Jorvik: a membrane-containing phage that will likely found a new family within Vinavirales

Pavol Bardy[1], Conor I.W. MacDonald[1], Roman Pantucek[2], Alfred A. Antson[1,3], Paul C.M. Fogg[3,4,5*]

[1]York Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York, YO10 5DD, United Kingdom

[2]Department of Experimental Biology, Faculty of Science, Masaryk University, Brno, 625 00, Czech Republic

[3]York Biomedical Research Institute, University of York, Wentworth Way, York, YO10 5NG, United Kingdom

[4]Biology Department, University of York, Wentworth Way, York, YO10 5DD, United Kingdom

[5]Lead contact

*Correspondence: paul.fogg@york.ac.uk

Summary

Although membrane-containing dsDNA bacterial viruses are some of the most prevalent predators in aquatic environments, we know little about how they function due to their intractability in the laboratory. Here, we have identified and thoroughly characterized a new type of membrane-containing bacteriophage, Jorvik, that infects the freshwater mixotrophic model bacterium *Rhodobacter capsulatus*. Jorvik is extremely virulent, can persist in the host integrated into the RuBisCo operon and encodes two experimentally verified cell wall hydrolases. Jorvik-like prophages are abundant in the genomes of alpha-proteobacteria, are distantly related to known viruses of the class *Tectiliviricetes,* and we propose they should be classified as a new family. Crucially, we demonstrate how widely used phage manipulation methods should be adjusted to prevent loss of virus infectivity. Our thorough characterization of environmental phage Jorvik provides important experimental insights about phage diversity and interactions in microbial communities that are often unexplored in common metagenomic analyses.

Introduction

Non-tailed bacteriophages are an understudied group of bacterial viruses[1–3]; an issue rooted in problems with sensitivity to chloroform, presence of proteins covalently bound to genomic DNA that hinder genome isolation and purification bias towards identification of tailed phages, e.g. non-tailed phages have a different buoyant density than used by most established protocols[4–7]. Nevertheless, recent microscopic, metagenomic and bioinformatic studies have revealed that non-tailed viruses are abundant in the environment, which emphasizes their potential importance and thus the necessity for further research[5,8,9].

Membrane-containing double-stranded (ds)DNA viruses with double jelly-roll fold capsids belonging to the PRD1-adenoviral lineage, *Bamfordvirae* kingdom, infect all domains of life[10–12]. Bacteria-infecting viruses of this kingdom belong to the class *Tectiliviricetes* and include some of the most prevalent predators of aquatic environments[5,8]. The protein shell of *Tectiliviricetes* phages is reinforced by its anchoring into an internal membrane, which is likely to provide extra stability against mechanical stress[13]. Based on the genome type, gene synteny and capsid characteristics, phages of this class are split into three families: *Tectiviridae, Autolykiviridae,* and *Corticoviridae*[14].

Out of the three families, *Tectiviridae* is the best-studied. Phages belonging to the *Tectiviridae* contain a linear dsDNA genome and create a transient membrane protrusion with a characteristic tubular shape for genome delivery[15]. Linear dsDNA phages belonging to the *Autolykiviridae* and *Tectiviridae* differ in their respective virion proteins, with the former being structurally similar to those of *Corticoviridae* phages. Unlike other phages of the lineage, *Autolykiviridae* phages often possess a broad host range, dominating the viral infection network of *Vibrionaceae* species[5]. In contrast to these two families, phages belonging to the *Corticoviridae* contain a circular dsDNA genome. Members of this family have been shown to infect marine bacteria[16] but only one, pseudoalteromonal phage PM2, has

been studied in detail[1,4,7,17,18]. Interestingly, double jelly-roll fold capsids may combine even

with ssDNA replication modules[19,20], and much of the lineage diversity remains to be

explored.

In this study, we identified a novel membrane-containing circular dsDNA phage,

Jorvik – named after the Viking designation of the city of York where this phage was

described. Jorvik is highly virulent and homologous prophages are abundant in

proteobacterial genomes. We performed experimental validation for several phage

components, critical for understanding how this group of phages function. Phylogenetic

analysis suggests that Jorvik-like phages form a new family-level taxon within the

Tectiliviricetes that shares a common ancestor with the *Autolykiviridae* and *Corticoviridae*

families as well as recently described fNo16-like phages[21]. We propose that all these phages

should be classified in the same order, *Vinavirales*, and we refer to them as such in the rest of

the manuscript. When working with an active virus stock, our experimental process focused

on the constant propagation of Jorvik with minimal culture manipulation. We suggest that this

approach can be applied to the cultivation and characterization of other membrane-containing

dsDNA viruses, allowing their hitherto unseen diversity to be revealed.

Results

Origin of phage Jorvik and its ability to spontaneously form more virulent variants

When cultivated in a defined RCV liquid medium, we observed that the *R. capsulatus* B10

strain spontaneously produced bacteriophage particles capable of infecting the *R. capsulatus*

SB1003, YW1, DE442 and St. Louis strains when they were grown in a rich YPS medium

(**Table 1**). However, even for a relatively short time, the phage particles were unstable in

liquid media, as demonstrated by a 90% decline in infective titre after just 24 h of incubation

in YPS (**Data S1.1**). The phage was thus routinely propagated on YPS agar plates with a soft

agar overlay, under aerobic conditions. For long-term storage, pieces of soft agar were taken

from confluently lysed plates by a sterile inoculation loop, transferred to an Eppendorf tube,

frozen and kept at -80°C.

The phage did not form plaques on cells growing under anaerobic conditions nor on

aerobic RCV plates. The optimal temperature range for plaque formation on YPS agar plates

was 20-34°C (**Figure S1A**). The plaques were 0.1 – 1.2 mm in diameter and had a turbid halo

with a less turbid centre (**Figure 1A**). After several rounds of passage with *R. capsulatus*

YW1, a plaque with a turbid halo and a completely clear centre appeared (**Figure 1B**). We

designated the phage Jorvik1, for the original turbid plaque variant, and Jorvik2, for the clear

plaque variant. It is notable that spontaneous induction of two phages from *R. capsulatus* B10

that were capable of infecting St. Louis was previously observed by Wall *et al.*[22]. These

phages were not characterized in any detail; therefore, we were not able to make any

comparison to phage Jorvik.

The titre of the phage was completely lost when the lysate was treated with

chloroform, which is indicative of membrane-containing phages. Surprisingly, the titre

declined substantially during repeated rounds of centrifugation and was negatively affected by

filtration (**Data S1.2**). Cryo-EM imaging showed that phage particles extracted from

confluently lysed plates were present in tightly packed formations rather than as individual

virions (**Figure 1C**), with most of the grid being free of particles. Tight arrangements have

been observed before when viral particles are at high density[27], however, here Jorvik only

reached relatively low titres in the range of 10^6 to 10^7 pfu.ml^{-1}. These results suggest that in

YPS media, Jorvik virions may behave as larger aggregates.

The Jorvik virion is an icosahedron with approximately 59 nm diameter, decorated with short spikes and containing a spherical density most likely corresponding to an inner membrane (**Figure 1D-E**). The genome of Jorvik1 consists of 8,762 bp circular dsDNA (**Figure 1F**). The GC content is 62.4 %, which is similar to the 66.6 % GC of the host strain *R. capsulatus* SB1003[23]. We identified 20 ORFs in the phage genome and predicted their function (**Table S1-2, Figure S2, Data S2**). The ORFs are encoded in three distinct operons, two on the forward strand and one on the reverse strand.

The 8,889 bp genome of Jorvik2 differs from that of Jorvik1 in two loci. The first is a duplication of a non-coding region (position 60-186) located between the phage integration attachment site, *attP,* and a gene predicted to encode a helix-turn-helix (HTH) domain-containing protein, Gp1. The second difference is a single nucleotide change inside the gene encoding Gp1 that results in a W56L single amino acid substitution. Since both differences were mapped within or near *gp1*, we conclude that the product of this gene plays a role in the lysis-lysogeny decision.

We did not detect tubulation morphologies resembling those reported for *Tectiliviridae* and *Autolykiviridae* phages, either for Jorvik particles in solution or attached to cell membranes[5,15]. Instead, we observed particles with ruptured virions from which DNA was escaping and cell membrane-attached particles that seem to lack part of the capsid at the attachment site (**Figure S1B-G**). It would be speculative to conclude that this rupturing is of biological significance, but nevertheless, these observations agree with the previously proposed hypothesis that the genome delivery mechanism of membrane-containing circular dsDNA phages differs from that of membrane-containing linear dsDNA phages[28].

134

Phage Jorvik1 integrates into the RuBisCO operon

136

The turbid zone of phage plaques is often caused by the integration of phage DNA into the

genome of the host bacterial cells, resulting in lysogenic immunity against superinfection[29,30].

To test if this was also the case for Jorvik1, the phage was plated on *R. capsulatus* St. Louis,

which produced the most turbid plaques. The turbid zone of a single plaque was picked with a

sterile tip and passaged on fresh plates to propagate phage-resistant mutants. Three individual

colonies were subsequently tested for spontaneous induction of the phage, for resistance to

Jorvik infection, and for the presence of Jorvik genes *gp3* (*M15)* and *gp15* (*slt)* by PCR (**Data**

S3.1). One of the three colonies tested, designated St. Louis C1 hereafter, produced positive

results for all three tests. After three passages, 10/10 St. Louis C1 colonies remained positive

for the presence of the virus, suggesting stable maintenance in the host. Interestingly, we were

unsuccessful in isolating a stable lysogen of the Jorvik2 variant when employing the same

approach, testing 20 colonies from different susceptible *R. capsulatus* strains.

149

Genomic DNA of both the parental *R. capsulatus* B10 strain and the lysogenic St.

Louis C1 strain was isolated and sequenced to identify the integration site of the phage. In

both St. Louis C1 and B10, the phage was integrated into the bacterial genome within the

RuBisCO cbbII operon between the genes encoding NAD(P)H-dependent quinone

oxidoreductase (*rcc01836*) and phosphoglucomutase (*rcc01837*). Jorvik does not possess a

predicted integrase gene but the *attP* and *attB* sequences show similarity to *dif* motifs

recognised by a cellular XerCD recombinase system[31,32], with the motif being conserved

among Jorvik-like phages (**Figure S3A**). The actual location of the *attB* in the genomes of

Jorvik-like phages seems variable. It was recently hypothesized that the host XerCD is

159 utilised for the integration of Gamma-proteobacterial membrane-containing dsDNA phages[21].

160 Our results suggest the same mechanism for Jorvik-like phages.

161

162 Phage Jorvik has a narrow host range and requires the pleiotropic regulator CtrA for
163 infection
164

165 Phages Jorvik1 and Jorvik2 had identical host ranges for the *R. capsulatus* strains tested here

166 (**Table 1**). Jorvik infected closely related strains, all of which had intact *attB* sites in their

167 genome (**Table 1**). Apart from the original B10 source strain and the St. Louis C1 lysogen,

168 PCR amplification of the *M15* and *slt* genes did not produce a product for any other strain

169 tested (**Data S3.1**); this suggests that there are no other lysogens of Jorvik among these

170 strains. Bioinformatic analysis revealed the presence of CRISPR-Cas spacers matching the

171 phage in diverse *R. capsulatus* isolates. These are strain A12 isolated in China, A52 and B41

172 isolated in Turkey and B6 isolated in the USA. The B6 strain showed to be resistant to

173 infection by Jorvik (**Table 1**), suggesting the system is functional.

