Laurinavicius, S., Sund, J., Roine, E., and Bamford, D.H. (2010) The single-stranded DNA genome
794 of novel archaeal virus halorubrum pleomorphic virus 1 is enclosed in the envelope decorated with
795 glycoprotein spikes. *J Virol* **84**: 788-798.
796 Pietilä, M.K., Demina, T.A., Atanasova, N.S., Oksanen, H.M., and Bamford, D.H. (2014) Archaeal viruses and
797 bacteriophages: comparisons and contrasts. *Trends Microbiol* **22**: 334-344.
798 Pietilä, M.K., Roine, E., Sencilo, A., Bamford, D.H., and Oksanen, H.M. (2016) Pleolipoviridae, a newly
799 proposed family comprising archaeal pleomorphic viruses with single-stranded or double-stranded DNA
800 genomes. *Arch Virol* **161**: 249-256.
801 Pietilä, M.K., Atanasova, N.S., Manole, V., Liljeroos, L., Butcher, S.J., Oksanen, H.M., and Bamford, D.H. (2012)
802 Virion architecture unifies globally distributed pleolipoviruses infecting halophilic archaea. *J Virol* **86**: 5067-
803 5079.
804 Pietilä, M.K., Laurinmäki, P., Russell, D.A., Ko, C.C., Jacobs-Sera, D., Hendrix, R.W. *et al.* (2013b) Structure of
805 the archaeal head-tailed virus HSTV-1 completes the HK97 fold story. *Proc Natl Acad Sci U S A* **110**: 10604-
806 10609.
807 Pietilä, M.K., Laurinmäki, P., Russell, D.A., Ko, C.C., Jacobs-Sera, D., Butcher, S.J. *et al.* (2013c) Insights into
808 head-tailed viruses infecting extremely halophilic archaea. *J Virol* **87**: 3248-3260.
809 Pope, W.H., Bowman, C.A., Russell, D.A., Jacobs-Sera, D., Asai, D.J., Cresawn, S.G. *et al.* (2015) Whole genome
810 comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. *Elife*
811 **4**: e06416.
812 Prangishvili, D., Bamford, D.H., Forterre, P., Iranzo, J., Koonin, E.V., and Krupovic, M. (2017) The enigmatic
813 archaeal virosphere. *Nat Rev Microbiol* **15**: 724-739.
814 Price, M.N., Dehal, P.S., and Arkin, A.P. (2010) FastTree 2--approximately maximum-likelihood trees for large
815 alignments. *PLoS One* **5**: e9490.
816 Pruesse, E., Peplies, J., and Glockner, F.O. (2012) SINA: accurate high-throughput multiple sequence
817 alignment of ribosomal RNA genes. *Bioinformatics* **28**: 1823-1829.
818 Rambaut, A. (2006) FigTree 1.4.3 - a graphical viewer of phylogenetic trees and a program for producing
819 publication-ready figures. In *http://treebioedacuk/software/figtree/*.
820 Raymann, K., Forterre, P., Brochier-Armanet, C., and Gribaldo, S. (2014) Global phylogenomic analysis
821 disentangles the complex evolutionary history of DNA replication in archaea. *Genome Biol Evol* **6**: 192-212.
822 Rodriguez-Brito, B., Li, L., Wegley, L., Furlan, M., Angly, F., Breitbart, M. *et al.* (2010) Viral and microbial
823 community dynamics in four aquatic environments. *ISME J* **4**: 739-751.
824 Rossler, N., Klein, R., Scholz, H., and Witte, A. (2004) Inversion within the haloalkaliphilic virus phi Ch1 DNA
825 results in differential expression of structural proteins. *Mol Microbiol* **52**: 413-426.
826 Roux, S., Enault, F., Hurwitz, B.L., and Sullivan, M.B. (2015) VirSorter: mining viral signal from microbial
827 genomic data. *PeerJ* **3**: e985.
828 Roux, S., Tournayre, J., Mahul, A., Debroas, D., and Enault, F. (2014) Metavir 2: new tools for viral
829 metagenome comparison and assembled virome analysis. *BMC Bioinformatics* **15**: 76.
830 Roux, S., Enault, F., Ravet, V., Colombet, J., Bettarel, Y., Auguet, J.C. *et al.* (2016) Analysis of metagenomic
831 data reveals common features of halophilic viral communities across continents. *Environ Microbiol* **18**: 889-
832 903.
833 Santos-Perez, I., Charro, D., Gil-Carton, D., Azkargorta, M., Elortza, F., Bamford, D.H. *et al.* (2019) Structural
834 basis for assembly of vertical single beta-barrel viruses. *Nat Commun* **10**: 1184.
835 Schägger, H., and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for
836 the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**: 368-379.
837 Sencilo, A., Paulin, L., Kellner, S., Helm, M., and Roine, E. (2012) Related haloarchaeal pleomorphic viruses
838 contain different genome types. *Nucleic Acids Res* **40**: 5523-5534.

