

1 **Phage inducible chromosomal minimalist island (PICMI), a family of**
2 **satellites of marine virulent phages**

3 Rubén Barcia-Cruz^{1,2}, David Goudenège^{1,3}, Jorge A. Moura de Sousa⁴, Damien Piel^{1,3}, Martial
4 Marbouty⁵, Eduardo P.C. Rocha⁴ and Frédérique Le Roux^{1,3#}

5 ¹Sorbonne Université, CNRS, UMR 8227, Integrative Biology of Marine Models, Station
6 Biologique de Roscoff, CS 90074, F-29688 Roscoff cedex, France.

7 ²Department of Microbiology and Parasitology, CIBUS-Faculty of Biology, Universidade de
8 Santiago de Compostela, Santiago de Compostela, Spain.

9 ³Ifremer, Unité Physiologie Fonctionnelle des Organismes Marins, ZI de la Pointe du Diable,
10 CS 10070, F-29280 Plouzané, France.

11 ⁴Institut Pasteur, Université Paris Cité, CNRS UMR3525, Microbial Evolutionary Genomics,
12 Paris, France.

13 ⁵Institut Pasteur, Université Paris Cité, Organization and Dynamics of Viral Genomes Group,
14 CNRS UMR 3525, Paris F-75015 France.

15 **#Corresponding author** : Frédérique Le Roux

16 Equipe Génomique des Vibrios, UMR 8227, Integrative Biology of Marine Models, Station
17 Biologique de Roscoff, CS 90074, F-29688, Roscoff cedex, France.

18 Tel: +33 2 98 29 56 47, frederique.le-roux@sb-roscoff.fr

19 **Keywords:** vibrio, marine viral particle, mobile genetic element, immune system

20

21 **SUMMARY**

22 Phage satellites are genetic elements that hijack the phage machinery for their own
23 dissemination. However, only few phage satellites have been characterized, and mechanisms
24 by which they influence microbial evolution in nature are unclear. Here we identify a new
25 family of satellites, the Phage Inducible Chromosomal Minimalist Island (PICMI), which is
26 broadly distributed in the marine bacteria *Vibrionaceae*. PICMI is characterized by reduced
27 gene content, does not encode genes for capsid remodeling and packages its DNA as a
28 concatemer. PICMI is integrated in the bacterial host genome at the end of the *fis* regulator and
29 encodes three core proteins necessary for excision and replication. PICMI is dependent on
30 virulent phage particles to spread to other bacteria and confers host protection from other
31 competitive phages, without interfering with its helper phage. The discovery of PICMI strongly
32 suggests that phages, including virulent ones, play important roles for mobility of phage defense
33 elements.

34

35 INTRODUCTION

36 Bacteriophages (or phages) are viruses that infect bacteria and may be the most diverse and
37 abundant biological entities in the ocean^{1,2}. Since most phages kill their hosts because of their
38 life cycles, phages are key players for promoting bacterial abundance and diversity. Phages are
39 themselves exploited by phage satellites, a class of mobile genetic elements that hijack the
40 phage machinery to promote their own dissemination while interfering with phage
41 reproduction³⁻⁶. Recently, several studies revealed that marine bacteria transduce a vast
42 diversity of chromosomal islands, many of which might be satellites. These elements were
43 named Virion Encapsidated Integrative Mobile Element (VEIME)⁷ and Tycheposons⁸. These
44 studies suggest that satellites play important roles in natural environments, however their
45 function and characteristics are still poorly understood.

46 Satellites may develop different strategies for hijacking the life cycle of a helper phage, and the
47 characterization of new satellites can reveal features that unite or separate different satellite
48 families⁹. Much of our knowledge about the lifestyle of phage satellites results from a limited
49 number of elements discovered in clinically relevant bacteria, such as P4-like satellites in
50 *Enterobacterales*¹⁰, the phage-inducible chromosomal islands (PICIs) in *Bacillales* and
51 *Gammaproteobacteria*^{3,11}, its closely related family capsid forming PICIs (cf-PICI) in
52 Proteobacteria and Firmicutes^{12,13}, and the phage-inducible chromosomal islands-like elements
53 (PLEs) in *Vibrio cholerae*^{14,15}. A common feature of all known satellites is their integration in
54 a specific site of the bacterial genome, where satellites remain until their excision is promoted
55 by the induction of a helper prophage, infection by a helper temperate phage, or infection by
56 the virulent helper phage ICP1. This process requires integrases and excision factors encoded
57 by the satellites⁶. The circularized extrachromosomal element then replicates extensively using
58 its own origin of replication and is packaged into viral particles using *pac* and/or *cos* packaging
59 systems¹⁶. Known satellites typically have genomes about one-third of the size of the helper
60 phages. Many satellites encode diverse mechanisms of capsid remodeling in order to fit their
61 smaller genomes, whilst excluding the larger genomes of their helper phages^{6,10,15,17}. Specific
62 features of each phage satellite family include the lifestyle of their helper phage, i.e. temperate
63 for PICI, cf-PICI and P4-like satellites vs. virulent for the helper phage of PLE named ICP1,
64 their genome size of an average 10, 9.5, 14, 18 kbp for P4-like satellites, PICI, cf-PICI and
65 PLE, gene repertoires and the mechanisms they use to subvert the host phage particles. The
66 exploration of known satellites raises the question of the minimal gene set required to excise
67 them from the genome, replicate as an extra-chromosomal element, and hijack the helper phage.

68 Whether the use of its machinery has a cost on helper phage reproduction is also crucial to
69 understand the interaction between the bacterial host, the phage, and the satellite.

70 The effect of satellites on phage reproduction varies across families, and sometimes even across
71 satellites of the same family. While P4-like satellites and PICIs only interfere partially with the
72 reproduction of their helper phages^{18,19}, PLEs completely abrogate the production of ICP1
73 progeny¹⁵. Finally, some P4-like satellites²⁰ and PICI²¹ encode hotspots of antiviral systems
74 protecting both the bacterial host and their helper phages from competing phages and other
75 mobile genetic elements. The associations of known phage satellites thus range from pure
76 parasitism to mutualism in relation to their bacterial and phage hosts. While it has become clear
77 that phage satellites are abundant and play diverse roles that affect their hosts, only a limited
78 number have been described, hindering our ability to fully understand the breadth of their
79 influence.

80 Given the abundance, diversity, and distribution of viruses in the ocean, identifying new phage
81 satellites and understanding their functions might be akin to “looking for a needle in a
82 haystack”. Characterizing and establishing the functions of new phage satellites requires
83 identification of the cognate helper phages and the cellular hosts to understand the parasitic life
84 cycle. In the present study, we addressed this challenge by taking advantage of bacteria from
85 the *Vibrionaceae* family, which is unique for its extensive culture and sequence coverage of
86 hosts and phages^{22,23}. The *Vibrionaceae* (vibrios) comprise a diverse group of bacteria that are
87 widespread within marine environments, encompassing human and animal pathogens^{24,25}.
88 Vibrios are easily cultured, allowing isolation of their infective phages, whole genome
89 sequencing, and inverse genetics.

90 We report the discovery of a new family of satellites that hijack virulent phages. We named this
91 family of satellites PICMI (Phage Inducible Chromosomal Minimalist Island) because of their
92 reduced size and gene content. When sequencing the genome of a virulent phage and its vibrio
93 host, we detected concatemeric repeat sequences of PICMI in virus particles. PICMI encodes
94 three core proteins necessary for excision and replication but is completely dependent on its
95 helper phage to generate the PICMI infective particles and mobilize across bacteria. Despite
96 this high dependency, PICMI does not strongly interfere with the fitness of its helper phage.
97 However, the satellite can confer host protection from other phages. PICMIs are broadly
98 distributed in the *Vibrionaceae* which encompasses the potential pathogens *V. cholerae*, *V.*

99 *parahaemolyticus* and *V. vulnificus*²⁴, suggesting an important role of this marine phage satellite
100 in *Vibrionaceae*.