174

175 To identify host genes that are required for successful phage infection, six available *R.*

176 *capsulatus* SB1003 gene knockout strains were tested against both Jorvik1 and Jorvik2

177 (**Table S3**)[33]. Of the six strains, only one carrying a deletion of the pleiotropic regulator *ctrA*

178 was consistently resistant to infection. The involvement of CtrA was verified by successful

179 restoration of infection by *in trans* complementation with a *ctrA* plasmid (**Table S3**). The

180 correct phosphorylation state of CtrA seems to be important for the phage infection because

181 the phospho-null (D51A) and phospho-mimetic (D51E) forms do not fully complement the

182 wildtype product; a similar effect was observed for *R. capsulatus* gene transfer agent

183 production[34]. We identified one putative CtrA-binding site, located in the intergenic region

184 between the structural and packaging operon near a putative regulatory repeat (**Figure S3B**).

185   It has been demonstrated that CtrA binds to the regulatory regions of several temperate-tailed

186   alpha-proteobacterial phages[35]. Our results support the hypothesis that CtrA is widely

187   involved in the regulation of lysogenic phages in the alpha-proteobacteria.

188

189   Growth characteristics show that phage Jorvik2 is highly virulent
190

191   Phage Jorvik adsorption kinetics followed a linear trend until approximately 60% of particles

192   had adsorbed; the adsorption rate for this 10-minute post-infection period was equal to

193   4.9×10^{-10} ml.min^{-1} (**Data S1.3**). This adsorption rate suggests that the phage recognises 10-

194   100 receptors per host cell[36]. Over a more prolonged adsorption period, a subfraction of

195   slower-adsorbing phage was identified. The adsorption efficiency of Jorvik reached 72 % by

196   the 20-minute post-infection time point (**Figure 2A**). Inefficient adsorption of a sub-

197   population is common in phages[37] and was reflected in the relatively broad rise period

198   observed in the phage growth curve (**Figure 2B**). The latent period of the phage is 80-95 min

199   and the titre peaks at ~130 min post-infection, with the inflexion point of the curve at around

200   110 min. These times are longer in comparison to *Alphatectivirus* phages infecting faster-

201   growing *E. coli* (inflexion point ~45 min)[38] or the *Corticovirus* phage PM2 that infects

202   *Pseudoalteromonas* (~70 min)[4], but shorter than that of tailed roseobacter phage RDJL phi1

203   (~140 min)[39]. To our knowledge, apart from Jorvik, RDJLphi1 is the only *Rhodobacteracaea*-

204   infecting phage with an estimated growth curve and more microbiological data are required to

205   assess what variability in propagation speed exists among these phages. For the growth

206   conditions tested, the burst size of Jorvik was estimated as 45 +/- 27 phage particles per

207   infected cell (**Data S1.4**).

208   Jorvik phage can effectively suppress growth of all three *R. capsulatus* propagation

209   strains – YW1, SB1003 and St. Louis (DE442 was not tested as it is an SB1003 derivative).

210   With an initial multiplicity of infection (MOI) in the range of 0.0005-0.005 (**Data S1.5**), the

bacterial density starts to decline 5-6 hours post-infection (**Figure 2C-E)**. Based on the phage

growth curve, we estimate this as the end of the third lytic cycle of the phage. After >20 hours

of incubation, phage-resistant cell growth is noticeable in all three strains (**Figure 2C-E**).

Since the estimated MOI required for the clearance of the cultures was strikingly low, we

tested the effect of cell growth state and the choice of titre method on the efficiency of phage

plating (**Data S1.2**). The data showed that both variables affected the titre estimation by

around two-fold, adjusting our MOI estimate to 0.001-0.01.

Phage virions survive repeated freeze-thaw and pH shifts

The stability of the Jorvik virions in response to temperature, freezing-thawing, pH and

salinity was tested. The phage was resistant to 30 min incubation at temperatures up to 45°C,

but the titre decreased rapidly when incubated at 47.5 and 50°C (**Figure 3A**). In the freeze-

thaw experiment the titre dropped to an average of 32% after a single cycle, compared to a

control incubated at room temperature (**Figure 3B**). Additional freeze-thaw cycles had only a

minor additional influence on the phage titre. Incubation for 24 h in YPS media at pH 5 to 9

did not affect phage stability compared to the pH 7 control. There was a slightly larger decline

in infectivity during incubation at pH=10 and a complete loss of titre in pH=4 media, even

after 90 minutes of incubation (**Figure 3C**). When diluted in media of increasing salinity, the

phage titre dropped to 33, 18 and 21% on average for YPS media with 0.1 M, 0.3 M and 0.6

M NaCl, respectively, compared to YPS media with 0 M NaCl. After the initial osmotic

shock, prolonged incubation in high salinity media did not lead to a larger decline of the titre

compared to the decline in salt-free media (**Figure 3D**).

Proteomic analysis confirmed an abundance of predicted structural proteins in the virion-enriched sample

237 To assess which gene products are part of the Jorvik virion and indirectly confirm their

238 function, an LC-MS/MS analysis was performed on a partially purified sample of Jorvik2 and

239 the relative abundance of the proteins was estimated (**Data S4**). The partially purified sample

240 was prepared by differential centrifugation as described in the Methods section. A thorough

241 purification was not possible due to the problems associated with different purification

242 procedures and prolonged storage described above. The structural operon of phage Jorvik

243 shares equivalent synteny and structural homology with the model phage PM2, thus we apply

244 the established phage PM2 nomenclature for the Jorvik gene products.

245   Gp9, which corresponds to the putative major capsid protein P2, was the most

246 abundant viral protein followed by the putative receptor-binding spike P1 (Gp14) and the

247 putative structural membrane protein P3 (Gp10). Structural proteins P1, P2 and P3 are also

248 the three most abundant proteins in phage PM2 virions[7,18]. Another 13 phage gene products

249 were identified in the sample, and all were present in comparable amounts to background host

250 bacterial proteins (**Data S4**). All 13 products were encoded in the structural and packaging

251 module of the genome. Four putative phage proteins were not detected by the MS analysis.

252 These include the HTH-domain protein Gp1, replicase Gp2, putative holin Gp20 and,

253 unexpectedly, a homologue of phage PM2 structural membrane protein P8 (Gp11). The

254 presence of the very short Gp4 in the sample could not be estimated, as the theoretically

255 acquired peptides after trypsin cleavage would be too short to analyse.

256

257 The phage encodes two lytic enzymes
258

259 Two gene products have predicted similarity to peptidoglycan hydrolases. One designated as

260 Slt is 231 amino acids (AA)-long, located in the structural module, and has a C-terminal

261 domain similar to the lytic transglycosylase P7 from phage PRD1 (HHpred; P27380, E-

262 value=9.7e-8, Score=71.57; region 102-225). The second designated as M15 is 184 AA-long,

263   located on the opposite strand packaging locus and is similar to peptidases from the M15

264   family (HHpred; PF08291.14, E-value=1.4e-21, Score=137; region 24-129). To confirm their

265   function in *R. capsulatus* cell wall digestion, the genes encoding these products were cloned,

266   overexpressed and the proteins purified using his-tag affinity chromatography. Both purified

267   proteins showed *R. capsulatus* cell wall degradation activity in zymogram and plate lysis

268   assays (**Figure 4, Data S3.2**).

269
270   Phylogenetic analysis classifies Jorvik-like phages as a novel family-level group within
271   the *Vinavirales* order
272
273   Phylogenetic analysis of two conserved proteins, the packaging ATPase P9 (Gp5) and the

274   major capsid protein P2 (Gp9) (**Data S5**), as well as Genome-BLAST Distance Phylogeny[40],

275   showed that phages Jorvik and *Marinomonas* phage YY form a unique family-level group

276   within the *Vinavirales* (**Figure 5).** This is consistent with the level of differences observed

277   during the manual inspection of genome synteny among the phages of *Vinavirales* (**Figure 6**).

278   The major distinction between Jorvik and other phages of *Vinavirales* lies in a different set of

279   regulatory genes, the replicase gene and lytic genes. The synteny of structural and packaging

280   loci is more conserved, with several Jorvik proteins identified as homologous phage PM2

281   counterparts using HHpred[41] (**Table S2**).

282
283   To estimate the prevalence of prophages similar to Jorvik integrated into sequenced bacteria

284   genomes, PSI-BLAST bioinformatic analysis of the two most conserved proteins, P9 and P2,

285   was used to identify sequences with both products encoded within a 10 kb window. In total,

286   849 prophage hits were identified, belonging to 737 unique bacterial strains and 436 unique

287   bacterial species (**Data S6.1**). These prophages were found predominantly in aquatic alpha,

288   beta and gamma-proteobacteria species (**Table S4**). We manually inspected contig sequences

289   from twelve selected representatives spread throughout the taxonomic lineages, and

290 confirmed that they all encoded additional PM2-like proteins in the P9/P2 locus (**Table S4**).

291 Clustering of P9 and P2 sequences showed the existence of putative prophage groups

292 different to the known phages of *Vinavirales*[42] (**Data S6.2-4**). Five prophages integrated in

293 the genomes of Alpha-proteobacteria clustered closely with Jorvik and *Marinomonas* phage

294 YY, whose host is likely misannotated (**Data S6.2-4, Table S1**).

Discussion

295
296

Jorvik-like phages represent a new family-level group of membrane-containing circular dsDNA phages

297
298
299

300 Phage Jorvik is a unique double jelly-roll virus. This double jelly-roll classification is

301 supported by the capsid diameter of around 59 nm, the presence of an inner membrane, the

302 decoration of pentons with spikes (**Figure 1D-E**) and bioinformatic comparison with the

303 archetypal *Corticovirus* phage PM2 (**Figure 6, Table S2**). To our knowledge, it is only the

304 second member of the order *Vinavirales* that has been characterized in detail, after phage

305 PM2[1,6,7,18]. Moreover, it is the first characterized freshwater phage of the group. This does not

306 come as a shock since our data demonstrate that established laboratory methods used for

307 phage propagation and storage are unsuitable for work with these types of phages. Indeed, if a

308 surreptitious lysogenic strain of phage Jorvik had not been available, the phage is likely to

309 have been lost a long time ago. The modifications to existing phage isolation and

310 manipulation methodologies, e.g. immediate processing of the isolation material without

311 storage, reduced centrifugation and filtration steps, propagation of the phage via soft agar

312 overlaying on a daily basis, long term storage of the phage in the form of frozen agar stabs,

313 can be applied to the discovery and study of other membrane-containing dsDNA phages.