839 Sencilo, A., Jacobs-Sera, D., Russell, D.A., Ko, C.C., Bowman, C.A., Atanasova, N.S. *et al.* (2013) Snapshot of
840 haloarchaeal tailed virus genomes. *RNA Biol* **10**: 803-816.
841 Sime-Ngando, T., Lucas, S., Robin, A., Tucker, K.P., Colombet, J., Bettarel, Y. *et al.* (2011) Diversity of virus-
842 host systems in hypersaline Lake Retba, Senegal. *Environ Microbiol* **13**: 1956-1972.
843 Svirskaite, J., Oksanen, H.M., Daugelavicius, R., and Bamford, D.H. (2016) Monitoring physiological changes
844 in haloarchaeal cell during virus release. *Viruses* **8**: 59.
845 Söding, J., Biegert, A., and Lupas, A.N. (2005) The HHpred interactive server for protein homology detection
846 and structure prediction. *Nucleic Acids Res* **33**: W244-248.
847 Tang, S.L., Nuttall, S., and Dyall-Smith, M. (2004) Haloviruses HF1 and HF2: evidence for a recent and large
848 recombination event. *J Bacteriol* **186**: 2810-2817.
849 Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V. *et al.* (2003) The COG
850 database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**: 41.
851 Triantaphyllidis, G.V., Abatzopoulos, T.J., and Sorgeloos, P. (1998) Review of the biogeography of the genus
852 Artemia (Crustacea, Anostraca). *Journal of Biogeography* **25**: 213-226.
853 Tschitschko, B., Erdmann, S., DeMaere, M.Z., Roux, S., Panwar, P., Allen, M.A. *et al.* (2018) Genomic variation
854 and biogeography of Antarctic haloarchaea. *Microbiome* **6**: 113.
855 Wang, J., Liu, Y., Liu, Y., Du, K., Xu, S., Wang, Y. *et al.* (2018a) A novel family of tyrosine integrases encoded
856 by the temperate pleolipovirus SNJ2. *Nucleic Acids Res* **46**: 2521-2536.
857 Wang, Y., Chen, B., Cao, M., Sima, L., Prangishvili, D., Chen, X., and Krupovic, M. (2018b) Rolling-circle
858 replication initiation protein of haloarchaeal sphaerolipovirus SNJ1 is homologous to bacterial transposases
859 of the IS91 family insertion sequences. *J Gen Virol* **99**: 416-421.
860 Weitz, J.S., Poisot, T., Meyer, J.R., Flores, C.O., Valverde, S., Sullivan, M.B., and Hochberg, M.E. (2013) Phage-
861 bacteria infection networks. *Trends Microbiol* **21**: 82-91.
862 Ventosa, A., de la Haba, R.R., Sanchez-Porro, C., and Papke, R.T. (2015) Microbial diversity of hypersaline
863 environments: a metagenomic approach. *Curr Opin Microbiol* **25**: 80-87.
864 Wichels, A., Biel, S.S., Gelderblom, H.R., Brinkhoff, T., Muyzer, G., and Schutt, C. (1998) Bacteriophage
865 diversity in the North Sea. *Appl Environ Microbiol* **64**: 4128-4133.
866 Vik, D.R., Roux, S., Brum, J.R., Bolduc, B., Emerson, J.B., Padilla, C.C. *et al.* (2017) Putative archaeal viruses
867 from the mesopelagic ocean. *PeerJ* **5**: e3428.
868 Villamor, J., Ramos-Barbero, M.D., Gonzalez-Torres, P., Gabaldon, T., Rossello-Mora, R., Meseguer, I. *et al.*
869 (2018) Characterization of ecologically diverse viruses infecting co-occurring strains of cosmopolitan
870 hyperhalophilic Bacteroidetes. *ISME J* **12**: 424-437.
871 Vos, M., Birkett, P.J., Birch, E., Griffiths, R.I., and Buckling, A. (2009) Local adaptation of bacteriophages to
872 their bacterial hosts in soil. *Science* **325**: 833.
873 Yamamoto, K.R., Alberts, B.M., Benzinger, R., Lawhorne, L., and Treiber, G. (1970) Rapid bacteriophage
874 sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification.
875 *Virology* **40**: 734-744.
876 Yutin, N., Backstrom, D., Ettema, T.J.G., Krupovic, M., and Koonin, E.V. (2018) Vast diversity of prokaryotic
877 virus genomes encoding double jelly-roll major capsid proteins uncovered by genomic and metagenomic
878 sequence analysis. *Virol J* **15**: 67.
879 Zhang, Z., Liu, Y., Wang, S., Yang, D., Cheng, Y., Hu, J. *et al.* (2012) Temperate membrane-containing halophilic
880 archaeal virus SNJ1 has a circular dsDNA genome identical to that of plasmid pHH205. *Virology* **434**: 233-241.
881 Zimmermann, L., Stephens, A., Nam, S.Z., Rau, D., Kubler, J., Lozajic, M. *et al.* (2018) A completely
882 reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J Mol Biol* **430**: 2237-2243.