101 RESULTS

102 Identification of a satellite, its cognate host, and helper phage from marine viral particles

103 We previously isolated and sequenced 49 phages infecting strains of *V. chagasii*²². One virulent
104 phage (115_E_34-1, named Φ 115 for simplicity) piqued our curiosity because its genome
105 sequence assembly revealed two contigs of 47,851 and 6,110 bp (Fig. S1). The number of
106 sequencing reads was 634,740 and 108,219 for the large and small contig respectively, with
107 only 16 hybrid reads, providing evidence for two distinct mobile genetic elements in the Φ 115
108 phage progeny. The large element corresponds to the genome of the Φ 115 phage. The small
109 element contains six genes encoding an integrase (*int*), a putative regulator (*alpA*), a putative
110 primase (*prim*) and three other genes of unknown function (Fig. 1A). Among the genes of
111 unknown function, one gene has four nucleotides overlap (ATGA) with the sequence of *int* and
112 was named *IOLG* for Integrase Over Lapping Gene. The two other genes were named *UP1* and
113 *UP2* for Unknown Protein 1 and 2. Using Phanotate, a tool dedicated to phage genome
114 annotation (see Methods), we identified six additional open reading frames (ORFs), (Fig. S2)
115 which were considered highly questionable because they encoded for proteins of 32 to 44 amino
116 acids. These doubtful ORFs and/or pseudogenes explain the existence of a large non-coding
117 region between *fis* and *UP2*. The 6,110 bp element was also found in the genome of the host
118 used to isolate the phage, *V. chagasii* 34_P_115 (herein named V115), integrated at the end of
119 the *fis* regulator gene and flanked by two direct repeats of 17 bp (Fig. S1). We thus assumed
120 that the 6,110 bp element is a phage satellite and the phage Φ 115 is its helper phage.

121 Despite the small size of the satellite (~1/8 the phage genome), we did not observe smaller-
122 sized capsids commensurate to its genome size, as described for other known families of phage
123 satellites (Fig. S3). We tested for the presence of physical contacts between the two genomes²⁶
124 to confirm that the satellite was located in viral particles lacking the phage genome. We applied
125 HiC²⁷ on different mixes of phage particles (see Methods). The result showed a clear absence
126 of physical contact between the DNAs of the phage Φ 115 and the satellite (Fig. S4), confirming
127 the exclusive packaging of the satellite in full-sized phage-like particles. To understand if the
128 satellite fills the capsid by packaging as a concatemer, we performed single-molecule nanopore
129 sequencing of Φ 115- encapsidated high-molecular weight DNA and found a fraction of the

130 viral particles contained a concatemer of 8 copies of the satellite of ~49 kbp size, similar to the
131 Φ 115 genome (Fig. S5, Table S1). Finally, we confirmed by Southern blot that concatemeric
132 repeat sequences of the satellite show DNA of similar size to that of the genome of phage Φ 115
133 packaged in viral particles (Fig. 1B). To estimate the percentage of Φ 115 particles that contain
134 the satellites instead of phage DNA, we first normalized the Illumina sequencing reads on
135 genome size ($634,740/47,851=13.26$ for the phage and $108,219/6,110=17.71$ for the satellite)
136 and next considered that eight copies of the satellite are packaged as concatemer in particles
137 ($17.71/8=2.21$). This led to 16% ($2.21/13.26*100$) of the population being hitchhiked by the
138 satellite. This estimation was further confirmed by nanopore sequencing (10 or 15% depending
139 on the replicate, Table S1) and qPCR of phage DNA (15%, Fig. S6). Altogether, these data
140 strongly suggest the identification of a new marine phage satellite. Due to its reduced size, we
141 named this satellite PICMI₁₁₅, for Phage Inducible Chromosomal Minimalist Island identified
142 in vibrio V115 and its cognate helper phage Φ 115.

143 **PICMI sustains minimal function of excision and replication**

144 Our cultivation-enabled model system has enabled us to dissect the various steps in the PICMI's
145 life cycle: (i) excision, (ii) replication, (iii) packaging (iv) transduction to a new host. Of these,
146 we were not able to identify *cos* or *pac* packaging sites in the helper phage genome, or any
147 homologs of genes involved in redirecting packaging that are characteristic of other satellite
148 families (i.e., *terS*, *sid*, *ppi*). However, Nanopore sequencing of the viral particles revealed
149 random extremities of the PICMI₁₁₅ concatemer (Fig. S5), which is indicative of a mechanism
150 of headful (*pac*-like) packaging.

151 Prior to exploring the induction of PICMI₁₁₅ by the helper phage, we generated Φ 115 viral
152 particles without the PICMI₁₁₅, by two passages in a V115 derivative lacking the entire
153 PICMI₁₁₅. The vibrio mutant was named Δ PICMI₁₁₅, and the phage progeny was named
154 Φ 115pure. The absence of the satellite in the Φ 115pure population was confirmed by nanopore
155 sequencing (Table S1), qPCR (Fig. S6), and Southern blot (Fig. 1B). To test for helper phage
156 dependent excision/circularization activation, we performed qPCR analysis with inward- and
157 outward-directed primers as shown in Fig.1A. Amplicons were obtained 15 minutes (min) after
158 adding the phage to the bacterial culture (Fig. 1C), the estimated time for its complete
159 adsorption and phage DNA injection in the host cytoplasm (Fig. S7). Excision/circularization
160 was observed exclusively in the presence of the helper phage (Fig. S8). Outward-directed

161 primers also amplify the junctions in the concatemer. A dramatic increase of the copy number
162 of the circular DNA was observed 30 min after phage addition (Fig. 1C), indicating intensive
163 replication. The increase in the amount of a DNA band at the size of the concatemer is observed
164 by Southern blot at 60 min (Fig. 1D), consistent with a plasmid-like rolling circle DNA
165 replication mechanism.

166 It is expected that PICMI₁₁₅ transduction requires the helper phage to adsorb on the recipient
167 host. The phage Φ 115 was previously described as having a narrow host range²², with only 2
168 out of 136 *V. chagasii* strains (V115 and V157) susceptible to Φ 115 infection and reproduction.
169 These two strains each encode one (identical) PICMI. Testing closer phylogenetic neighbors to
170 the original host V115, we found that the phage Φ 115 adsorbs to the strain V511 without
171 producing progeny (Fig. S9), probably due to intracellular defense mechanisms^{23,28}. V511 does
172 not carry PICMI and shows 100% identity with the *fis* gene of V115 and V157. We thus
173 assumed that the vibrio strain V511, in addition to the V115 derivative lacking PICMI₁₁₅
174 (Δ PICMI₁₁₅), could be used as recipient for transduction assays. We first inserted a
175 chloramphenicol resistance marker (Cm^R) downstream of the *prim* gene of PICMI₁₁₅ (Fig. 1A)
176 and infected this strain with Φ 115pure to produce lysates of viral particles with the Cm^R-
177 encoding PICMI₁₁₅. The introduction of the Cm^R cassette slightly increased the copies number
178 of PICMI₁₁₅ in phage particles (Fig. S6). We thus assumed that the excision, replication, and
179 packaging functions of the PICMI₁₁₅-Cm^R satellite were intact. We next used this lysate to
180 infect the two susceptible hosts, Δ PICMI₁₁₅ and V511, and selected for chloramphenicol
181 resistant cells that acquired the PICMI₁₁₅-Cm^R satellite. Transductants were obtained at a
182 multiplicity of infection (MOI) from 0.1 to 0.0001 for the recipient Δ PICMI₁₁₅, and MOI from
183 10 to 0.01 for V511. PICMI₁₁₅-Cm^R was transduced at higher frequency when Δ PICMI₁₁₅ was
184 the recipient (transductants in CFU.ml⁻¹/ phages in PFU.ml⁻¹ ~ 6.10⁻³) relative to V511 (10⁻⁶)
185 (Fig. 1E, upper panel). We confirmed by PCR that the integration of PICMI₁₁₅-Cm^R occurred
186 at the end of the *fis* gene in all tested transductants (Fig. 1E, lower panel). Altogether our results
187 showed that PICMI₁₁₅ is activated by a virulent phage, replicated by rolling circle, packaged,
188 and transduced as a concatemer. With roughly 15% of viral particles that contain the satellite,
189 the question arises whether PICMI₁₁₅ inflicts a cost on its helper phage.

190 Most known satellites interfere at least partially with the reproduction of their helper phage,
191 although this effect can vary broadly between families^{15,18,19}. We thus quantified the extent to
192 which PICMI₁₁₅ interferes with its helper phage Φ 115. We compared the titer of phages

193 produced by the bacterial strains V115 wild type (wt), Δ PICMI₁₁₅, and two clones of Δ PICMI₁₁₅
194 +PICMI₁₁₅-Cm^R after infection by Φ 115pure. No significant differences (ANOVA, P=0.75, F-
195 test) were observed between the strains (Fig. 1F), showing that PICMI₁₁₅ does not strongly
196 impact the reproductive fitness of its helper phage.