314

The most striking finding was that repeated rounds of centrifugation as well as filtration led to

a rapid decline of the phage titre. Since centrifugation and filtration of samples is a common

procedure when isolating and handling environmental phages[43,44], many membrane-

containing dsDNA phages could easily be lost during purification. Also, due to practical

considerations, environmental samples are likely to be stored for various lengths of time prior

to analysis. As demonstrated here for Jorvik, storage at 4°C for even one day leads to an order

of magnitude loss of titre. Thus, the few membrane-containing dsDNA phages isolated so far

are likely to be a gross underestimate of their true abundance. Interestingly, even though the

survival of the phage in laboratory conditions was low, the virions were quite robust against

pH shifts and freeze-thawing cycles. This contrasts with phage PM2 where one cycle of

freeze-thawing led to >99% loss of titre[7]. The virions also seem to favour low salt conditions

reflecting the freshwater origin of this phage.

A good estimate of the environmental abundance of phages can be inferred from the presence

of homologous prophages residing in bacterial genome sequences. In the case of phage

Jorvik, more than 800 similar putative prophage sequences were found throughout the

proteobacteria phylum mainly in aquatic isolates. These included previously characterised

phages belonging to *Vinavirales*, which group outside of Jorvik-like phage cluster (**Figure 5,**

Data S5, Data S6.2), suggesting Jorvik-like phages form a separate taxon at the same level as

current *Corticoviridae* and *Autolykivirdae* families. Similar observations were recently made

for fNo16-like phages infecting *Vibrio* species[21], further suggesting that a major redefinition

of taxonomy within the order *Vinavirales* is required.

The phage Jorvik2 is capable of clearing liquid cultures of *Rhodobacter* strains in several

hours even when present at an MOI as low as 0.001 (**Figure 4C-E**). This makes Jorvik2

340 highly virulent even when compared to potentially therapeutic phages[45,46], and thus likely to

341 be ecologically relevant. While we corrected the MOI based on our estimate of the efficiency

342 of plating, one needs to consider that other unknown conditions may influence the real

343 efficiency of plating and the measured MOI might still be underestimated. Interestingly, the

344 fast-growing *R. capsulatus* strain St. Louis was capable of escaping clearance by the phage in

345 one biological replicate (**Figure 2E**). This suggests that cells entering later growth phases

346 become less susceptible to infection, which was supported by the observation that plaque

347 production was hampered when the phage was plated with late log-phase to stationary phase

348 cells compared to early log-phase cells (**Data S1.2**). No plaque formation nor clearing of the

349 culture was observed when cells were infected in RCV defined media or during anaerobic

350 cultivation. In all of these conditions, *R. capsulatus* cells tend to produce more extracellular

351 polysaccharide[47,48], which may hinder the phage's ability to infect.

352

353 Jorvik structural operon
354

355 The phage Jorvik genome is 8,762 bp, which is the smallest genome of the known membrane-

356 containing dsDNA phages, and contains three distinct operons. The largest operon is a

357 structural operon that is organized with equivalent synteny to other members of *Vinavirales*

358 (**Figure 6**). The major protein products are P1 (receptor-binding spike), P2 (major capsid

359 protein), P3, P7 and P8 (inner membrane protein), and P10 (unknown function). Interestingly,

360 the putative inner membrane protein P8 was the only PM2 homologue that was not identified

361 as part of the virion by mass spectrometry analysis. P8 is short and hydrophobic and thus hard

362 to detect using MS methods, therefore its absence could be a false negative and its presence in

363 the sample cannot be ruled out. Another interesting protein is a putative N-acetyltransferase

364 designated Gp12, which is encoded between P8 and P10. The protein is conserved among the

365 phages of *Vinavirales*, with the notable exception of phage PM2 (**Figure 6**)[42]. MS analysis

366   detected only a very low amount of Gp12 in the phage-enriched sample, suggesting that it is

367   not part of the virion but rather it is necessary for the acetylation of other structural

368   components, similar to the function of Gp14 from *Salmonella* phage P22[49].

369

370   Homologues of phage PM2 conserved inner membrane proteins P5 and P6 could not be

371   identified in Jorvik phages. In the genome, they are substituted with Slt and a hypervariable

372   region encoding several short hydrophobic proteins Gp16-19 (**Table S1**). Since Gp16-19 are

373   present in a structural module and were detected in the virion-enriched sample, they could be

374   structural virion membrane proteins. Their low sequence conservation might be a result of co-

375   evolution with bacterial host factors, for example, they could be involved in virus-host

376   membrane fusion. A putative holin gene, *gp20*, is located at the very end of the structural

377   operon. A homologue of this gene product from *Marinomonas* phage YY has predicted

378   similarity to the *R. capsulatus* gene transfer agent holin[50] (HHpred; PF11351.11, E-

379   value=0.03, Score=37.17; region 4-109 out of 118 AA).

380

381   Jorvik replication operon
382

383   The replication operon encodes two gene products. The first is the HTH domain-containing

384   protein Gp1, which we propose is involved in the lysis-lysogeny decision of the phage.

385   The Gp1 protein of virulent phage variant Jorvik2 carries a mutation in a key structural

386   residue and a duplication of the putative upstream regulatory region. It is likely that these

387   changes promote the lytic replication, which explains why Jorvik2 forms clear plaques and

388   why we could not isolate a stable lysogen of this variant.

389

390   The second protein encoded in the locus is the putative replication protein Gp2. Gp2

391   substantially differs from the replicases of characterised phages of the *Tectiliviricetes* (**Figure**

392   **6, Figure S2D**), having only 212 amino acid residues and two distinctive domains. In

393   comparison, both PM2 and PRD1 replicases have more than 500 residues. The N-terminal

394   domain of Jorvik replicase (34-116 AA) is similar to the lytic replication protein of coliphage

395   P1 (HHpred; P19654, E-value=2.1e-5, Score=55.9); while the C-terminal domain does not

396   resemble any characterised proteins. Due to the short length of Gp2, we speculate that

397   additional host factors are required for Jorvik genome replication and that the C-terminal

398   domain might be involved in the interaction with those factors. A recent metagenomic

399   analysis reported that the major distinguishing factor of *Vinavirales* phages is the type of

400   replicase protein[42], which appears to coevolve with a specific host to allow for optimal

401   function.

402

403   Jorvik packaging operon
404

405   The reverse strand packaging module encodes the packaging ATPase P9, putative endolysin

406   M15, P15-like repressor Gp7 and two hypothetical proteins, Gp4 and Gp6. The function of

407   Gp7 is predicted from the fact that its N-terminal domain shows high similarity to known

408   repressors including P15 of phage PM2 (HHpred; Q9XJS8, E-value=3.8e-6, Score=59; region

409   12-61 out of 149 AA) and cI repressor of phage lambda (HHpred; P03034, E-value=4.9e-5,

410   Score=51; region 5-58 out of 149 AA). The C-terminal domain is rich in prolines, and its

411   function is unknown. Upstream of this gene lies an intergenic region (sequence 3263-3448)

412   containing a putative CtrA-binding site (3265-3276), a non-coding repeat sequence identified

413   by blastn (3337-3384), and two promoters for expression of two oppositely facing operons

414   (**Figure 1F**). In phage PM2, the P15 repressor affects expression from two oppositely facing

415   promoters[51] and Gp7 might work similarly. We believe CtrA, Gp7 and this intergenic region

416   play a major regulatory role of the phage life cycle, but the exact molecular mechanism of the

417   regulation remains to be determined.

418

The function of protein Gp6 may lie in virion membrane assembly. The *gp6* gene precedes the

packaging ATPase ORF, which is an equivalent location to the membrane assembly factor gp-

h of other *Vinavirales* phages. It also possesses a weak similarity to protein P10, the gp-h

homologue in phage PRD1[42] (**Figure S2C**). Moreover, the protein was abundant in the

partially purified phage sample. This can be explained by the incorporation of the protein in

mature virions or virion assembly intermediates. The conservation of a virion membrane

assembly factor across genome packaging modules of *Tectiliviricetes* phages is striking and

more research is required to explain if and how the factor is involved in the packaging

process. The existence and potential function of the last hypothetical product of the operon,

Gp4, remains unknown.

429

Jorvik lytic enzymes

431

Phage Jorvik encodes two lytic enzymes. Slt (Gp15) is encoded in the structural operon while

M15 (Gp3) is encoded in the packaging operon. Both Slt and M15 were identified in the

phage-enriched sample by MS, with Slt being more abundant. Many phages encode two PG

hydrolases, a virion-associated hydrolase and an endolysin[52,53]. The former is required for

creating a pore in the cell wall during the injection of DNA into the host. The latter is required

for the release of infectious virions after their assembly. In the phages of *Vinavirales*,

functional homologues of Slt in the same genome position can be identified (**Figure 6**)[42]. In

contrast, M15 homologues were found only among Jorvik-like phages infecting Alpha-

proteobacteria and as such are unique. Based on the genome position and MS data, we

hypothesize Slt is a virion-associated hydrolase while M15 is an endolysin, analogous to the

P7 and P15 hydrolases of phage PRD1[54].

443

The isolation and detailed characterization of phage Jorvik prove that membrane-containing dsDNA phages of *Vinavirales* are widespread predators of aquatic niches. Modifications to classical laboratory methodologies and approaches described here can be applied in the cultivation of other membrane-containing dsDNA bacterial viruses and may drive a revolution in their study.

Limitations of the study

The instability of the phage in laboratory conditions might have affected the outcomes of some microbiological assays such as phage growth curve and burst size determination, leading to an underestimation of the real values. Also, due to this instability only a partially purified sample of the phage could be analysed by mass spectrometry, thus the virion association of individual components needs to be interpreted carefully. The study did not target the molecular mechanisms of phage integration and operon regulation, the mechanisms proposed were based on similarity with homologous proteins and the behaviour of mutant phenotypes. Due to very low sequence similarity among proteins of Tectiliviricetes, the support of several phylogenetic tree branches is low. For this reason, we did not provide any extensive interpretation of the topology of trees.

482
483 ## Author contributions

484 P.B. and P.F. designed the experiments; P.B, C.M. and P.F. conducted the experiments; P.B.,

485 C.M., R.P., A.A and P.F. performed the data analysis; P.B., C.M., R.P., A.A and P.F. wrote

486 the paper.

487

488 ## Declaration of interest

489 The authors declare no competing interests.

490

Main figures and legends

Figure 1: Morphological and genomic characteristics of phage Jorvik. A-B) Morphology of Jorvik1 (A) and Jorvik2 (B) plaques on the lawn of *R. capsulatus*, strain YW1. The scale bar is 2 mm. C) An aggregate of the phage obtained from a lysed plate wash, imaged by cryo-EM. The black dots are gold fiducials, scale bar represents 200 nm. D-e) Native capsids of the phage imaged using cryo-EM (D) and cryo-ET reconstruction (E) with head spikes depicted by blue arrows, scale bar is 50 nm. F) Genome map of the phage variant Jorvik1. The duplicated region of Jorvik2 adjacent to *gp1* is highlighted in black and the single nucleotide change in *gp1* is depicted by black dot. Genes encoding structural proteins (medium blue), acetyltransferase (light blue), hydrolases (light green), unknown function (yellow), holin (dark green), DNA-binding protein (magenta), replicase (orange), ATPase (dark blue) and P15-like repressor (dark magenta) are depicted. Genes shown in bold have sequence homologues in the database, white dots within ORF arrows depict predicted transmembrane regions. Putative regulatory repeat (grey), terminators (loop symbol) and *attP* site (X symbol) are highlighted. Operons are depicted in shades of orange in the centre of the map. See also **Figures S1-3** and **Data S2, S4** and **S7**.