883

884

885 **Table and Figure legends**

886

887 **Table 1.** Viruses from Lake Retba

888

889 **Table 2.** Strains isolated from the Lake Retba samples

890

891 **Table 3.** Virus purification by PEG-NaCl precipitation, rate zonal, equilibrium and differential

892 ultracentrifugation

893 **Figure 1.** Maximum likelihood phylogenetic tree of the isolated Lake Retba strains based on the partial 16S

894 rRNA gene sequences. Isolates from Lake Retba (LR) are highlighted in red. Sequences were aligned using

895 MUSCLE (Edgar, 2004) and the maximum likelihood tree was constructed using the FastTree2 program (Price

896 *et al.*, 2010). The numbers above the branches represent bootstrap support values from 100 replicates. The

897 scale bar represents the number of substitutions per site.

898

899 **Figure 2.** Transmission electron microscopy of purified viruses (A) HRPV10, (B) HRPV11, (C) HRPV12 (D)

900 HFTV1. (A-C) staining with uranyl acetate; (D) staining with phosphotungstic acid. HFTV1 particles devoid of

901 DNA are indicated by arrows. Bars, 100 nm.

902

903 **Figure 3.** Lipid and protein analysis of virions. (A-C) A thin-layer chromatogram of lipids extracted from virus

904 particles purified by PEG-NaCl precipitation, rate zonal (in sucrose) and equilibrium (in CsCl) centrifugation

905 and concentrated by differential centrifugation (A) HRPV10, (B) HRPV11, and (C) HRPV12 and their

906 corresponding host strains. The corresponding band of each lipid species are marked by 1-4. The major lipid

907 species of *Haloarcula hispanica* (Hh) are indicated on the right and their positions marked by the Roman

908 numerals as follows: PG, phosphatidylglycerol (I); PGP-Me, phosphatidylglycerophosphate methyl ester (II);

909 PGS, phosphatidylglycerosulfate (III); TGD, triglycosyl glycerodiether (IV). (D-G) SDS-PAGE analysis of the

910 purified viruses (D) HRPV10, (E) HRPV11, (F) HRPV12, (G) HFTV1. Molecular mass marker is shown (M, kDa).