197 **AlpA is a key regulator of PICMI activation**

198 Having established that PICMI₁₁₅ is a phage satellite, we further analyzed the role of each
199 PICMI₁₁₅-encoded gene in the various aspects of the satellite's lifestyle. Each of the six genes
200 (Fig. 1A) was deleted in V115, and the mutants were compared to the wild-type host for
201 excision, packaging, and transduction. This revealed that *alpA*, *int*, and *IOLG* are necessary for
202 PICMI₁₁₅ excision induced by the helper phage (Fig. 2A; Fig. S10). The deletion of *prim* results
203 in a lower number of circularized PICMI₁₁₅ copies, relative to the wild type, indicating that the
204 primase is involved in PICMI₁₁₅ replication (Fig. 2A; Fig. S10). Accordingly, the number of
205 viral particles that contain PICMI₁₁₅ in phage progenies was strongly reduced in Δ *alpA*, Δ *int*,
206 Δ *IOLG*, and Δ *prim* deletion mutants (Fig. 2B, Fig. S6). It also led to a much lower transduction
207 of PICMI₁₁₅-Cm^R, below our limit of detection (Table 1). When expressed in trans in the V115
208 derivative Δ *int*, *int*, with or without the overlapping gene *IOLG*, restored the helper phage-
209 induced excision of PICMI₁₁₅ (Fig. 2C). The expression of *int/IOLG* also complemented the
210 Δ *IOLG* deletion (Fig. 2C). Notably, the expression of *alpA* in trans was sufficient to induce
211 PICMI₁₁₅ excision even in the absence of phage (Fig.2C), and the copy number of circularized
212 satellites slightly increased 30 min post infection (Fig. S11).

213 AlpA is predicted to act as a DNA-binding regulator, and previous work suggested that it is a
214 transcriptional regulator³ and/or an excisionase²⁹. In PICMI₁₁₅, the expression of *alpA* from a
215 plasmid did not alter the expression of the satellites' genes, suggesting that *alpA* induction of
216 PICMI₁₁₅ excision is not mediated by transcriptional regulation (Fig. S12A, B). We used
217 ColabFold³⁰ to search for structural similarities with known excisionases and found strong
218 similarities with the TorI regulator in *E. coli* and Xis excisionase from *Streptomyces*
219 *ambofaciens* (Fig. S12C). We conclude that among the three genes essential for excision, *int*
220 and *IOLG* are constitutively expressed, *alpA* is activated by phage, and the three proteins are
221 involved in the formation of the excision complex.

222

223 **The intensity of PICMI₁₁₅ activation depends on the phage used as helper**

224 The induction of the satellite can be highly specific to the helper phage(s)⁵. We searched for
225 other phages that could infect the strain V115 to establish whether induction of PICMI₁₁₅ is also
226 specific to the helper phages Φ 115. We used viruses from seawater collected at the same oyster
227 farm four years later (see Methods) and isolated seven phages infecting the host V115. The
228 probe directed against phage Φ 115 also detected the newly isolated phages by Southern blot,
229 suggesting that they are genetically related (Fig. 3A). Except for Φ 27, PICMI₁₁₅ was detected
230 in the progenies resulting from the infection of all other phages, although at a much lower
231 quantity than in the Φ 115pure infection (Fig. 3A). The phage Φ 27 was unable to induce a
232 detectable excision of PICMI₁₁₅ (Fig. 3B). We thus compared the expression of the PICMI
233 genes upon Φ 115pure (Fig. 3C) and Φ 27 infection (Fig. 3D). This revealed that 15 min post
234 infection by Φ 115pure, presumably as soon as the phage is injected in the cytoplasm (Fig. S7),
235 only *alpA*, *UPI* and *prim* were upregulated, and thus defined as an early regulon. An increase
236 of transcripts for the remaining PICMI₁₁₅ genes (*int*, *IOLG* and *UP2*) was observed after 60
237 min. This resulted from the increase of the number of copies of the PICMI₁₁₅ genome and not
238 necessarily through gene activation. Infection by phage Φ 27, which did not lead to PICMI₁₁₅
239 induction, has no reproducible effect on PICMI₁₁₅ gene expression (Fig. 3D). Altogether our
240 results demonstrate that Φ 115 spurs strong induction of PICMI₁₁₅ and that the ability to induce
241 PICMI is related to the ability to efficiently induce the early PICMI₁₁₅ regulon.

242 **PICMI-like elements are broadly distributed in the *Vibrionaceae***

243 Our description of the PICMI minimalism was based on a single model system identified in our
244 collection, raising questions about the size, distribution, and diversity of PICMI-like elements
245 in bacterial genomes. We thus built a MacSyFinder model³¹ to allow the automatic
246 identification of this element in bacterial genomes. Our PICMI₁₁₅ prototype is integrated in the
247 V115 host genome at the end of the *fis* regulator gene and contains only six genes. Among
248 those, we showed that *int*, *alpA*, and *prim* are necessary for PICMI₁₁₅ lifestyle. In accordance
249 with the experimental data, we set the presence and colocalization of *int*, *alpA*, *prim*, and *fis*
250 genes as mandatory in the model. We then used it to search for PICMI-like elements in all
251 Genbank bacterial complete genomes (v243, 05/26/2021) and identified 135 elements (Table
252 S2). From this list, the 67 satellites in *Vibrionaceae* genomes have a significantly smaller size
253 (average 6.7 kbp, unpaired t test P<0.0001) (Fig. S13). We thus extended our search for such

254 elements to a much larger albeit specific dataset of 19185 *Vibrionaceae* genomes (NCBI
255 Assembly database, 02/16/2023). We identified a total of 97 elements, broadly distributed in
256 diverse *Vibrionaceae* species (Fig. 4, Table S3). We never detected more than one PICMI-like
257 element in these genomes, contrasting with P4-like satellites in *E. coli* genomes, that can
258 contain up to three P4-like satellites¹⁰. Pairwise alignment of the DNA sequences of PICMI-
259 like satellites permitted grouping them into 35 distinct PICMI subfamilies (>90% global
260 pairwise nucleotide identity) (Fig. S14, Table S3). Up to seven subfamilies could be detected
261 in one species (*V. cholerae*) (Fig. S14). In most cases, the distribution of PICMI subfamilies
262 coincides with the host species phylogeny. This is expected for mobile genetic elements
263 transduced by vibriophages that, for the vast majority, have a narrow host range^{22,23,32}.
264 However, PICMI₁₁₅ was detected in two strains of *V. chagasii* (V115 and V157) isolated during
265 the same time series sampling in France²² and in a *V. toranzoniae* strain isolated from cultured
266 clams in NW Spain (cmf 13.9) (Fig. 4 and Fig. S15). Another *V. toranzoniae* isolate from
267 seawater in SE Spain (96-373) carries a different PICMI subfamily. This incongruence between
268 the subfamilies of PICMI and the bacterial hosts suggests that this satellite might be horizontally
269 transferred between diverse *Vibrio* species.

270 We then analyzed the genes that are frequent across the PICMI variants. As for PICMI₁₁₅, large
271 non-coding regions might be explained by the presence of false ORF and/or pseudogenes (Fig.
272 S16). Since these genes are not predicted to be functional, they were not considered in our
273 analysis. The core genes encoding the integrase, the AlpA regulator and the primase were
274 identified in all the elements, as expected, as they were used to identify them. We showed above
275 that a gene encoding unknown function overlapping *int* on four nucleotides (*IOLG*) is essential
276 for the excision of the PICMI₁₁₅. Among the 35 PICMI subfamilies, 21 carry a *IOLG* homolog
277 that overlaps *int* over 1, 4 or 13 nucleotides, ATGA being the most frequent overlapping site
278 (n=14) (Fig. 4 and Fig. S14). Four elements carry a gene contiguous to the *int* gene (*ICG*) and
279 three do not carry a gene between *int* and *alpA*. The *IOLG* or *ICG* were grouped in 13 distinct
280 gene families (>20% protein identities, 80% LmaxRap), all of them of unknown function. The
281 remaining seven PICMI elements were more divergent in gene repertoires and gene order. In
282 six of them, *alpA* was in two or three copies. These loci may thus have been the result of
283 multiple events of integration, gene loss, and recombination with other satellites. *UPI*
284 homologs were found in 12 subfamilies of PICMI (Fig. 4 and Fig. S14), in line with our
285 observations that *alpA*, *UPI*, and *prim* form the early regulon activated by the helper phage
286 (Fig. 3). In 26 other PICMI subfamilies, single genes were also present between *alpA* and *prim*,

287 forming eight distinct families, each encoding an unknown function. On the nine remaining
288 subfamilies, *alpA* was adjacent to *prim*. Altogether, our analysis revealed that PICMI-like
289 elements have small genomes, are widely distributed in the *Vibrionaceae*, and encode a limited
290 number of genes that are essential for its lifestyle.