Figure 2: Growth characteristics of phage Jorvik2. A) Adsorption curve on the propagating strain YW1 (red) and phage-resistant strain P12F (black). On the y-axis, values of plaque-forming units (PFUs) for specific time points relative to PFUs at time point 0 are shown. B) Growth curve of the phage on the propagating strain YW1 (red) and sterile media control (black). On the y-axis, PFUs for specific time points relative to PFUs at time point 0 are shown. C-E) Growth curves of strains YW1 (C), SB1003 (D) and St. Louis (E) with (red) and without the phage (black) are shown. The multiplicity of infection was below 0.01. The raw data are present in **Data S1.3-5.**

Figure 3: Phage survivability under different conditions. Individual measurements of biological replicates are plotted as points, their averages are plotted as Bars. A) Heating of the phage lysate for 30 min. B) Freeze-thawing of the phage lysate. C) Incubation of the phage in the medium with different pH. D) Incubation of the phage in the medium with different NaCl concentrations. The raw data are present in **Data S1.1** and **Data S1.6-8.**

Figure 4: Peptidoglycan hydrolysing activity of phage-encoded enzymes. A) SDS-PAGE gel of the lytic protein samples – lysozyme positive control (Lys), putative peptidase M15 (M15), putative soluble lytic transglycosylase (Slt) and negative controls protease 3C (3C) and bovine serum albumin (BSA). B) Zymogram with *R. capsulatus* SB1003 cell walls embedded in the gel matrix. The raw gels are present in **Data S3.2.** C) Plate lysis assay of protein samples dropped onto a layer of water agar with embedded cell walls.

Figure 5: Phylogenetic tree of phage Jorvik generated by Genome-BLAST Distance Phylogeny of the predicted amino acid sequence using the web server VICTOR [40]. The pseudo-bootstrap support values from 100 replications are shown next to nodes, with an average support equal to 90 %. The scale bar defines the branch length, scaled in terms of the D6 distance formula. The legend on the right defines different taxonomy rankings of the phages with the colour coding differentiating between different taxon groups that were assigned within a rank. The colour of the name depicts approved or officially proposed family ranks – *Autolykiviridae* (teal), Jorvik-like (magenta), *Corticoviridae* (green), fNo16-like (orange) and *Tectiviridae* (blue). See also **Table S4** and **Data S5-6.**

Figure 6: Genome synteny of membrane-containing dsDNA bacteriophages. Names of the phages with GenBank accession numbers are shown on the left. The colour bar on the right shows HHpred score for the pairwise alignment between corresponding homologues of Jorvik and PM2 phage, with raw values present in **Table S2.** The scale bar at the bottom shows the length of the genome in base pairs. Phages are grouped according to the replicase gene type. Genes are coloured according to their putative function as explained in the legend.

Main tables and legends

Table 1: Interaction between phage Jorvik and the tested host strain. See also **Table S3**.

Rhodobacter capsulatus strain	Plaque-forming ability[a]	Integrated phage[b]	CRISPR spacers match	*attB* in the bacterial genome	GenBank accession	Reference
SB1003	+	-	-	+	NC_014034	[23]
DE442	+	-	-	+	NZ_AYPR00000000	[24]
B10	-	+	-	+	JAOTPJ000000000	[25], this study
St. Louis	+	-	-	+	NZ_VIBE00000000	[25]
St. Louis C1	-	+	-	+	NA	this study
YW1	+	-	-	+	NZ_AYPY00000000	[24]
B6	-	-	+	+	NZ_AYQA0000000	[24]
YW2	-	-	-	+	NZ_AYPZ00000000	[24]
2.4.1	-	-	NA	NA	NA	[26]
H9	-	-	NA	NA	NA	[25]
P12F1	-	-	NA	NA	NA	[25]
SP18	-	-	NA	NA	NA	[25]
SP36	-	-	NA	NA	NA	[25]
Rhodobacter sphaeroides strain SCJ	-	-	NA	NA	NA	[25]

[a]same results were observed for variants Jorvik1 and Jorvik2; [b]tested by positive PCR amplification of genes Slt and M15, spontaneous induction of YW1-infecting phage particles when cultivated in RCV; +, positive result; -, negative result; NA, data not available

STAR+METHODS

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Paul Fogg (paul.fogg@york.ac.uk).

Materials availability

All unique reagents or materials generated in this study will be made available on request by the lead contact, but we may require a completed materials transfer agreement if there is potential for commercial application.

Data and code availability

- Nucleotide sequence data reported are available in DDBJ/ENA/GenBank under the accession number OP588643 for *Rhodobacter* phage Jorvik, the Third Party Annotation Section of the DDBJ/ENA/GenBank databases under the accession number TPA: BK062760 for *Marinomonas* phage YY. *Rhodobacter capsulatus* B10 strain Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAOTPJ000000000. The version described in this paper is version JAOTPJ010000000. All datasets are publicly available as of the date of publication. All AlphaFold2-predicted models which are discussed within the manuscript are attached as **Data S7**.
- This paper does not report original code.
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial strains

The bacterial strains used in this study are summarized in **Table 1** and **Table S3**. In addition, several *Escherichia coli* strains used for cloning are specified in the description of the

574 corresponding methods. *Rhodobacter* strains were grown either in minimal media RCV[71] or

575 rich media YPS[22], *E. coli* strains were grown in LB. *Rhodobacter* strains complemented for

576 the production of CtrA were cultivated in the presence of 10 μg.ml^{-1} of kanamycin. Unless

577 otherwise stated, *Rhodobacter* strains were incubated aerobically at 30°C and *E. coli* strains at

578 37°C. For anaerobic cultivation, bacteria were incubated in sealed 15 ml screw-cap glass

579 tubes filled with YPS up to the rim or YPS plates in Anaerocult® A jars (Merck Millipore).

580 The samples were incubated at 30°C in an illumination cabinet, placed 30 cm from three 40

581 W light-emitting tubes (Panasonic FL40SS・W/37c).

582

583 Bacteriophage Jorvik
584

585 The phage was propagated using YPS soft agar overlaying. The overnight culture of the

586 propagation strain was mixed with phage lysate, then 3 ml of soft agar (YPS + 0.7% w/v agar)

587 were added and overlaid on a YPS agar (1.5% w/v) plate. The plate was incubated at 30°C

588 overnight. The soft agar containing phage particles was harvested using a bacterial loop and

589 resuspended in YPS media. The tube was vortexed vigorously and centrifuged at 6000 RCF

590 for 4 min to separate the agar from the phage solution. Repeated or high-speed centrifugation

591 was avoided as we observed that these procedures led to substantial loss of plaque forming

592 units. Filtration of phage lysates was not performed as this also led to a large reduction of

593 phage titre; instead, the supernatant was transferred to a clean tube using a 30G gauge needle,

594 stored in the fridge and used for any experiments within 8 h. For pH and salt stability assays,

595 soft agar containing the phages was resuspended in 5-10 times of excess of YPS by passing

596 through a 30G gauge needle and directly used for experiments. This led to one order of

597 magnitude higher titres, but with the presence of residual agar and cell debris in a lysate.

598

599 Unless otherwise stated, the phage titre was estimated by a drop assay. Here, 10 µl of diluted

600 samples were dropped onto freshly solidified soft agar containing early log phase culture of

601 *Rhodobacter capsulatus* YW1 (OD$_{600nm}$ = 0.1-0.25) and left with an open cover for 7-10

602 minutes in a flow hood to dry. For estimating the titre from soft agar overlay, the host cell

603 culture was mixed with diluted phage samples before the addition of soft agar and

604 subsequently plated. For the estimation of efficiency of plating for bacterial cultures at

605 different growth phases, OD$_{600nm}$ of log and stationary phase cultures was measured, cells

606 diluted to OD$_{600nm}$=0.10 in YPS and 250 µl of the culture was used for phage plating either by

607 soft agar overlay or drop assay as stated.

METHOD DETAILS

Standard laboratory handling stability assay

611 For the chloroform treatment, 5 % v/v of chloroform were added to the lysate. The sample

612 was vortexed, incubated for 10 minutes on a bench, centrifuged at 6000 RCF for 4 min, and

613 100 ul of the supernatant from the top of the tube was transferred to a new tube and titred. For

614 the phage filtering, Merck Millex®GP PES Membrane 0.22 µm (Merck Millipore), Minisart

615 NML Plus cellulose acetate membrane 0.45 µm (Sartorius), yellow syringe filters PES

616 membrane 0.45 µm (Starlab), Nylon 66 membrane 0.2 µm (Supelco) and Durapore® PVDF

617 membrane 0.22 µm (Merck Millipore) filters were used to filter 2 ml of phage lysate. For

618 centrifugation experiments, 500 ul of phage lysate were spun at 6000 RCF for 4 min, and 100

619 ul of the supernatant from the top of the tube were subsequently transferred to a fresh tube

620 and titred.

621

Isolation of lysogenic strain

624 Phage lysate was mixed with the *Rhodobacter capsulatus* St. Louis overnight culture and

625 plated using a soft agar overlay. After the overnight incubation, several plaques were picked

626   with a sterile tip and streaked onto fresh plates and incubated for two days to get individual

627   phage-resistant colonies. Single colonies were restreaked onto a fresh plate and subsequently

628   inoculated into RCV. After two days of incubation in RCV, the cells were spun, and the

629   supernatant was dropped onto St. Louis WT strain to discover phage-inducing lysogens.

630   Colonies positive for phage production were subsequently tested for the presence of the phage

631   genes *gp15* (*slt)* and *gp3* (*M15)* by PCR. To verify the colonies do not come from the parental

632   B10 strain, after obtaining the lysogen sequence, the average nucleotide identity between St.

633   Louis C1, B10 and St. Louis WT strains was calculated using OrthoANIu web tool [70]. For St.

634   Louis C1 and B10, contig sequences obtained from Illumina sequencing were used as an

635   input. For St. Louis WT, contig sequences deposited under GenBank project

636   NZ_VIBE00000000.1 were used. This resulted in 99.95 and 99.47 % average nucleotide

637   identity when St. Louis C1 was compared to St. Louis WT and B10 respectively.

638

639   Growth characteristics
640

641   The growth of different strains in presence of phage was determined using OD_{600nm}

642   measurement on Spectrostar Nano Microplate Reader (BMG LabTech). The strains were

643   grown to OD_{600nm} 0.2 to 0.5 and diluted to a final OD_{600nm}=0.1 in 200 μl of media containing

644   between 2 and 7×10^4 phage particles per well (**Data S1.5**). The plate was incubated at 30°C,

645   and the OD_{600nm} was measured at 30 min intervals for 20 h. The plates were shaken orbitally

646   at 100 rpm for 10 s before each measurement. The range of phage titre rather than specific

647   titre value was used for growth experiments because of the inherent laboratory instability, a

648   fresh phage lysate was used for each experiment without previous knowledge of its precise

649   titre.