911

912 **Figure 4.** Mung bean nuclease (MBN) analyses of (A) φX174 single-stranded (ss) and double-stranded (ds)

913 genomic DNA (B) HRPV10, (C) HRPV11, and (D) HRPV12. MBN amounts used are indicated as units (U) per

914 1 µg of DNA. All reactions contained 300 ng of DNA. Molecular mass marker (M) is indicated in kb.

915

916 **Figure 5.** HRPV10, HRPV11, HRPV12 and the members of the family *Pleolipoviridae*. (A) Phylogenomic tree

917 was constructed using the Genome BLAST Distance Phylogeny (GBDP) strategy implemented in VICTOR

918 (Meier-Kolthoff and Goker, 2017). The numbers above branches are GBDP pseudo-bootstrap support values

919 from 100 replications. Clades corresponding to the genera *Alphapleolipovirus*, *Betapleolipovirus* and

920 *Gammapleolipovirus* are colored light blue, beige, and grey, respectively. (B) Genomic comparison of

921 pleolipoviruses depicted in panel A. Homologous genes are indicated with the same colors.

922

923 **Figure 6.** Genomic comparison of HFTV1, HRTV-4, eHP-4, eHP-15 and eHP-1. Open reading frames (ORFs)

924 are depicted as arrows indicating the directionality of transcription. When possible, the predicted functions

925 are indicated above the corresponding ORFs. Shading connecting the ORFs indicates the amino acid

926 sequence identity between the corresponding protein products; the color key is provided at the bottom of

927 the figure. Abbreviations: TerS and TerL, small and large subunits of the terminase, respectively; CBD,

928 carbohydrate-binding domain; PAPS, phosphoadenosine phosphosulfate; MCM, minichromosome

929 maintenance helicase; PCNA, proliferating cell nuclear antigen; MTase, methyltransferase.

930

1

2

Table 1. Viruses from Lake Retba

Virus	Original host strain	Origin of the virus	Plaque morphology and diameter	Virus stock titer (pfu/ml)	Chloroform sensitivity[a]	Nonidet P-40 sensitivity[b]	Triton X-100 sensitivity[c]	Virion morphology
Halorubrum pleomorphic virus 10 (HRPV10)	*Halorubrum* sp. LR2-17	Sample LR2	Hazy, 3-5 mm	~1×10^{11}	Resistant	Sensitive, titer drops 11 logs	Sensitive, titer drops 11 logs	Pleomorphic
Halorubrum pleomorphic virus 11 (HRPV11)	*Halorubrum* sp. LR2-12	Sample LR2	Very hazy, 5-10 mm	~5×10^{11}	Sensitive, titer drops 2-3 logs	Sensitive, titer drops 8 logs	Sensitive, titer drops 11 logs	Pleomorphic
Halorubrum pleomorphic virus 12 (HRPV12)	*Halorubrum* sp. LR1-23	Sample LR1	Hazy, 5-8 mm	~1×10^{11}	Sensitive, titer drops ~1 log	Sensitive, titer drops 7 logs	Sensitive, titer drops 10 logs	Pleomorphic
Haloferax tailed virus 1 (HFTV1)	*Haloferax* sp. LR2-5	Sample LR2	Clear, 2-4 mm	~1×10^{12}	Resistant	Resistant	Resistant	Icosahedral, long non-contractile tail