291 **Identification of a new defense system in PICMI₁₁₅**

292 In spite of the small size of PICMIs, all subfamilies have accessory genes and some of them
293 encode for known phage defense systems (Table S3), namely Restriction modification systems
294 type I and II, a retron type II, and Paris type I³³. These systems are located in the locus between
295 *prim* and *fis*, suggesting that this might be a hotspot for the acquisition of anti-viral defense
296 genes, akin to the locus between the integrase and *Psu* in P4-like satellites²⁰. This also suggests
297 that PICMIs can provide viral defense mechanisms to their host, as observed both in P4-like
298 satellites and in PICI²¹. Consistent with this, we observed that the presence of PICMI₁₁₅ in V511
299 (transductants) greatly affected the infection outcome of this bacteria by the phage Φ 511 (Fig.
300 5A), showing an antiviral effect of the PICMI₁₁₅.

301 We hypothesized that PICMI₁₁₅ immunity is mediated at least in part by a new defense system,
302 localized between *prim* and *fis* genes and encoded by the gene *UP2*. To test this hypothesis, we
303 cloned *UP2* under the control of its native promoter in a plasmid and transferred it through
304 conjugation to 46 other *V. chagasii* strains that are susceptible to at least one phage²². As a
305 control, the same plasmid expressing GFP was transferred to the strains. Among the 90 possible
306 host and phage combinations that led to the production of phage progeny, eight combinations,
307 involving eight different phages, were affected by UP2 (Fig. 5B). Six out of the eight phages
308 affected by UP2 (Fig. 5C) belong to the same family, as defined by Virus Intergenomic Distance
309 Calculator (VIRIDIC³⁴) with pairwise identities ranging from 55 to 70%. Within this family
310 (Fig. S17A), only the helper phage Φ 115 was not affected by UP2 (Fig. S17B). *V. chagasii*
311 strain V511 was susceptible to a member of this VIRIDIC family, Φ 511 (Fig. 5C) and, to a
312 minor extent, to genetically more diverse phages Φ 168 and Φ 177 (Fig. S17B). We found that
313 UP2 influences the production of all three phages infecting the strain V511 (Fig. 5B and Fig.
314 S17B). UP2 anti-viral activity seemed, however, dependent on the V511 genetic background,
315 as it was not observed for the combination involving phages Φ 177 and Φ 168 and other hosts,
316 including the “original host” that was used to isolate the phage (Fig. 5B). It is noteworthy that
317 the effect of the complete PICMI₁₁₅ element on the infection of V511 by Φ 511 was much more
318 pronounced than UP2 alone, suggesting epistatic effects within this genetic element. In the

319 strain carrying UP2, 60 min after the addition of Φ 511, the phage titer in the culture did not
320 change (Fig. 5C, fold change 10^0) in contrast to the GFP control (fold change 10^2). In the strain
321 carrying the full PICMI₁₁₅ the phage titer decreased by $\sim 10^3$ (Fig. 5A). We conclude that PICMI
322 protects the bacterial host from non-helper phage. This protection relies at least in part on a
323 novel UP2 defense system, whose activity is dependent on the host background.

324 **DISCUSSION**

325 Here, we report a new family of phage satellites that is packaged as concatemers in viral
326 particles. This results in a packaged molecule of DNA with a size similar to that of the helper
327 phage. As a result, PICMI does not require gene(s) involved in re-shaping the capsid size. The
328 PICMI family is among the smallest of phage satellites, with PICMI₁₁₅ being the smallest such
329 element with demonstrated activity. At the other end of the spectrum, cf-PICI¹² produce their
330 own capsids dedicated to the exclusive packaging of their genome. Their minimalist gene
331 repertoire seems dedicated to genes for excision and integration, DNA replication, and anti-
332 viral defenses toward competitors of the helper phage. The small size of the PICMI implies a
333 high dependency on the helper phage for activation, packaging, and release of the particles in
334 the bacterial lysate. In our model system of PICMI₁₁₅, vibrio V115 and phage Φ 115, this
335 dependency is not accompanied by any significant cost for the helper phage production. This
336 finding fits with previous results for cf-PICI, in which costs for the helper phages were also
337 insignificant¹².

338 The discovery of the PICMI family, and more specifically the mechanistic characterization of
339 the PICMI₁₁₅ life cycle, validates the previous hypothesis that capsid size reduction is not a
340 common strategy for the marine satellites⁷. Furthermore, PICMI lacks identifiable packaging
341 genes, suggesting that it does not affect the composition of the viral particle, beyond packaging
342 it with its own DNA. What are the advantages of packaging multiple copies of the satellite
343 within native full-sized helper phage capsids? First, non-remodeled capsids might guarantee
344 optimal interactions with the helper phage tail. Second, it diminishes the number of functions
345 that must be encoded by the PICMI genome. Third, having multiple copies of the satellite
346 injected by the viral-like particle into the cell could increase the expression of satellite genes
347 that are necessary for its integration (gene dosage), thus increasing the frequency of integration
348 after transduction. Finally, multiple copies of the extrachromosomal element could potentially
349 reduce the efficiency of host defense. As larger satellite size results in lower numbers of copies
350 packaged, this finding also suggests a potential tradeoff between the acquisition of accessory

351 genes and the efficiency of transduction. This could explain why the genes present in the *prim-*
352 *fts* hotspot region are restricted in number and subject to high turnover. A higher efficiency of
353 transduction by polyploidization underscores a feature that makes PICMI unique. Indeed,
354 PLE's transduction is severely reduced when packaged into ICP1-size capsids as concatemer
355 or 6-7 PLE genomes relative to small remodeled capsids⁹. Future work will be needed to
356 determine how the PICMI element is integrated or maintained as a single copy in the genome
357 as we never detected more than one PICMI-like element in *Vibrionaceae* genomes.

358 PICMI induction requires the infection by its helper phage. Early after infection, the regulon
359 encoding *alpA*, *UPI*, and *prim* is activated. The role of *UPI* is unknown, but it is noteworthy
360 that *UPI* orthologs were found in 12 out of 35 subfamilies of PICMI and distributed in diverse
361 *Vibrio* species (*V. cholerae*, *V. fluvialis*, *V. vulnificus* and *V. chagasii*) (Fig. S14). While *UPI*
362 is not essential for PICMI activation and spread, its maintenance at 100% frequency in these
363 subfamilies suggests it encodes for a trait that is under strong selection. *Prim*, which is
364 necessary for efficient replication of the satellite encodes a putative RNA polymerase that
365 synthesizes short fragments of RNA, which are then used as primers by the DNA polymerase.
366 The *prim* gene is also found in the vast majority of known satellites, in accordance with a core
367 and essential function of this protein in the lifestyle of satellites¹³. AlpA appears as the key
368 regulator of the switch from latency (integrated) to activation (excised) of PICMI. Indeed, its
369 expression is necessary and sufficient to trigger the activation of the satellite in the absence of
370 the helper phage. Our results suggest that the PICMI *int* gene is constitutively expressed and
371 that phage-induced *alpA* is required for the formation of a functional excision complex. Hence,
372 several relevant questions need to be further addressed, such as: I) how does the helper phage
373 activate the early regulon? II) How do AlpA, the integrase, and probably IOLG interact to
374 catalyze the excision of the satellite? The understanding of *alpA* induction mechanism by the
375 helper phage Φ 115 could benefit from analysis of the genomes of closely related phages that
376 infect the V115 strain but induce little or no PICMI.