650

651 The growth curve of phage was determined using an approach similar to Kropinski, 2018[72].

652 YW1 strain was used as a host. Log phase cultures (OD_{600nm} = 0.2-0.5) diluted to OD_{600nm} =

653 0.1 and 2×10^4 to 7×10^4 phages were added to 3 ml of final culture volume. The mixture

654 was then incubated at 30°C and gently vortexed before sampling each time point. Ten and

655 hundred-fold dilutions of the infected cells were prepared one hour post-infection, which was

656 still in the lag phase of the phage. For sampling, 200 μl aliquots were spun at 12 000 RCF for

657 2 min and plated with log-phase YW1 strain using soft agar overlaying. Plaques were counted

658 the next day. The burst size was calculated as the difference between the averages of plaque-

659 forming units before and after the lytic burst.

660

661 For the adsorption assay, an approach similar to Kropinski, 2009[36] was employed. In brief,

662 100 μl of between 1×10^5 and 9×10^5 phage was added to 900 μl of the log-phase grown

663 culture, with a final OD_{600nm} of 0.1. The mixture was incubated at 30°C, shaking at 90 rpm.

664 The aliquots were diluted 20 times in ice-cold YPS, spun at 12 000 RCF for 2 min and 200 μl

665 were plated with a 100 μl of log-phase YW1 culture using the soft agar overlay. Plaques were

666 counted the next day. To estimate the adsorption rate, the equation from Kropinski, 2009[36]

667 was used.

668

669 Stability characteristics
670

671 For the heat stability assay, 250 μl of phage lysate was incubated in a heat block set at a

672 specific temperature for 30 min. The titre of the phage was then estimated using a soft agar

673 drop assay. For the freeze-thawing experiment, 250 μl of phage lysate with the titre between 1

674 $\times 10^5$ and 1×10^7 PFU.ml^{-1} was frozen at -80°C for 30 min and then thawed at room

675 temperature for another 30 min. After the last thawing cycle, the phage titre of all samples

676 was estimated by a soft agar drop assay. For the pH and salinity experiment, YPS media was

677 modified to different salt concentrations and pH by the addition of NaCl and HCl/NaOH

678 respectively. Precise pH was determined using pH electrode InLab Easy BNC (Mettler

679 Toledo) and adjusted as necessary. The phage lysate with the titre between 1×10^5 and $1 \times$

680 10^7 PFU.ml^{-1} was diluted 1:19 with the modified YPS media and incubated for 1.5 h at room

681 temperature. After plating the phage, the lysates were transferred to the fridge and the titre

682 was estimated again after an additional 22.5 h (24 h in total) of incubation in the modified

683 YPS media. All the growth and stability experiments were done in three independent

684 biological replicates, utilising different overnight cultures and different phage lysates.

685

686 DNA isolation and sequencing
687

688 The genome of Jorvik1 was isolated from 18 ml of lysate with a titre of 1×10^7 PFU.ml^{-1}. The

689 phage was pelleted by the addition of PEG8000 to a final concentration of 10% (w/v),

690 incubation at room temperature for 10 min and centrifugation at 10,000 RCF for 10 min. The

691 supernatant was discarded, and the pellet was resuspended in 500 μl phage buffer (50 mM

692 NaCl, 10 mM Tris pH7, 10 mM MgSO$_4$). Cellular nucleic acids were removed by digestion

693 with 0.2 mg.ml^{-1} DNaseI and 0.02 mg.ml^{-1} RnaseA and incubation at 37°C for 1 h. EDTA

694 was added to a final concentration of 50 mM, then encapsidated DNA was extracted using the

695 "Purification of Nucleic Acids by Extraction with Phenol:Chloroform protocol" [73]. Purified

696 DNA was linearized with blunt ends using EcoRV (New England Biolabs) and inserted into

697 the ZeroBlunt Topo vector (Thermo Fisher Scientific). The genome of Jorvik2 was isolated

698 from 6 ml of a lysate with the titre 4×10^6 in a similar manner. The integrity and amount were

699 sufficient for direct sequencing. The genomic DNA of strains B10 and St. Louis C1 were

700 isolated from 1.5 ml of stationary aerobic culture grown in YPS according to the published

701 manual[73]. All the samples were sequenced using Illumina sequencers (HiSeq/NovaSeq) and a

702   250 bp paired-end protocol. Genome sequencing was provided by MicrobesNG

703   (http://www.microbesng.com).

704

705   Estimation of phage integration site in St. Louis C1
706

707   The raw reads of St. Louis C1 were searched for the integration site by filtering for reads

708   containing the first and last 17 nt of the NODE_41 contig sequence, which contained the

709   circular sequence of the phage, using the "grep" command. This resulted in 596 reads,

710   including the reverse complements. The sequence upstream to the start of the NODE_14

711   contig sequence and downstream to the end of the sequence was analysed in Vim text editor,

712   resulting in two populations, one (479 reads) corresponding to the phage circular sequence

713   and the other (91 reads) corresponding to the bacterial integration site in *cbbII* operon. The

714   upstream/downstream sequence of the remaining 26 reads was shorter than 6 nt and these

715   were not assessed.

716

717   *R. capsulatus* genes knockouts
718

719   Knockouts in *R. capsulatus* were created by gene transfer agent mediated gene replacement[34].

720   pCM66T plasmid constructs were created with a gentamicin resistance cassette flanked by

721   500-1000 bp of DNA from either side of the target gene (**Table S5**). Assembly was achieved

722   by a one-step, four-component NEBuilder reaction and transformation into NEB 10-beta cells

723   (New England Biolabs). All oligonucleotides were obtained from IDT and designed with an

724   optimal annealing temperature of 60°C when used with Q5 DNA Polymerase (New England

725   Biolabs). All cloning reactions were carried out with NEBuilder according to the

726   manufacturer's guidelines (New England Biolabs).

727

728   Deletion constructs were introduced into the *E. coli* conjugation strain S17-1 and then

729   transferred to the *R. capsulatus* gene transfer agent hyperproducer strain DE442 by solid

730   phase conjugal mating. One millilitre aliquots of overnight cultures of the *E. coli* S17-1 donor

731   and *Rhodobacter* recipient strains were centrifuged at 5000 RCF for 1 min, washed with 1 ml

732   YPS medium, centrifuged again and resuspended in 200 µl YPS. Ten microlitres of

733   concentrated donor and recipient cells were mixed and spotted onto YPS agar or spotted

734   individually as negative controls. Plates were incubated overnight at 30°C. Spots were

735   scraped, suspended in 100 µl YPS broth and plated on YPS + 100 µg ml^{-1} rifampicin

736   (counter-selection against *E. coli*) + 10 µg ml^{-1} kanamycin (plasmid selection). Plates were

737   incubated overnight at 30°C then restreaked onto fresh agar to obtain single colonies.

738

739   A standard gene transfer agent bio-assay[74] was carried out to replace the intact chromosomal

740   gene with the deleted version using 10 µg.ml^{-1} gentamicin for selection. Successful knockouts

741   were confirmed by Sanger sequencing (Eurofins Genomics).

742

743   Bioinformatics and data visualization
744

745   ORFs were identified using *ab initio* prediction of Prokka 1.14.5[56] and GeneMarkS 4.28[57]

746   with the "phage" algorithm. The annotation and circular maps were done in Artemis[58]. The

747   predicted function of protein was based on primary sequence similarity using blastp[59], the

748   comparison of profile hidden Markov models using HHpred[41] and predicted tertiary structure

749   similarity using a combination of AlphaFold2[60,61] and DALI search[62]. The terminators were

750   predicted using the Genome2D web tool

751   [http://genome2d.molgenrug.nl/g2d_pepper_transterm.php].

752

753   Psi-blast search against bacterial and viral taxids using matrix BLOSUM=45 was performed

754   using the less conserved major capsid protein P2 sequence until convergence[59,63]. Then, a list

755   of nucleotide seqIDs encoding proteins found in all iterations was determined using NCBI

756   batch Entrez search against identical protein groups database

757   [https://www.ncbi.nlm.nih.gov/sites/batchentrez]. Unique Gis encoded by the longest

758   nucleotide sequence were filtered and a list of proteins encoded within the sequences was

759   generated using NCBI e-utilities[63]. This list served as a subject for psi-blast with the

760   packaging ATPase P9 until convergence using matrix BLOSUM=45. The position in the

761   nucleotide sequence of these new hits was retrieved using a batch Entrez search and compared

762   to the original major capsid protein P2 hits. Hits originating from the same nucleotide

763   sequence with less than 10000 nt from each other were filtered using awk script and

764   considered putative prophage sequences. The taxonomy information of the hits was retrieved

765   using the NCBI tax identifier

766   [https://www.ncbi.nlm.nih.gov/Taxonomy/TaxIdentifier/tax_identifier.cgi]. The putative

767   CtrA-binding site and *dif* motifs were identified using consensus sequences and regular

768   expression search tool [https://regex101.com/].

769

770   Phylogeny analysis
771

772   Phylogeny analysis was performed according to Grybchuk *et al.*[75]. In brief, a list of

773   homologous proteins to those from Jorvik was generated by PSI-BLAST. The search was

774   performed using the matrix BLOSUM=45 and the default settings until converging. The

775   taxonomy database used for the search was *Tectiliviricetes* in the case of the major capsid

776   protein P2, Alpha-proteobacteria (filtered out after the convergence) and *Tectiliviricetes* in the

777   case of the packaging ATPase P9. The complete sequences were then aligned using MAFFT

778   v7.487 e-ins-I algorithm[64]. The uninformative columns were cropped using trimAl v1.4

779   algorithm automated1[65] and the phylogenetic tree was computed by maximum likelihood

780   method using software iqtree2 v2.0.6[66] with a bootstrap test. A similar approach was

781   performed using P9 and P2 sequences from **Data S6.1** as an input, with the final tree

782   computed using fast bootstrap test. The trees were visualized using FigTree v.1.4.4.

783   [http://tree.bio.ed.ac.uk/software/figtree/].

784

785   For the Genome-BLAST Distance Phylogeny, genomic sequences classified to

786   *Autolykiviridae*, *Corticoviridae* were retrieved from GenBank and together with the sequence

787   of Jorvik used as a nucleotide input to VICTOR web server

788   [https://ggdc.dsmz.de/victor.php][40]. The phylogeny was performed on a predicted amino acid

789   sequence under settings recommended for prokaryotic viruses, using a recommended greed-

790   with-trimming algorithm of formula D6.