4   [a] assayed by plaque assay in the presence of 20% (v/v) chloroform

5   [b] assayed by plaque assay in the presence of 0.1 % (v/v) Nonidet P-40

6   [c] assayed by plaque assay in the presence of 0.1 % (v/v) Triton-X-100

7

1

2   **Table 2. Strains isolated from the Lake Retba samples**

Strains	Origin[a]	16S rRNA sequence Acc. No; length (bp)	Efficiency of plating[b]			
			HRPV10	HRPV11	HRPV12	HFTV1
Haloferax sp. LR1-5	LR1	MG462733; 1443				
Haloferax sp. LR1-14	LR1	MG462735; 1529				
Haloferax sp. LR1-18	LR1	MG462745; 1443				
Haloferax sp. LR1-19	LR1	MG462736; 1443				
Haloferax sp. LR1-24	LR1	MG462739; 1443				
Haloferax sp. LR2-5	LR2	MG462740; 1443				1 (H)
Haloferax sp. LR2-16	LR2	MG462742; 1443				
Halomicroarcula sp. LR2-15	LR2	MG462749; 1442				
Halorubrum sp. LR1-6	LR1	MG462734; 1440				
Halorubrum sp. LR1-15	LR1	MG462744; 1440		$\sim7\times10^{-4}$		
Halorubrum sp. LR1-21	LR1	MG462746; 1440		$\sim2\times10^{-1}$		
Halorubrum sp. LR1-22	LR1	MG462737; 1440				
Halorubrum sp. LR1-23	LR1	MG462738; 1440			1 (H)	$\sim5\times10^{-8}$
Halorubrum sp. LR2-4	LR2	MG462747; 1436				
Halorubrum sp. LR2-12	LR2	MG462741; 1441	$\sim1\times10^{-3}$	1 (H)	$\sim2\times10^{-4}$	
Halorubrum sp. LR2-13	LR2	MG462748; 1440		$\sim9\times10^{-3}$		
Halorubrum sp. LR2-17	LR2	MG462750; 1440	1 (H)			
Halorubrum sp. LR2-19	LR2	MG462751; 1440				
Halorubrum sp. LR2-20	LR2	MG462743; 1441				

3   a. LR1, Lake Retba sample 1; LR2, Lake Retba sample 2

4   b. The sensitivities of the archaeal strains to isolated viruses (Table 1) are shown as efficiency of plating (EOP)
5   measured as plaque forming units. For the original host (marked by H), the EOP was set to a value of 1. EOPs
6   on others strains are relative to the EOP of the original host.

7

1

Table 3. Virus purification by PEG-NaCl precipitation, rate zonal (in sucrose), equilibrium (in CsCl) and differential ultracentrifugation

Virus	Number of the infections purified viruses (total pfus)[a]	Recovery of the infectious purified viruses[b] (%)	Yield of the purified viruses in protein (total mg of protein)[c]	Specific infectivity of the purified viruses (pfu / mg of protein)
HRPV10	~2×10^{13}	~11	~0.8	~3×10^{13}
HRPV11	~7×10^{13}	~15	~1.5	~5×10^{13}
HRPV12	~8×10^{12}	~8	~0.4	~2×10^{13}
HFTV1	~3×10^{9}	~0.0005	~1.9	~2×10^{9}

[a] Total pfus (purified viruses) obtained from a liter of virus stock

[b] Calculated based on the total pfus in the starting material (virus stocks; see the virus stock titers in Table 1) and the final sample (purified viruses)