377 PICMI is the second family of identified satellites induced by a virulent phage. In contrast to
378 PLE, PICMI does not alter the production of its helper phage. We showed that PICMI can
379 confer immunity toward other virulent phages, and we identified a new defense system (UP2)
380 encoded by the satellite. The phage range of UP2 activity appears very narrow, and many
381 phages susceptible to this system are phylogenetically related to the helper phages. Hence, by
382 protecting the bacteria if the phage is not a helper, PICMI is first protecting itself. This system

383 promotes the stable coexistence of both helper phage and satellites within the bacterial
384 populations.

385 Satellite, helper phage, and bacterial host interactions are highly specific. With such a narrow
386 host range, how do the right combinations of phages, satellites, and bacterial hosts interact in
387 the marine environment? We speculate that blooms of specific vibrio strains can dramatically
388 increase the abundances of specific phages and satellites. When colonizing an animal host such
389 as oyster, vibrios can reach a higher density that might favor physical contact and promote
390 phage infection and satellite transduction. The distribution of a satellite is expected to adhere
391 closely to the distribution of the helper phages at a small temporal and spatial scales due to their
392 total dependency.

393 We recently highlighted that many phage defense genes are encoded on large genomic islands,
394 named phage defense elements, but the mechanisms of transfer of these elements remained
395 unexplored^{23,28}. Around 0.6% of marine viral particles (3.2×10^{26} globally) are packaged
396 satellites⁷, and the discovery of PICMI-mediated immunity strongly suggests that phages,
397 including virulent ones, play an important role in the mobility of the phage defense elements.
398 A common view is that only virulent phages should be used for phage therapy to limit horizontal
399 gene transfer. However, our data suggest that this idea should be taken with caution because
400 PICMI₁₁₅ was efficiently transduced by a virulent phage. Indeed, the discovery of PICMI₁₁₅ and
401 its helper virulent phage underscores the importance of understanding the interactions between
402 virulent phages and the mobile genetic elements encoded by their bacterial hosts.

403 **AUTHOR CONTRIBUTIONS**

404 FLR conceived the study, supervised the project and secured funding. RBC, DP, MM and FLR
405 conducted the experiments. DG, JMS and EPCR performed the genomic analyses. RBC, DG,
406 JMS, DP, MM, EPCR, and FLR analysed the data. FLR, RBC, JMS and EPCR wrote the
407 manuscript.

408 **ACKNOWLEDGEMENTS**

409 We thank Agnes Thierry, Céline Loot, Francois-Xavier Barre and Mélanie Blokesch for
410 valuable suggestions. We thank Sophie Le Panse (MERIMAGE, Roscoff), Gwen Tanguy,
411 Erwan Legeay (GENOMER, Roscoff), Karine Cahier and Yannick Labreuche for technical
412 assistance. We thank Jenna Sternberg from Life Science Editors for help with the Manuscript.

413 This work was supported by fundings from the European Research Council (ERC) under the
414 European Union's Horizon 2020 research and innovation program (grant agreement No
415 884988, Advanced ERC Dynamic) to FLR, from the Agence Nationale de la Recherche (ANR-
416 20-CE35-0014 « RESISTE ») to EPCR and FLR. R.B.-C. acknowledges the Spanish Ministerio
417 de Ciencia e Innovación for his FPI predoctoral contract (BES-2017-079730). EPCR lab was
418 funded by Equipe FRM (Fondation pour la Recherche Médicale): EQU201903007835. This
419 work used the computational and storage services (TARS cluster) provided by the IT
420 department at Institut Pasteur, Paris.

421 **DECLARATION OF INTERESTS**

422 Authors declare no competing interests.

423

424 TITLES AND LEGENDS TO FIGURES

425 **Figure 1 Life cycle of PICMI₁₁₅**

426 **A-**Schematic representation of PICMI₁₁₅ integrated between the *zntR* and *fis* genes. Arrows
427 depict inward- and outward-directed primers used to detect the integration site after excision
428 (grey) or excised/circularized PICMI₁₁₅ (black). For transduction experiments (D) PICMI₁₁₅
429 was marked by Cm^R (brown triangle). Brown forward and grey reverse primers were used to
430 control of the integration of PICMI₁₁₅ at the end of *fis* gene.

431 **B-** Estimation of phage and PICMI DNA size. DNA extracted from phage Φ 115 or as control
432 Φ 115pure, were separated on an agarose gel (left panel, SYBR Green stained), and Southern
433 blotted (right panel) with PICMI₁₁₅ or Φ 115 probes. M: molecular marker (Smart ladder
434 Eugentec).

435 **C and D-**The dynamics of excision and replication were explored by qPCR (C) and Southern
436 blot (D). Bar charts show the mean fold change +/- SEM. from three independent experiments
437 (individual dots). In the southern blot C, S and P indicate the concatemeric and single form of
438 PICMI₁₁₅ and phage genome respectively.

439 **E-** To determine transduction frequencies, the phage Φ 115pure was produced from a derivative
440 of V115 carrying a Cm^R marked PICMI₁₁₅ (see A). Δ PICMI₁₁₅ and V511 were used as recipient
441 cells. Results (upper panel) indicate the ratio between the titer of PICMI₁₁₅-Cm^R (CFU/ml) and
442 the titer of phage (PFU/ml). Bar charts show the mean +/- SEM from three independent
443 experiments (individual dots). The integration of PICMI₁₁₅-Cm^R at the end of the *fis* gene was
444 confirmed by PCR (SYBR Green gel stained, lower panel).

445 **F-** Interference with the reproduction of the helper phage. The strain V115 wild type, a
446 derivative lacking the full PICMI₁₁₅ and two clones (c1 and c2) of transductants carrying
447 PICMI₁₁₅-Cm^R, were infected by Φ 115pure at MOI10 for 60 minutes. Bar charts show the mean
448 fold change of phage titer +/- s.d. from three independent experiments (individual dots).
449 Differences between treatments are not statistically significant (ANOVA, P=0.75, F-test).

450

451 **Figure 2. Genes involved in PICMI₁₁₅ activation**

452 **A-** Fold change of the copies number for phage, empty integration site and circularized
453 PICMI₁₁₅, 30 minutes post-infection by the phage F115pure. Bar charts show the mean +/-
454 SEM. from three independent experiments (individual dots).

455 **B-** Phage progenies were produced using DPICMI₁₁₅, PICMI₁₁₅-Cm^R and derivatives lacking a
456 single gene (e.g. Dint). Phage DNAs were separated on an agarose gel and Southern blotted
457 with PICMI₁₁₅ or F115 probes. M: molecular marker (Smart ladder Eugentec). These phages
458 were also used for transduction experiments (Table 1).

459 **C-** For complementation assays, the genes necessary for PICMI₁₁₅ excision, *int/IOLG* and *alpA*
460 or, as control, *gfp* were cloned under the control of the arabinose inducible. Circular from of
461 PICMI₁₁₅ was detected by classical PCR and gel stained (upper panel) or qPCR (fold change
462 30 min post infection, lower panel). Bar charts show the mean +/- SEM. from three independent
463 experiments (individual dots). The expression of *alpA* is sufficient to induce DPICMI₁₁₅
464 activation in the absence of phage (upper panel, time 0 min), explaining the similar fold change
465 between the *Dalp* mutant and its complemented derivative (lower panel).

466 **Figure 3. The activation of PICMI₁₁₅ is helper phage specific.**

467 **A-** PICMI₁₁₅ DNA was less or not detected in F115 genetically related phages. DNA from viral
468 particles were extracted and separated on SYBR Green stained gel and Southern blotted with
469 an PICMI₁₁₅ or F115 probes. M: molecular marker (Smart ladder Eugentec).

470 **B-** An efficient induction of the satellite is specific to the helper phage F115. The vibrio V115
471 carrying a Cm^R marked PICMI₁₁₅ was infected with the diverse phages at a MOI of 10 for the
472 indicated time, PCR amplicon corresponding to the circularized and concatemeric form of
473 PICMI₁₁₅ were visualized on agarose gel.

474 **C and D-** The vibrio V115 was infected with F115pure (**C**) or F27 (**D**). Each of the six genes
475 from PICMI₁₁₅, the two flanking genes *fis*, *zntR* and the house keeping gene *gyrA* were detected
476 by qRT-PCR. Bar charts (same color code than in Fig. 1A) show the mean +/- SEM fold change
477 from three independent experiments (individual dots).