791

792   Proteomic analysis
793

794   The phage sample for proteomic analysis was prepared in the same way as for the genome

795   isolation. After pelleting and resuspending in B buffer, extra polishing was performed. Here,

796   most remaining bacteria were separated by 6 000 RCF/4 min/8°C spin. Phage particle

797   aggregates were enriched from the supernatant by an additional 17 000 RCF/25 min/8°C spin,

798   resuspended in B buffer to final A_{280nm} ~ 2 and used for LC-MS/MS analysis performed by

799   The Centre of Excellence in Mass Spectrometry, University of York. Here, the phage-

800   enriched sample was run 1 cm into a NuPAGE 10% BisTris gel (Invitrogen) and stained with

801   SafeBLUE stain (NBS Biologicals). The stained gel segment was excised, destained, reduced

802   with DTE and alkylated with iodoacetamide. Protein was in-gel digested using Promega

803   sequencing grade trypsin and incubation overnight at 37°C. The resulting peptides were

804   analysed by LC-MS/MS over a 1 h acquisition with elution from a 50 cm EN C18 PepMap

805 column (Thermo Fisher Scientific) driven by an mClass UPLC (Waters) onto an Orbitrap

806 Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) operated in DDA mode. MS1

807 spectra were acquired at high resolution in the Orbitrap mass analyser. MS2 spectra were

808 acquired in TopSpeed mode in the linear ion trap after HCD fragmentation.

809

810 LC-MS data in .raw format were imported into Progenesis QI for peak picking with tandem

811 mass lists exported in .mgf format for database searching using Mascot. Data were searched

812 against the combined Jorvik and NCBI *R. capsulatus* YW1 databases appended with common

813 proteomic contaminants. Peptide matches were filtered to 1% FDR using the Percolator

814 algorithm and as determined empirically by comparison to a reverse database search.

815 Accepted peptide identifications were imported into QI and associated with the LC-MS

816 chromatographic data. Results were further filtered to require a minimum of two unique

817 peptides per protein identification. As a metric for relative abundance between proteins the

818 Top3 approach has been used, whereby the best three responding peptide precursor ion areas

819 for each protein are extracted and compared.

820

821 Electron microscopy
822

823 For cryo-EM grids, a confluently lysed plate of phage was washed with 2 ml of RCV using a

824 pipette and the media was spun at 4000 RCF for 6 min. The pellet, containing aggregates of

825 virus particles as well as particles attached on host membranes, was resuspended in the

826 residual amount of liquid, mixed with 6 nm colloid gold fiducials (Auriol), applied onto glow-

827 discharged 200 mesh copper Quantifoil® grids and plunge frozen in liquid ethane. The grids

828 were imaged using Glacios TEM (Thermo Fisher Scientific) operated at 200 kV at the cryo-

829 EM facility of the University of York.

830

Tomography data collection and reconstruction

Cryo-electron tomography data were acquired at ×73,000 magnification and -5 μm defocus using -60,60° dose symmetric tilt series acquisition with 3° increment and a total dose of 2.9 $e^-\text{Å}^{-2}$ per each tilt image. The tilt images were aligned using fiducial alignment in IMOD 4.11.20 Tomography package Etomo[67]. Subsequently, the tilt series were 6 times binned, reconstructed into 3D tomograms and filtered using the non-linear anisotropic diffusion algorithm of Etomo.

Production of the lytic enzymes

The *M15* and *slt* genes were cloned using HiFi assembly (NEB) of PCR-generated products with Lic+ plasmid[55]. The primers used for amplification are summarised in **Table S5**. The assembled plasmids were transformed into *E. coli* Stellar cells (Takara Bio) and individual colonies were verified for the presence of the insert by sequencing (Eurofins Genomics GATC). The plasmids were subsequently isolated from verified colonies and transformed to expression strain *E. coli* Rosetta (DE3) pLysS (Novagen). The expression was induced with 0.1 mM IPTG at 25°C (M15) and 20°C (Slt) respectively. The pellets were resuspended in buffer B_M15 (Tris 20 mM pH=7.5, NaCl 100 mM, $CaCl_2$=5mM, Imidazole = 20 mM) and B_Slt (Tris 20 mM pH=8.5, NaCl 400 mM, $CaCl_2$=5mM, Imidazole = 10 mM), sonicated using Sonoplus HD2070 (Bandelin) and purified using HisTrap 5ml FF column (GE Healthcare) and step imidazole gradient. The fraction containing proteins was pooled, buffer exchanged until reaching the original imidazole concentration and the his-tag was cleaved overnight in presence of 1mM DTT and 20 μg of 3C protease. His tag and the protease were removed from the protein samples by 20 min incubation with Super Ni-NTA resin (Generon).

857 Zymogram and the plate lysis assay

858

859 Zymogram and plate lysis assays were performed according to Benesik *et al., 2018* [52]. In

860 brief, the analysed proteins were diluted in B_M15 or B_Slt buffer to a concentration of 0.6

861 mg.ml^{-1}, based on Bradford protein assay calibrated on BSA standard. For the zymogram,

862 protein samples were loaded into SDS-PAGE gels, containing 15 % acrylamide and 200 µl of

863 *R. capsulatus* SB1003 crude peptidoglycan, prepared according to Fogg *et al., 2012* [76]. The

864 gel ran at 150 V for 75 minutes, was washed three times for 15 minutes in fresh distilled

865 water, placed in a fresh container with distilled water and incubated at 30°C/40 rpm for 3

866 hours. The gel was then stained with a solution of 0.1% Methylene Blue and 0.001 M KOH

867 for 3 hours and destained in distilled water. For the plate lysis assay, 500 µl of *R. capsulatus*

868 SB1003 crude peptidoglycan were incorporated into 4 ml of 0.4% (w/v) dH$_2$O agar and

869 poured on a clean petri dish. Protein samples were then dropped onto the plate and incubated

870 for 3-12 hours at 30°C until clear zones of peptidoglycan degradation appeared.

871

872 QUANTIFICATION AND STATISTICAL ANALYSIS

873 The structural figures were created using UCSF ChimeraX (*67*). All the charts were plotted

874 using the ggplot2 system (*68*) of the package R [https://www.r-project.org/].

875

876

Supplemental Data titles and legends

Data S1. Raw data of microbiological assays, related to Figures 2 and 3. S1.1) Estimation of phage Jorvik stability during incubation in media with a different pH. S1.2) Phage Jorvik stability during common laboratory manipulations. S1.3) Raw data related to the phage adsorption assay. S1.4) Raw data related to the phage growth curve. S1.5) Raw data related to *R. capsulatus* growth curve with and without the addition of phage Jorvik. S1.6) Estimation of phage Jorvik stability during heating. S1.7) Estimation of phage Jorvik stability during freeze-thawing. S1.8) Estimation of phage Jorvik stability during incubation in media with a different salt concentration.

Data S2. Identified sequences of putative gene products from Jorvik-like prophages, which are not annotated in GenBank, related to Figure 1 and Table S1.

Data S3. Images of gels related to Figure 4. S3.1) Colony PCR amplification of Jorvik genes M15 and slt performed on Rhodobacter strains. a) Demonstration of the presence of the annotated genes in Rhodobacter capsulatus strain St. Louis C1. b-c) Verification of slt gene presence in 4/4 individual colonies of R. capsulatus B10 strain and 10/10 individual colonies of St. Louis C1. d) Verification that both slt and M15 are present in two individual St. Louis C1 colonies. e-f) PCR amplification of M15 (e) and slt (f) from a bank of Rhodobacter strains. Ladders 100bp and 1kb+ are from New England Biolabs. S3.2) Images of gels related to the zymogram experiment. a) Slt sample is composed of two bands as a result of an ineffective cleavage of the his-tag by 3C protease. The red arrow highlights the uncleaved product and the grey arrow the cleaved product. The expected molecular weights of the product are shown in numbers. Gel lanes descriptions: M, marker PageRulerTM 10-250 kDa (Thermo Fisher Scientific); 1, metal affinity purified Slt with the sample loaded in higher (1a) and lower (1b) concentration; 2, purified Slt with 3C protease after overnight incubation; 3, pellet bead fraction of purified Slt with 3C protease; 4, a supernatant fraction of purified Slt with 3C protease with the sample loaded in higher (4a) and lower (4b) concentration. b) The complete image of the zymogram from which insets are shown in Figure 4.

Data S4. MS analysis of partially purified Jorvik2 virions, related to Figure 1. The proteins are ordered according to their abundance estimated from the area of the top three peptides identified for the protein. ID Filter = Mascot Percolator to 1% FDR and QI min 2 peptides. Phage proteins are highlighted in grey.

Data S5. Phylogenetic trees of phage Jorvik based on conserved protein sequences, related to Figure 5. S5.1) Phylogenetic trees of phage Jorvik based on the sequence of major capsid protein P9. Maximum likelihood tree based on PSI-BLAST searches in the taxid "viruses". The confidence estimates of branches are shown next to nodes, with values >90% shown in cyan. The low confidence separation of branches near the tree centre of the tree is enclosed in blue oval. The colour of the name depicts approved or officially proposed family ranks – Autolykiviridae (teal), Jorvik-like (magenta), Corticoviridae (green), fNo16-like (orange), Tectiviridae (blue), Turriviridae (gold) and unassigned (gray). The scale bar defines the branch length as the number of substitutions per site. All the used sequences for computing the tree are present in **Data S5.2.** S5.2) List of sequences used for building the phylogeny

923   trees of Jorvik based on the packaging ATPase P9. S5.3) Phylogenetic trees of phage Jorvik
924   based on the sequence of major capsid protein P2. Maximum likelihood tree based on PSI-
925   BLAST searches in the taxid "viruses". The confidence estimates of branches are shown next
926   to nodes, with values >90% shown in cyan. The low confidence separation of branches near
927   the tree centre of the tree is enclosed in blue oval. Jorvik-like phages are highlighted in bold.
928   The colour of the name depicts approved or officially proposed family ranks – Autolykiviridae
929   (teal), Jorvik-like (magenta), Corticoviridae (green), fNo16-like (orange) and unassigned
930   (gray). The scale bar defines the branch length as the number of substitutions per site. All
931   the used sequences for computing the tree are present in **Data S5.4**. S5.4) List of sequences
932   used for building the phylogeny trees of Jorvik based on the major capsid protein P2.
933
934   **Data S6. Analysis of sequences from nucleotide databases which encode homologs of**
935   **Jorvik proteins P2 and P9 in proximity, related to Figure 5.** S6.1) List of sequences from
936   nucleotide databases which encode homologs of Jorvik proteins P2 and P9 in close
937   proximity. S6.2) Clustering of conserved proteins of putative prophages similar to Jorvik. a-b)
938   unrooted radial trees of P9 protein (a) and P2 protein (b) based on putative prophage
939   sequences from **Data S6.1**. Positions of characterised phages are highlighted. The individual
940   tree files are attached in **Data S6.3** for P9 and **Data S6.4** for P2 respectively. S6.3) Raw tree
941   file of P9 sequences, plotted in **Data S6.2a**. S6.4) Raw tree file of P2 sequences, plotted in
942   **Data S6.2b**.
943
944   **Data S7. Files of AlphaFold2-generated models discussed in the manuscript with plots of**
945   **their predicted local distance difference test (pLDDT) per residue, related to Figures 1 and**
946   **S2.**
947

References

1. Mäntynen, S., Sundberg, L.-R., Oksanen, H., and Poranen, M. (2019). Half a Century of Research on Membrane-Containing Bacteriophages: Bringing New Concepts to Modern Virology. Viruses *11*, 76. 10.3390/v11010076.