[c] Total mg of protein (purified viruses) obtained from a liter of virus stock

Figure 1. Maximum likelihood phylogenetic tree of the isolated Lake Retba strains based on the partial 16S rRNA gene sequences. Isolates from Lake Retba (LR) are highlighted in red. Sequences were aligned using MUSCLE (Edgar, 2004) and the maximum likelihood tree was constructed using the FastTree2 program (Price et al., 2010). The numbers above the branches represent bootstrap support values from 100 replicates. The scale bar represents the number of substitutions per site.
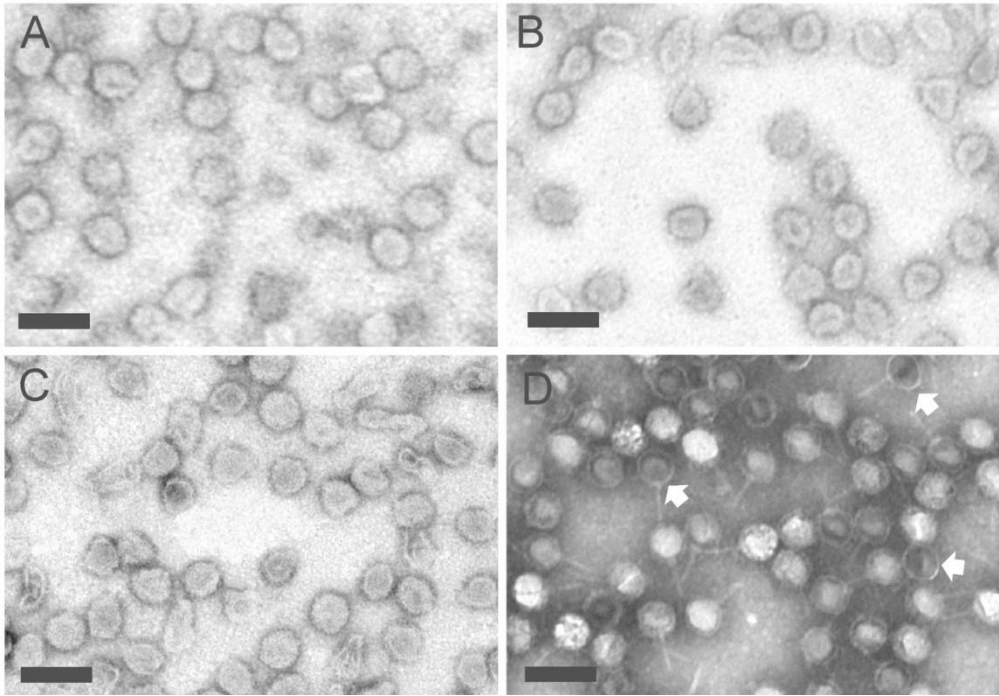
146x226mm (300 x 300 DPI)

Figure 2. Transmission electron microscopy of purified viruses (A) HRPV10, (B) HRPV11, (C) HRPV12 (D) HFTV1. (A-C) staining with uranyl acetate; (D) staining with phosphotungstic acid. HFTV1 particles devoid of DNA are indicated by arrows. Bars, 100 nm.
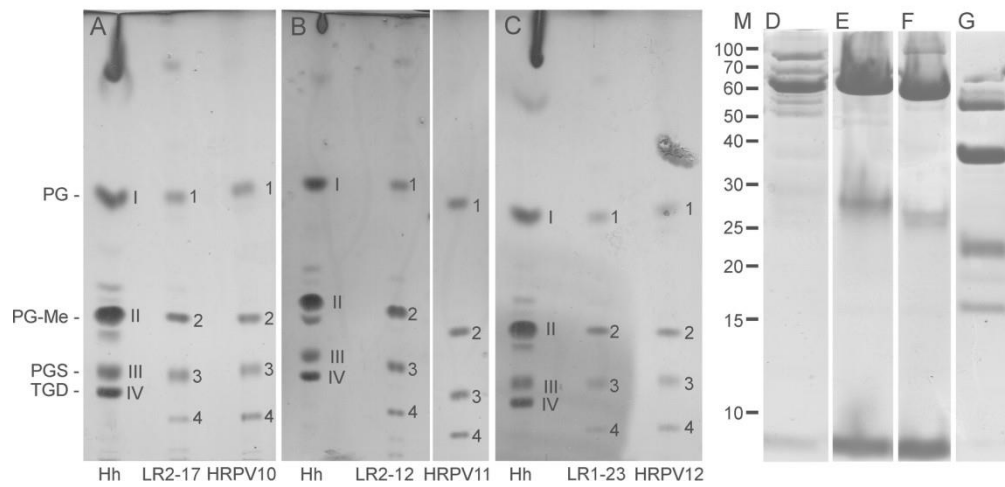
110x76mm (300 x 300 DPI)