478

479 **Figure 4. PICMI-like satellite distribution and gene content in the *Vibrionaceae*.**

480 Phylogenetic persistent core tree and genomic representation of the 97 PICMI elements found
481 in *Vibrionaceae* (GenBank 01-27-23 containing 19189 organisms). Genus, super clades or of
482 species names are indicated in the grey boxes. "P." correspond to *Photobacterium* genus,
483 Harveyi, Splendidus are super clades encompassing several *Vibrio* species. The PICMI₁₁₅
484 element is pinpointed by bold strain name (34_P_115) and by an asterisk. Solid colors indicate
485 core PICMI₁₁₅ genes. Grey colors indicate accessory and singleton PICMI-like genes defined
486 using reciprocal best-hit with 20% identity for 50% coverage.

487 **Figure 5. PICMI₁₁₅ and UP2 confer host immunity to specific phages.**

488 **A-** PICMI₁₁₅ greatly affected the infection outcome by the phage Φ 511. *Vibrio* strain V511 (wt)
489 and transductants carrying the full satellite (PICMI₁₁₅-Cm^R, two clones c1 and c2) were infected
490 with F511 at an MOI 10 for 60 minutes. Bar charts show the mean fold change of phage titer
491 +/- SEM from three independent experiments (individual dots).

492 **B-** UP2 encodes a novel defense system which activity depends on the host genetic background.
493 A plasmid carrying the gene *UP2* under the control of its native promoter or, as control, the *gfp*
494 under the constitutive promoter P_{LAC}, were transferred to diverse *V. chagasii* strains. Tenfold
495 dilutions of a phage were spotted on susceptible strain (black and red squares). Rows represent
496 sequenced *Vibrio* strains ordered according to the Maximum Likelihood persistent genome
497 phylogeny of *V. chagasii* (n=46) Columns represent phages (n=48) ordered by VIRIDIC
498 clustering dendrogram. Change in susceptibility between UP2 and GFP strain strains are
499 indicated by a red square.

500 **C-** The susceptibility to UP2 of six phages that belong to the same VIRIDIC family than F115
501 was tested using their original host, i.e. the host used to isolate these phages. To this aim fold
502 change of phage titer (as determined in A) was compared between GFP vs UP2 carrying host.
503

504 **Table 1. PICMI₁₁₅ transfer by Φ 115 phage produced from different donors**

505

Donor strain	PICMI₁₁₅ titre^a
PICMI ₁₁₅ -Cm ^R	5,85E+03 ± 1,62E+03
<i>Δint</i> -Cm ^R	<1
<i>ΔIOLG</i> -Cm ^R	<1
<i>ΔalpA</i> -Cm ^R	<1
<i>ΔUP1</i> -Cm ^R	4,78E+03 ± 1,98E+03
<i>Δprim</i> -Cm ^R	<1
<i>ΔUP2</i> -Cm ^R	4,88E+03 ± 1,52E+03

506

507 ^aPICMI₁₁₅ titer/ml of lysate, using vibrio Δ PICMI₁₁₅ as recipient strain. The titer of phages
508 produced from the different donors was 10⁶ PFU/ml. The means and standard deviations from
509 four independent experiments are presented.

510

511 **METHODS**

512 **RESOURCE AVAILABILITY**

513 **Lead contact**

514 Further information and requests for resources and reagents should be directed to and will be
515 fulfilled by the Lead Contact, Frédérique Le Roux: fleroux2014@gmail.com.

516 **Materials availability**

517 Strains, phages, and plasmids generated in this study are available upon request and without
518 restrictions from the lead contact upon request.

519 **Data and code availability**

520 Accession numbers of vibrio and phage genomes isolated and sequenced in²² are listed in the
521 Key Resources table. This paper does not report original code. All programs used to analyze
522 genomes were previously reported and are freely available online (see key resources table). The
523 MacSyFinder models used to identify PICMI are available upon request and can be used with
524 the MacSyFinder to make novel analysis. Any additional information required to reanalyze the
525 data reported in this paper is available from the lead contact upon request.

526 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

527 **Bacterial strains and growth conditions**

528 Phages and bacterial strains used in this study are listed in Table S4 and S5 respectively. Strains
529 used or established for the genetic approach are presented in Table S6. *Vibrio. chagasii* isolates
530 were grown in marine agar (MA, Difco) or marine broth (MB) at RT with gentle agitation.
531 *Escherichia coli* strains were grown at 37°C in Luria-Bertani (LB, Difco) agar or in LB broth
532 with shaking (250 r.p.m.). Chloramphenicol (Cm; 5 or 25µg/ml for *V. chagasii* and *E. coli*,
533 respectively), thymidine (0.3 mM) and diaminopimelate (0.3 mM) were added as supplements
534 when necessary (all chemicals from Sigma-Aldrich). Induction of the P_{BAD} promoter was
535 achieved by the addition of 0.2% L-arabinose to the growth media, and conversely, was
536 repressed by the addition of 1% D-glucose. Conjugation between *E. coli* and vibrios were
537 performed at 30°C as described previously³⁵ with the exception that we used TSA-2 (Tryptic
538 Soy Agar supplemented with 1.5% NaCl) instead of LB for mating and selection. Briefly,
539 overnight cultures of donor and recipient were diluted at 1:100 in culture media without
540 antibiotic and grown up an OD_{600nm} of 0.3. The mating was performed on TSA-dap using a

541 donor/recipient ratio of 5/1. Counter-selection of *ΔdapA* donor was done by plating on a TSA
542 devoid of diaminopimelic acid (DAP) and supplemented with antibiotic.

543 **Phage isolation, high titer stock, and titration**

544 New phages infecting V115 were isolated from concentrated seawater viruses sampled in
545 summer 2021 in the same oyster farm and using the same protocol than in^{22,23}. A volume of 100
546 μl of an overnight (ON) culture of bacterial host and 20 μl of viruses were directly plating on a
547 bottom agar plate (1.5% agar, in MB) and 3.5 ml molten top agar (55°C, 0.4% agar, in MB)
548 were added to form host lawns in overlay and allow for plaque formation³⁶. Plaque plugs were
549 first eluted in 500 μl of MB for 24 hours at 4°C, 0.2-μm filtered to remove bacteria, and re-
550 isolated twice on V115 for purification before storage at 4°C and, after supplementation of 25%
551 glycerol at -80°C. High titer stocks (>10¹¹ PFU/ml) were generated by confluent lysis in agar
552 overlays³⁶. To determine the titer of phage, bacterial lawns were prepared by mixing 100 μl of
553 on overnight culture of cells with top agar and poured onto plates. Then, tenfold dilutions of
554 phage were spotted on plate, which were incubated at RT for 24 h.

555 **METHOD DETAILS**

556 **Plasmid construction**

557 The primers and plasmids used or established in this study are listed in Table S7 and S8
558 respectively. For the preparation of quantitative PCR (qPCR) standards, each amplicon was
559 PCR amplified using the RedTaq polymerase (VWR) and cloned in the plasmid pCR2.1 using
560 the TOPO-TA CloningTM Kit (Invitrogen).

561 For vibrio mutagenesis, cloning was performed using Herculase II fusion DNA polymerase
562 (Agilent) for PCR amplification and the Gibson Assembly Master Mix (New England Biolabs,
563 NEB) for insert-plasmid assembly, according to the manufacturer instructions. All cloning was
564 confirmed by digesting plasmid minipreps with specific restriction enzymes and/or sequencing
565 (Eurogentec).

566 **Nucleic acid extraction, amplification, and southern blot**

567 Prior to DNA extraction, phage suspensions (5 ml, >10¹¹ PFU/ml) were concentrated to
568 approximately 500 μl on centrifugal filtration devices (30 kDa Millipore Ultra Centrifugal
569 Filter, Ultracel UFC903024) and washed with 1/100 MB to decrease salt concentration.

570 Alternatively, phages were concentrated using PEG 8000 1X and NaCl 1M, incubated ON at
571 4°C, centrifuged 30 min at 4500 rpm, and the pellet was resuspended in 500 µl SM buffer (NaCl
572 100 mM, MgSO₄·7H₂O 8 mM, Tris-Cl 50 mM). The concentrated phages were next treated for
573 30 min at 37°C with 10 µl of DNase (Promega) and 2.5 µl of RNase (Macherey-Nagel) at
574 1000 unit and 3.5 mg/ml, respectively. The nucleases were inactivated by adding EDTA (20
575 mM, pH 8). DNA extraction encompassed a first step of protein lysis (0.02 M EDTA pH 8.0,
576 0.5 mg/ml proteinase K, 0.5% sodium dodecyl sulfate) for 30 min incubation at 55°C, a phenol
577 chloroform extraction, and an ethanol precipitation. Bacterial DNA was extracted using the
578 Wizard Genomic DNA Purification Kit (Promega).