2. Krishnamurthy, S.R., Janowski, A.B., Zhao, G., Barouch, D., and Wang, D. (2016). Hyperexpansion of RNA Bacteriophage Diversity. PLoS Biol. *14*, e1002409. 10.1371/journal.pbio.1002409.

3. Hopkins, M., Kailasan, S., Cohen, A., Roux, S., Tucker, K.P., Shevenell, A., Agbandje-McKenna, M., and Breitbart, M. (2014). Diversity of environmental single-stranded DNA phages revealed by PCR amplification of the partial major capsid protein. ISME J. *8*, 2093–2103. 10.1038/ismej.2014.43.

4. Kivelä, H.M., Männistö, R.H., Kalkkinen, N., and Bamford, D.H. (1999). Purification and Protein Composition of PM2, the First Lipid-Containing Bacterial Virus To Be Isolated. Virology *262*, 364–374. 10.1006/viro.1999.9838.

5. Kauffman, K.M., Hussain, F.A., Yang, J., Arevalo, P., Brown, J.M., Chang, W.K., VanInsberghe, D., Elsherbini, J., Sharma, R.S., Cutler, M.B., et al. (2018). A major lineage of non-tailed dsDNA viruses as unrecognized killers of marine bacteria. Nature *554*, 118–122. 10.1038/nature25474.

6. Espejo, R.T., and Canelo, E.S. (1968). Properties of bacteriophage PM2: a lipid-containing bacterial virus. Virology *34*, 738–747. 10.1016/0042-6822(68)90094-9.

7. Kivelä, H.M., Kalkkinen, N., and Bamford, D.H. (2002). Bacteriophage PM2 Has a Protein Capsid Surrounding a Spherical Proteinaceous Lipid Core. J. Virol. *76*, 8169–8178. 10.1128/JVI.76.16.8169-8178.2002.

8. Brum, J.R., Schenck, R.O., and Sullivan, M.B. (2013). Global morphological analysis of marine viruses shows minimal regional variation and dominance of non-tailed viruses. ISME J. *7*, 1738–1751. 10.1038/ismej.2013.67.

9. Yutin, N., Bäckström, D., Ettema, T.J.G., Krupovic, M., and Koonin, E.V. (2018). Vast diversity of prokaryotic virus genomes encoding double jelly-roll major capsid proteins uncovered by genomic and metagenomic sequence analysis. Virol. J. *15*, 67. 10.1186/s12985-018-0974-y.

980 10. Abrescia, N.G.A., Bamford, D.H., Grimes, J.M., and Stuart, D.I. (2012).
981   Structure unifies the viral universe. Annu. Rev. Biochem. *81*, 795–822.
982   10.1146/annurev-biochem-060910-095130.

983 11. Woo, A.C., Gaia, M., Guglielmini, J., Da Cunha, V., and Forterre, P. (2021).
984   Phylogeny of the Varidnaviria Morphogenesis Module: Congruence and
985   Incongruence With the Tree of Life and Viral Taxonomy. Front. Microbiol. *12*,
986   704052. 10.3389/fmicb.2021.704052.

987 12. Bamford, D.H. (2003). Do viruses form lineages across different domains of life?
988   Res. Microbiol. *154*, 231–236. 10.1016/S0923-2508(03)00065-2.

989 13. Azinas, S., Bano, F., Torca, I., Bamford, D.H., Schwartz, G.A., Esnaola, J.,
990   Oksanen, H.M., Richter, R.P., and Abrescia, N.G. (2018). Membrane-containing
991   virus particles exhibit the mechanics of a composite material for genome
992   protection. Nanoscale *10*, 7769–7779. 10.1039/C8NR00196K.

993 14. Walker, P.J., Siddell, S.G., Lefkowitz, E.J., Mushegian, A.R., Adriaenssens, E.M.,
994   Dempsey, D.M., Dutilh, B.E., Harrach, B., Harrison, R.L., Hendrickson, R.C., et
995   al. (2020). Changes to virus taxonomy and the Statutes ratified by the
996   International Committee on Taxonomy of Viruses (2020). Arch. Virol. *165*,
997   2737–2748. 10.1007/s00705-020-04752-x.

998 15. Peralta, B., Gil-Carton, D., Castaño-Díez, D., Bertin, A., Boulogne, C., Oksanen,
999   H.M., Bamford, D.H., and Abrescia, N.G.A. (2013). Mechanism of Membranous
1000   Tunnelling Nanotube Formation in Viral Genome Delivery. PLoS Biol. *11*,
1001   e1001667. 10.1371/journal.pbio.1001667.

1002 16. Krupovič, M., and Bamford, D.H. (2007). Putative prophages related to lytic
1003   tailless marine dsDNA phage PM2 are widespread in the genomes of aquatic
1004   bacteria. BMC Genomics *8*, 236. 10.1186/1471-2164-8-236.

1005 17. Joo, Y.J., Kim, H.-J., Lee, J.Y., and Kim, J. (2004). Biochemical quantitation of
1006   PM2 phage DNA as a substrate for endonuclease assay. J. Microbiol. Seoul
1007   Korea *42*, 99–102.

1008 18. Abrescia, N.G.A., Grimes, J.M., Kivelä, H.M., Assenberg, R., Sutton, G.C.,
1009   Butcher, S.J., Bamford, J.K.H., Bamford, D.H., and Stuart, D.I. (2008). Insights
1010   into Virus Evolution and Membrane Biogenesis from the Structure of the
1011   Marine Lipid-Containing Bacteriophage PM2. Mol. Cell *31*, 749–761.
1012   10.1016/j.molcel.2008.06.026.

1013   19. Laanto, E., Mäntynen, S., De Colibus, L., Marjakangas, J., Gillum, A., Stuart,
1014   D.I., Ravantti, J.J., Huiskonen, J.T., and Sundberg, L.-R. (2017). Virus found in a
1015   boreal lake links ssDNA and dsDNA viruses. Proc. Natl. Acad. Sci. *114*, 8378–
1016   8383. 10.1073/pnas.1703834114.

1017   20. Kejzar, N., Laanto, E., Rissanen, I., Abrishami, V., Selvaraj, M., Moineau, S.,
1018   Ravantti, J., Sundberg, L.-R., and Huiskonen, J.T. (2022). Cryo-EM structure of
1019   ssDNA bacteriophage ΦCjT23 provides insight into early virus evolution. Nat.
1020   Commun. *13*, 7478. 10.1038/s41467-022-35123-6.

1021   21. Kalatzis, P.G., Mauritzen, J.J., Winther-Have, C.S., Michniewski, S., Millard, A.,
1022   Tsertou, M.I., Katharios, P., and Middelboe, M. (2023). Staying below the
1023   Radar: Unraveling a New Family of Ubiquitous "Cryptic" Non-Tailed
1024   Temperate Vibriophages and Implications for Their Bacterial Hosts. Int. J. Mol.
1025   Sci. *24*, 3937. 10.3390/ijms24043937.

1026   22. Wall, J.D., Weaver, P.F., and Gest, H. (1975). Gene transfer agents,
1027   bacteriophages, and bacteriocins of Rhodopseudomonas capsulata. Arch.
1028   Microbiol. *105*, 217–224. 10.1007/BF00447140.

1029   23. Strnad, H., Lapidus, A., Paces, J., Ulbrich, P., Vlcek, C., Paces, V., and
1030   Haselkorn, R. (2010). Complete genome sequence of the photosynthetic purple
1031   nonsulfur bacterium Rhodobacter capsulatus SB 1003. J. Bacteriol. *192*, 3545–
1032   3546. 10.1128/JB.00366-10.

1033   24. Ding, H., Moksa, M.M., Hirst, M., and Beatty, J.T. (2014). Draft Genome
1034   Sequences of Six Rhodobacter capsulatus Strains, YW1, YW2, B6, Y262, R121,
1035   and DE442. Genome Announc. *2*, e00050-14. 10.1128/genomeA.00050-14.

1036   25. Weaver, P.F., Wall, J.D., and Gest, H. (1975). Characterization of
1037   Rhodopseudomonas capsulata. Arch. Microbiol. *105*, 207–216.
1038   10.1007/BF00447139.

1039   26. Leidenfrost, R.M., Wappler, N., and Wünschiers, R. (2020). Draft Genome
1040   Assembly of Rhodobacter sphaeroides 2.4.1 Substrain H2 from Nanopore Data.
1041   Microbiol. Resour. Announc. *9*, e00414-20. 10.1128/MRA.00414-20.

1042   27. Duyvesteyn, H.M.E., Ginn, H.M., Pietilä, M.K., Wagner, A., Hattne, J., Grimes,
1043   J.M., Hirvonen, E., Evans, G., Parsy, M.-L., Sauter, N.K., et al. (2018). Towards
1044   in cellulo virus crystallography. Sci. Rep. *8*, 3771. 10.1038/s41598-018-21693-3.

1045   28. Kivelä, H.M., Daugelavičius, R., Hankkio, R.H., Bamford, J.K.H., and Bamford,
1046   D.H. (2004). Penetration of Membrane-Containing Double-Stranded-DNA

1047   Bacteriophage PM2 into Pseudoalteromonas Hosts. J. Bacteriol. *186*, 5342–5354.
1048   10.1128/JB.186.16.5342-5354.2004.

1049 29. Gudlavalleti, B.S., Phung, T., Barton, C.L., Becker, A., Graul, B.L., Griffin, J.T.,
1050   Hays, C.J., Horn, B., Liang, D.R., Rutledge, L.M., et al. (2020). Whole genome
1051   sequencing identifies an allele responsible for clear vs. turbid plaque
1052   morphology in a Mycobacteriophage. BMC Microbiol. *20*, 148. 10.1186/s12866-
1053   020-01833-4.

1054 30. Bondy-Denomy, J., Qian, J., Westra, E.R., Buckling, A., Guttman, D.S.,
1055   Davidson, A.R., and Maxwell, K.L. (2016). Prophages mediate defense against
1056   phage infection through diverse mechanisms. ISME J. *10*, 2854–2866.
1057   10.1038/ismej.2016.79.

1058 31. Carnoy, C., and Roten, C.-A. (2009). The dif/Xer Recombination Systems in
1059   Proteobacteria. PLoS ONE *4*, e6531. 10.1371/journal.pone.0006531.

1060 32. Huber, K.E., and Waldor, M.K. (2002). Filamentous phage integration requires
1061   the host recombinases XerC and XerD. Nature *417*, 656–659.
1062   10.1038/nature00782.

1063 33. Brimacombe, C.A., Ding, H., and Beatty, J.T. (2014). Rhodobacter capsulatus
1064   DprA is essential for RecA-mediated gene transfer agent (RcGTA) recipient
1065   capability regulated by quorum-sensing and the CtrA response regulator. Mol.
1066   Microbiol. *92*, 1260–1278. 10.1111/mmi.12628.

1067 34. Fogg, P.C.M. (2019). Identification and characterization of a direct activator of
1068   a gene transfer agent. Nat. Commun. *10*, 595. 10.1038/s41467-019-08526-1.