Figure 3. Lipid and protein analysis of virions. (A-C) A thin-layer chromatogram of lipids extracted from virus particles purified by PEG-NaCl precipitation, rate zonal (in sucrose) and equilibrium (in CsCl) centrifugation and concentrated by differential centrifugation (A) HRPV10, (B) HRPV11, and (C) HRPV12 and their corresponding host strains. The corresponding band of each lipid species are marked by 1-4. The major lipid species of *Haloarcula hispanica* (Hh) are indicated on the right and their positions marked by the Roman numerals as follows: PG, phosphatidylglycerol (I); PGP-Me, phosphatidylglycerophosphate methyl ester (II); PGS, phosphatidylglycerosulfate (III); TGD, triglycosyl glycerodiether (IV). (D-G) SDS-PAGE analysis of the purified viruses (D) HRPV10, (E) HRPV11, (F) HRPV12, (G) HFTV1. Molecular mass marker is shown (M, kDa).
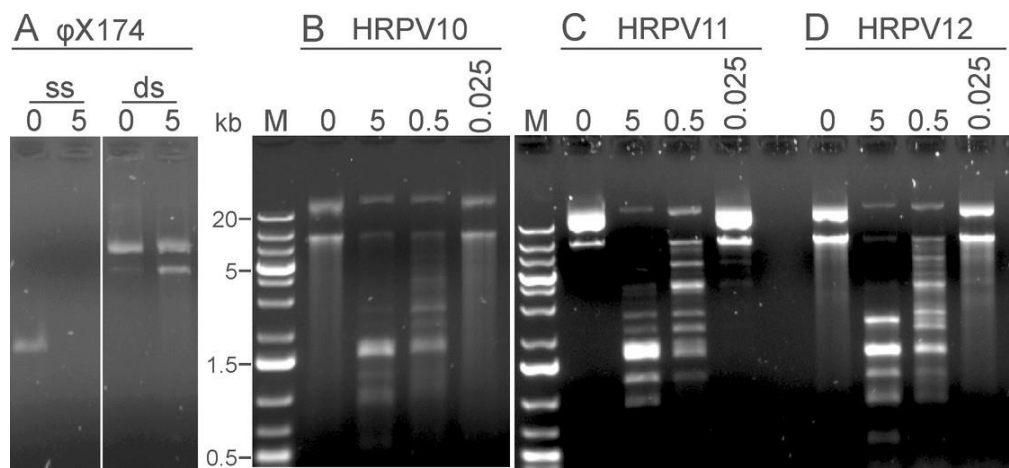
156x76mm (300 x 300 DPI)

Figure 4. Mung bean nuclease (MBN) analyses of (A) φX174 single-stranded (ss) and double-stranded (ds) genomic DNA (B) HRPV10, (C) HRPV11, and (D) HRPV12. MBN amounts used are indicated as units (U) per 1 μg of DNA. All reactions contained 300 ng of DNA. Molecular mass marker (M) is indicated in kb.
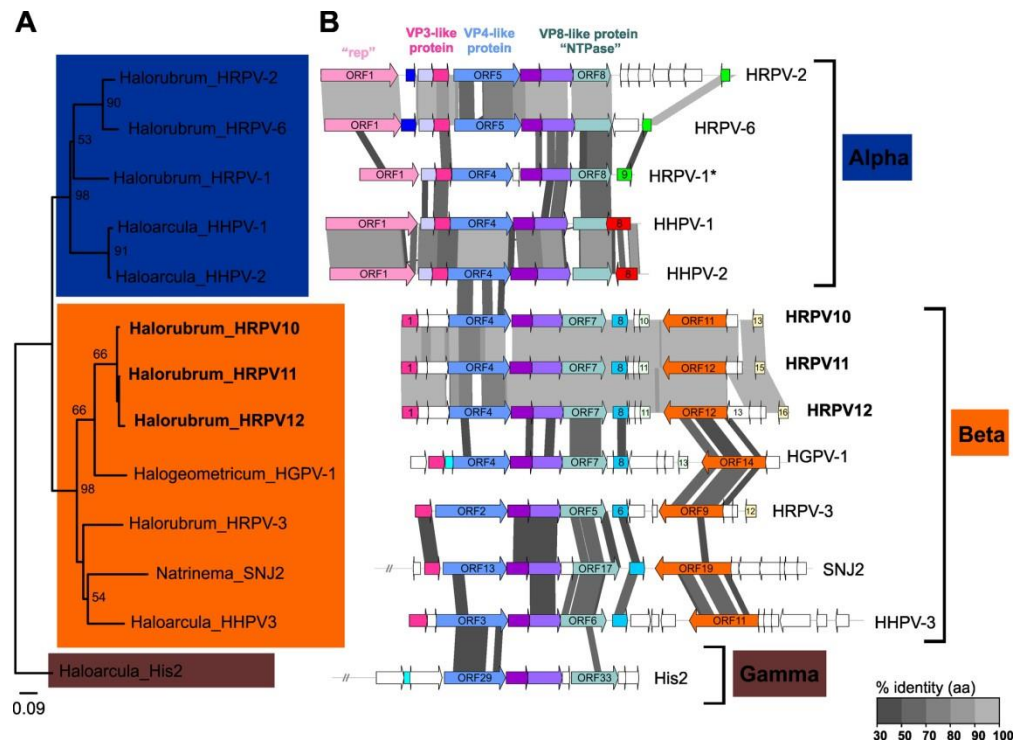
94x42mm (300 x 300 DPI)