579 RNA was extracted with TRIzol™ Reagent (Sigma-Aldrich) and High Pure RNA Isolation Kit
580 (Roche), treated by TURBO DNase (Ambion) and reverse transcribed using the Transcriptor
581 First Strain cDNA Synthesis Kit (Roche).

582 Classical PCRs were performed using the RedTaq (WVR) and amplicons were visualized by
583 SYBR Green stained (Sigma) agarose gel electrophoresis (1 to 2% agarose). qPCR and qRT-
584 PCR was performed using LightCycler 480 SYBR Green I Master (Roche). The thermal cycling
585 conditions were 95°C for 10 min, followed by 40 cycles of 95°C for 10 s, 60°C for 20 s and
586 72°C for 25 s, then 1 cycle of 95°C for 5 s, 65°C for 1 min and 95°C for 15 s. Standard curves
587 were constructed using serial dilutions of plasmid, leading to the number of DNA copies per 20
588 ng of DNA. Number of copies for phage, empty integration site, and circularized PICMI₁₁₅ were
589 normalized by the number of copies of vibrio (*gyrB*) per sample. For the qRT-PCR, the resulting
590 copies number were normalized on *gyrA* for each sample. To determine fold change, samples
591 collected at different time post-infection were compared to the sample before adding phages.

592 For Southern blot, DNA samples were run on 0.7% agarose gel at 100V for one hour. Then, the
593 DNA was transferred to Nylon membranes (Hybond-N+; Amersham Life Science) using
594 standard methods. DNA was detected using a DIG-labelled probe (Digoxigenin-11-dUTP
595 alkali-labile), anti-DIG antibody (Anti-Digoxigenin-AP Fab fragments) and Chemiluminescent
596 detection with CSPD following the instructions of the kit (all products and kits from ROCHE).

597 **Construction of HiC libraries, sequencing, and analysis.**

598 1 ml of different mix of high titer stocks (>10¹¹ PFU/ml) of phages (mix1: 1 ml of Φ115, mix2:
599 1 ml of Φ191, mix3: 500 µl of Φ115 + 500 µl of Φ191) were fixed in a 5 ml Eppendorf tube
600 by adding formaldehyde (Sigma-Aldrich, ref - F4775, Formalin 35-36.5% plus methanol 15%)
601 to a final concentration of 3% and incubated at RT for 1 hour under gentle agitation. The

602 reaction was stopped by adding glycine (stock = 2.5 M) to a final concentration 0.125 M and
603 incubated at RT for 20 min under gentle agitation. Fixed particles were then centrifuged at
604 16,000 x g for 20 minutes at 4°C. Supernatant was discarded, resuspended in 1 ml of PBS 1X,
605 and recentrifuged at 16,000 x g for 20 minutes at 4°C. The supernatant was again discarded
606 carefully and the pellet was resuspended in 45 µl of Tris 10 mM pH 7.5. The HiC libraries were
607 then constructed using the ARIMA Kit (Arima Genome-Wide HiC+ Kit). HiC genomic
608 libraries were then processed for sequencing as previously described³⁷ and were sequenced on
609 Nextseq550 apparatus (2 x 35 bp). Contact maps were generated using Hicstuff³⁸ (bowtie2 -
610 very sensitive local mode – mapping quality of 30) and a reference FastA files containing the 3
611 phage genomes. Contact maps were then binned at 1kb resolution, balanced, and displayed
612 using Hicstuff.

613 **Electron microscopy**

614 Following concentration on centrifugal filtration devices (Millipore, amicon Ultra centrifugal
615 filter, Ultracel 30K, UFC903024), 20 µl of the phage concentrate were adsorbed for 10 min to
616 a formvar film on a carbon-coated 300 mesh copper grid (FF-300 Cu formvar square mesh Cu,
617 delta microscopy). The adsorbed samples were negatively contrasted with 2% Uranyl acetate
618 (EMS, Hatfield, PA, USA). Imaging was performed using a Jeol JEM-1400 Transmission
619 Electron Microscope equipped with an Orious Gatan camera at the platform MERIMAGE
620 (Station Biologique, Roscoff, France).

621 **Vibrio mutagenesis**

622 PICMI labelling (PICMI₁₁₅-Cm^R) was performed by cloning the 500bp end of UP2 gene in the
623 suicide plasmid pSW23T³⁹. To inactivate UP2 or *prim* gene and at the same time label the
624 PICMI derivatives (Δ prim and Δ UP2 -Cm^R), a 500bp internal region of the gene was cloned in
625 the suicide plasmid pSW23T. After conjugative transfer, selection of the plasmid-borne drug
626 marker (Cm^R) resulted from integration of pSW23T in the target region by a single crossing-
627 over. The integration of the suicide plasmid was verified by PCR.

628 Gene deletion was performed by cloning 500bp fragments flanking the gene in the pSW7848T
629 suicide plasmid⁴⁰. This pSW23T derivative vector encodes the *ccdB* toxin gene under the
630 control of an arabinose-inducible and glucose-repressible promoter, P_{BAD}³⁵. Selection of the
631 plasmid-borne drug marker on Cm and glucose resulted from integration of pSW7848T in the
632 genome. The second recombination leading to pSW7848T elimination was selected on
633 arabinose-containing media. Mutants were screened by PCR using external primers.

634 For the complementation experiments, the genes necessary for PICMI₁₁₅ excision, *int/IOLG*
635 and *alpA*, or a *gfp* control were cloned under the control of the conditional P_{BAD} promoter in a
636 P15A-*ori*-based replicative vector. The plasmids were transferred by conjugation into the
637 mutants. Strains were grown to mid-exponential phase in the presence of 0.2% arabinose
638 (activation of P_{BAD}) and then infected with Φ 115pure for 30 min. To explore its anti-phage
639 activity, UP2 under its native promoter was cloned in a pMRB plasmid⁴¹, and the same plasmid
640 expressing the GFP was used as control.

641 We were not able to delete the complete PICMI₁₁₅ by allelic exchange using the pSW7848T
642 suicide plasmid. As an alternative, we cloned the *alpA* gene of the *V. chagasii* or *V. aestuarianus*
643 under the control of P_{BAD} promoter in a P15A-*ori*-based replicative vector (Spec^R), assuming
644 that the expression of *alpA* in trans is sufficient to induce PICMI₁₁₅ excision even in the absence
645 of phage (Fig. 2C) and the *alpA* from the two *Vibrio* species are interchangeable for PICMI₁₁₅
646 induction (Table S9). The plasmid was transferred by conjugation to V115. The Spec^R
647 conjugant was grown overnight in MB with Spec and arabinose, serially diluted and plated on
648 TSA-2 (TSB-2 with agar). A total of 48 colonies were screened by PCR to identify V115
649 derivatives that lack the PICMI (Δ PICMI). We obtained clones with the mutation (8/48) only
650 when using *alpA* from *V. aestuarianus*.

651 **PICMI induction**

652 The vibrio strain was grown to mid-exponential phase in Marine broth (OD=0.3) and infected
653 under static condition with the phage at a multiplicity of infection (MOI) of 10 otherwise
654 indicated. At each time point, an aliquot of the culture was centrifuged, the supernatant was
655 filtered at 0.2 μ m and the titer of phages was determined by drop spotting serial dilutions of the
656 supernatant on the host lawn. To determine fold change, the titer of phage in the lysate was
657 compared to the same amount of phage added in the culture media without bacteria. Total RNA
658 and/or DNA was extracted from the bacterial pellet.

659 **Adsorption estimation.**

660 Phage adsorption experiments were performed as previously described⁴². Phages were mixed
661 with exponentially growing cells (OD 0.3; 10⁷ CFU/ml) at a MOI of 0.01 and incubated at RT
662 without agitation. At different time points, 250 μ l of the culture was transferred in a 1.5 ml tube
663 containing 50 μ l of chloroform and centrifuged at 14,000 rpm for 5 min. The supernatant was
664 10-fold serially diluted and drop spotted onto a fresh lawn of a sensitive host to quantify the

665 remaining free phage particles. In this assay, a drop in the number of infectious particles at 15
666 or 30 min indicated bacteriophage adsorption.