1069 35. Mascolo, E., Adhikari, S., Caruso, S.M., deCarvalho, T., Folch Salvador, A.,
1070   Serra-Sagristà, J., Young, R., Erill, I., and Curtis, P.D. (2022). The
1071   transcriptional regulator CtrA controls gene expression in Alphaproteobacteria
1072   phages: Evidence for a lytic deferment pathway. Front. Microbiol. *13*, 918015.
1073   10.3389/fmicb.2022.918015.

1074 36. Kropinski, A.M. (2009). Measurement of the rate of attachment of
1075   bacteriophage to cells. Methods. Mol. Biol. *501*, 151-5. 10.1007/978-1-60327-
1076   164-6_15.

1077 37. Storms, Z.J., Arsenault, E., Sauvageau, D., and Cooper, D.G. (2010).
1078   Bacteriophage adsorption efficiency and its effect on amplification. Bioprocess
1079   Biosyst. Eng. *33*, 823–831. 10.1007/s00449-009-0405-y.

1080 38. Bamford, D.H., Rouhiainen, L., Takkinen, K., and Söderlund, H. 1981
1081   Comparison of the Lipid-containing Bacteriophages PRD1, PR3, PR4, PR5 and
1082   L17. J. Gen. Virol. *57*, 365–373. 10.1099/0022-1317-57-2-365.

1083 39. Zhang, Y., and Jiao, N. (2009). Roseophage RDJLΦ1, Infecting the Aerobic
1084   Anoxygenic Phototrophic Bacterium Roseobacter denitrificans OCh114. Appl.
1085   Environ. Microbiol. *75*, 1745–1749. 10.1128/AEM.02131-08.

1086 40. Meier-Kolthoff, J.P., and Göker, M. (2017). VICTOR: genome-based phylogeny
1087   and classification of prokaryotic viruses. Bioinformatics *33*, 3396–3404.
1088   10.1093/bioinformatics/btx440.

1089 41. Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M.,
1090   Gabler, F., Söding, J., Lupas, A.N., and Alva, V. (2018). A Completely
1091   Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its
1092   Core. J. Mol. Biol. *430*, 2237–2243. 10.1016/j.jmb.2017.12.007.

1093 42. Yutin, N., Rayko, M., Antipov, D., Mutz, P., Wolf, Y.I., Krupovic, M., and
1094   Koonin, E.V. (2022). Varidnaviruses in the Human Gut: A Major Expansion of
1095   the Order Vinavirales. Viruses *14*, 1842. 10.3390/v14091842.

1096 43. Hyman, P. (2019). Phages for Phage Therapy: Isolation, Characterization, and
1097   Host Range Breadth. Pharmaceuticals *12*, 35. 10.3390/ph12010035.

1098 44. Poxleitner M., Pope W., Jacobs-Sera D., Sivanathan V., and Hatfull G. (2018).
1099   Chapter 5: Isolation. In Sea phage discovery guide. Howard Hughes Medical
1100   Institute.

1101 45. Sharma, S., Datta, S., Chatterjee, S., Dutta, M., Samanta, J., Vairale, M.G.,
1102   Gupta, R., Veer, V., and Dwivedi, S.K. (2021). Isolation and characterization of
1103   a lytic bacteriophage against Pseudomonas aeruginosa. Sci. Rep. *11*, 19393.
1104   10.1038/s41598-021-98457-z.

1105 46. Henry, M., Lavigne, R., and Debarbieux, L. (2013). Predicting In Vivo Efficacy
1106   of Therapeutic Bacteriophages Used To Treat Pulmonary Infections.
1107   Antimicrob. Agents Chemother. *57*, 5961–5968. 10.1128/AAC.01596-13.

1108 47. Brimacombe, C.A., Stevens, A., Jun, D., Mercer, R., Lang, A.S., and Beatty, J.T.
1109   (2013). Quorum-sensing regulation of a capsular polysaccharide receptor for the
1110   Rhodobacter capsulatus gene transfer agent (RcGTA). Mol. Microbiol. *87*, 802–
1111   817. 10.1111/mmi.12132.

48. Sherlock, D., and Fogg, P.C.M. (2022). Loss of the Rhodobacter capsulatus Serine Acetyl Transferase Gene, cysE1, Impairs Gene Transfer by Gene Transfer Agents and Biofilm Phenotypes. Appl. Environ. Microbiol. *88*, e00944-22. 10.1128/aem.00944-22.

49. Bazinet, C.W., and King, J. (1985). A late gene product of phage P22 affecting virus infectivity. Virology *143*, 368–379. 10.1016/0042-6822(85)90377-0.

50. Westbye, A.B., Leung, M.M., Florizone, S.M., Taylor, T.A., Johnson, J.A., Fogg, P.C., and Beatty, J.T. (2013). Phosphate concentration and the putative sensor kinase protein CckA modulate cell lysis and release of the Rhodobacter capsulatus gene transfer agent. J. Bacteriol. *195*, 5025–5040. 10.1128/JB.00669-13.

51. Männistö, R.H., Grahn, A.M., Bamford, D.H., and Bamford, J.K.H. (2003). Transcription of Bacteriophage PM2 Involves Phage-Encoded Regulators of Heterologous Origin. J. Bacteriol. *185*, 3278–3287. 10.1128/JB.185.11.3278-3287.2003.

52. Benešík, M., Nováček, J., Janda, L., Dopitová, R., Pernisová, M., Melková, K., Tišáková, L., Doškař, J., Žídek, L., Hejátko, J., et al. (2018). Role of SH3b binding domain in a natural deletion mutant of Kayvirus endolysin LysF1 with a broad range of lytic activity. Virus Genes *54*, 130–139. 10.1007/s11262-017-1507-2.

53. Rodríguez-Rubio, L., Martínez, B., Donovan, D.M., Rodríguez, A., and García, P. (2013). Bacteriophage virion-associated peptidoglycan hydrolases: potential new enzybiotics. Crit. Rev. Microbiol. *39*, 427–434. 10.3109/1040841X.2012.723675.

54. Rydman, P.S., and Bamford, D.H. (2002). The Lytic Enzyme of Bacteriophage PRD1 Is Associated with the Viral Membrane. J. Bacteriol. *184*, 104–110. 10.1128/JB.184.1.104-110.2002.

55. Fogg, M.J., and Wilkinson, A.J. (2008). Higher-throughput approaches to crystallization and crystal structure determination. Biochem. Soc. Trans. *36*, 771–775. 10.1042/BST0360771.

56. Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics *30*, 2068–2069. 10.1093/bioinformatics/btu153.

57. Besemer, J., Lomsadze, A., and Borodovsky, M. (2001). GeneMarkS: a self-training method for prediction of gene starts in microbial genomes.

1146 Implications for finding sequence motifs in regulatory regions. Nucleic Acids
1147 Res. *29*, 2607–2618. 10.1093/nar/29.12.2607.

58. Carver, T., Harris, S.R., Berriman, M., Parkhill, J., and McQuillan, J.A. (2012).
    Artemis: an integrated platform for visualization and analysis of high-
    throughput sequence-based experimental data. Bioinformatics *28*, 464–469.
    10.1093/bioinformatics/btr703.

59. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W.,
    and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of
    protein database search programs. Nucleic Acids Res. *25*, 3389–3402.
    10.1093/nar/25.17.3389.

60. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O.,
    Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly
    accurate protein structure prediction with AlphaFold. Nature *596*, 583–589.
    10.1038/s41586-021-03819-2.

61. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and
    Steinegger, M. (2022). ColabFold: making protein folding accessible to all. Nat.
    Methods *19*, 679–682. 10.1038/s41592-022-01488-1.

62. Holm, L. (2020). DALI and the persistence of protein shape. Protein Sci. Publ.
    Protein Soc. *29*, 128–140. 10.1002/pro.3749.

63. Kans, J. (2022). Entrez Direct: E-utilities on the Unix Command Line. In Entrez
    Programming Utilities Help. NCBI, 85-186.

64. Katoh, K., and Standley, D.M. (2013). MAFFT Multiple Sequence Alignment
    Software Version 7: Improvements in Performance and Usability. Mol. Biol.
    Evol. *30*, 772–780. 10.1093/molbev/mst010.

65. Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a
    tool for automated alignment trimming in large-scale phylogenetic analyses.
    Bioinformatics *25*, 1972–1973. 10.1093/bioinformatics/btp348.

66. Minh, B.Q., Nguyen, M.A.T., and von Haeseler, A. (2013). Ultrafast
    Approximation for Phylogenetic Bootstrap. Mol. Biol. Evol. *30*, 1188–1195.
    10.1093/molbev/mst024.

67. Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996). Computer
    Visualization of Three-Dimensional Image Data Using IMOD. J. Struct. Biol.
    116, 71–76. 10.1006/jsbi.1996.0013.

1179 68. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris,
1180   J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in
1181   visualization and analysis. Protein Sci. Publ. Protein Soc. *27*, 14–25.
1182   10.1002/pro.3235.

1183 69. Wickham H. (2016). ggplot2, Elegant Graphics for Data Analysis. Springer.

1184 70. Lee I., Kim Y. O., Park S. C., and Chun J. (2016). OrthoANI: an improved
1185   algorithm and software for calculating average nucleotide identity. Int. J. Syst.
1186   Evol. Microbiol. *66*, 1100–1103. 10.1099/ijsem.0.000760.

1187 71. Ormerod, J.G., Ormerod, K.S., and Gest, H. (1961). Light-dependent utilization
1188   of organic compounds and photoproduction of molecular hydrogen by
1189   photosynthetic bacteria; relationships with nitrogen metabolism. Arch.
1190   Biochem. Biophys. *94*, 449–463. 10.1016/0003-9861(61)90073-x.

1191 72. Kropinsky A. M. (2018). Practical Advice on the One-Step Growth Curve.
1192   Methods. Mol. Biol. *1681*, 41-47. 10.1007/978-1-4939-7343-9_3

1193 73. Sambrook, J., and Russell, D. W. (2001). Commonly Used Techniques in
1194   Molecular Cloning. In Molecular cloning. Cold Spring Harbor Laboratory Press.

1195 74. Hynes, A., and Lang, A. (2013). Rhodobacter capsulatus Gene Transfer Agent
1196   (RcGTA) Activity Bioassays. BIO-Protoc. *3*, e317. 10.21769/BioProtoc.317.

1197 75. Grybchuk, D., Macedo, D.H., Kleschenko, Y., Kraeva, N., Lukashev, A.N.,
1198   Bates, P.A., Kulich, P., Leštinová, T., Volf, P., Kostygov, A.Y., et al. (2020). The
1199   First Non-LRV RNA Virus in Leishmania. Viruses *12*, 168. 10.3390/v12020168.

1200 76. Fogg, P.C.M., Westbye, A.B., and Beatty, J.T. (2012). One for All or All for One:
1201   Heterogeneous Expression and Host Cell Lysis Are Key to Gene Transfer Agent
1202   Activity in Rhodobacter capsulatus. PLOS ONE *7*, e43772.
1203   10.1371/journal.pone.0043772.

1204