Figure 5. HRPV10, HRPV11, HRPV12 and the members of the family *Pleolipoviridae*. (A) Phylogenomic tree was constructed using the Genome BLAST Distance Phylogeny (GBDP) strategy implemented in VICTOR (Meier-Kolthoff and Goker, 2017). The numbers above branches are GBDP pseudo-bootstrap support values from 100 replications. Clades corresponding to the genera *Alphapleolipovirus*, *Betapleolipovirus* and *Gammapleolipovirus* are colored light blue, beige, and grey, respectively. (B) Genomic comparison of pleolipoviruses depicted in panel A. Homologous genes are indicated with the same colors.
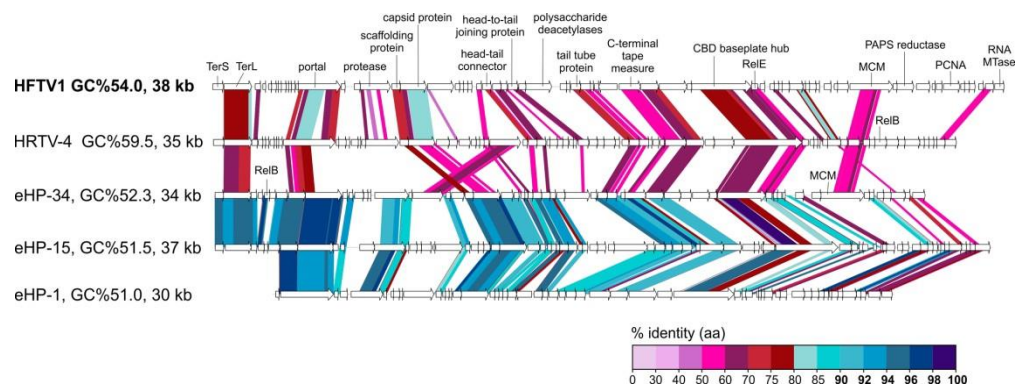
168x123mm (300 x 300 DPI)

Figure 6. Genomic comparison of HFTV1, HRTV-4, eHP-4, eHP-15 and eHP-1. Open reading frames (ORFs) are depicted as arrows indicating the directionality of transcription. When possible, the predicted functions are indicated above the corresponding ORFs. Shading connecting the ORFs indicates the amino acid sequence identity between the corresponding protein products; the color key is provided at the bottom of the figure. Abbreviations: TerS and TerL, small and large subunits of the terminase, respectively; CBD, carbohydrate-binding domain; PAPS, phosphoadenosine phosphosulfate; MCM, minichromosome maintenance helicase; PCNA, proliferating cell nuclear antigen; MTase, methyltransferase.

191x71mm (300 x 300 DPI)