667 **PICMI transduction**

668 The number of PICMI particles were quantified using the transduction titering assay. Briefly,
669 lysates were produced by infecting V115 derivatives carrying PICMI₁₁₅-Cm^R and derivatives
670 by Φ 115pure. A 1:100 dilution (in fresh MB broth) of an overnight recipient strain was grown
671 until an OD₆₀₀ of 0.3 was reached. Then, 100 ml of the recipient culture was dispatched in a 96
672 well plate, infected by addition of 10 μ l of PICMI lysate serial dilutions prepared with MB for
673 1H at RT. The different mixtures of culture-PICMI-Cm^R were plated out on TSA-2 plates
674 containing chloramphenicol. LBA plates were incubated at RT for 24 h, and the number of
675 colonies formed (transduction particles present in the lysate) were counted and represented as
676 the colony forming units (CFU/ml). PCRs were performed to confirm the integration of PICMI
677 at the end of the *fis* gene.

678 ***In silico* prediction and analysis of PICMI-like element**

679 The PICMI-like elements were searched using two datasets: 1) the bacterial division of
680 GenBank release v243 (5/26/2021) that contains 24,243 complete genomes, including 456
681 genomes of the *Vibrionaceae* family and 2) the NCBI Assembly database (2/16/2023) with
682 19,185 *Vibrionaceae* genomes available but with variable assembly quality. SatelliteFinder
683 (v0.9.1)¹³ was used on both datasets with a dedicated PICMI model defined by four mandatory
684 genes encoding: the integrase (PF00589.25, PF00239.24, PF07508.16), AlpA (PF05930.15,
685 PF12728.10), the primase (DUF3987 with PF13148.9 and DUF5906 with PF19263.2) and the
686 Fis regulator (PF02954.22). The resulting elements were then filtered by excluding those with
687 *int*, *alpA*, and *fis* localized in different contigs, those predicted to belong to other families (PICI,
688 cf-PICI, P4 and PLE), and those integrated into a gene showing a lower identity with *fis*. The
689 genomic region starting with the *fis* gene and ending with the direct repeat upstream the *int* gene
690 was extracted, aligned with FAMSA (v1.6.2)⁴³, and each PICMI subfamily was defined by a
691 pairwise nucleic identity $\geq 90\%$ (Table S3 and Fig. S14).

692 The PICMI genes were clustered in families using mmseqs2 (v14.7e284) reciprocal best-hit⁴⁴
693 with 20% identity and 50% coverage thresholds (Table S3 and Fig. S14). Phage defense systems
694 were annotated using Defense-Finder (v1.0.8 and models v1.1.0)³³ and phage structural genes
695 were annotated using PhANNs (v1.0.0)⁴⁵ with a threshold score ≥ 7 .

696 The functional annotation of genes used multiple approaches, i.e. tblastn similarity searches on
697 GenBank, InterPro domain prediction⁴⁶, and ProtNLM annotation
698 [<https://www.uniprot.org/help/ProtNLM>]. The 3D protein structures predicted by ColabFold
699 (v1.5.2-patch)³⁰ were compared to publicly available protein structures using the Foldseek
700 search server (v5)⁴⁷.

701 Comparative genomics were performed using PanACoTA workflow (v1.4.0)⁴⁸. Persistent
702 genes were defined as present in single copy in at least 90% genomes with a minimum of 30%
703 protein identity. Protein sequences of each family were first aligned and concatenated.
704 Phylogenetic reconstruction was done using iqtree2 (v2.0.3)⁴⁹ with 1000 bootstraps and GTR
705 model. Genome plots were generated using dedicated python scripts based on the ‘DNA
706 Features Viewer’ library ([https://github.com/Edinburgh-Genome-
707 Foundry/DnaFeaturesViewer](https://github.com/Edinburgh-Genome-Foundry/DnaFeaturesViewer)).

708 We clustered phages using VIRIDIC (v1.1, default parameters)³⁴. Intergenomic similarities
709 were identified using BLASTN pairwise comparisons. Viruses’ assignment into genera ($\geq 70\%$
710 similarities) and species ($\geq 95\%$ similarities) ranks follows the International Committee on
711 Taxonomy of Viruses (ICTV) genome identity thresholds. We used PHANOTATE v1.5.0⁵⁰ for
712 the syntactic annotation of phage $\Phi 115$ and, after removing restriction on gene size, analysis of
713 large non-coding regions within PICMIs.

714 **Nanopore genome assembly and analysis**

715 The Nanopore sequencing library was prepared using Native Barcoding genomic DNA (EXP-
716 NBD104) and Ligation Sequencing Kit 1D (SQK-LSK109) and sequenced using MinION flow
717 cell R9.4.1 at the platform GENOMER (Station Biologique, Roscoff, France).

718 Demultiplexing and base calling of raw nanopore sequencing data (Table S1) was performed
719 using Guppy software (v6.1.1, --flowcell FLO-MIN106 --kit SQK-LSK109). The base called
720 sequences were used as input for genome assembly performed using FLYE (v2.9)⁵¹ RAVEN
721 (v1.4.0)⁵² and NECAT (v0.0.1)⁵³ with default parameters.

722 The comparisons of Illumina and Nanopore assemblies of both $\Phi 115$ and PICMI₁₁₅ were
723 performed using pairwise FAMSA alignment (v1.6.2)⁴³. FLYE was selected for further analysis
724 because it showed the highest similarity with previous Illumina sequencing. Nanopore PICMI
725 reads were analyzed using a dedicated script based on FAMSA alignment to precisely
726 determine read start and end on an artificial 13 copies concatemer reference.

727

728 **REFERENCES**

- 729 1. Breitbart, M., Bonnain, C., Malki, K., and Sawaya, N.A. (2018). Phage puppet masters
730 of the marine microbial realm. *Nat Microbiol* 3, 754-766. 10.1038/s41564-018-0166-y.
- 731 2. Suttle, C.A. (2005). Viruses in the sea. *Nature* 437, 356-361. 10.1038/nature04160.
- 732 3. Fillol-Salom, A., Martinez-Rubio, R., Abdulrahman, R.F., Chen, J., Davies, R., and
733 Penades, J.R. (2018). Phage-inducible chromosomal islands are ubiquitous within the
734 bacterial universe. *Isme J* 12, 2114-2128. 10.1038/s41396-018-0156-3.
- 735 4. Fillol-Salom, A., Miguel-Romero, L., Marina, A., Chen, J., and Penades, J.R. (2020).
736 Beyond the CRISPR-Cas safeguard: PICI-encoded innate immune systems protect
737 bacteria from bacteriophage predation. *Curr Opin Microbiol* 56, 52-58.
738 10.1016/j.mib.2020.06.002.
- 739 5. Ibarra-Chavez, R., Hansen, M.F., Pinilla-Redondo, R., Seed, K.D., and Trivedi, U.
740 (2021). Phage satellites and their emerging applications in biotechnology. *FEMS*
741 *Microbiol Rev* 45. 10.1093/femsre/fuab031.
- 742 6. Penades, J.R., and Christie, G.E. (2015). The Phage-Inducible Chromosomal Islands: A
743 Family of Highly Evolved Molecular Parasites. *Annu Rev Virol* 2, 181-201.
744 10.1146/annurev-virology-031413-085446.
- 745 7. Eppley, J.M., Biller, S.J., Luo, E., Burger, A., and DeLong, E.F. (2022). Marine viral
746 particles reveal an expansive repertoire of phage-parasitizing mobile elements. *Proc*
747 *Natl Acad Sci U S A* 119, e2212722119. 10.1073/pnas.2212722119.
- 748 8. Hackl, T., Laurenceau, R., Ankenbrand, M.J., Bliem, C., Cariani, Z., Thomas, E.,
749 Dooley, K.D., Arellano, A.A., Hogle, S.L., Berube, P., et al. (2023). Novel integrative
750 elements and genomic plasticity in ocean ecosystems. *Cell* 186, 47-62 e16.
751 10.1016/j.cell.2022.12.006.
- 752 9. Boyd, C.M., Subramanian, S., Dunham, D.T., Parent, K.N., and Seed, K.D. (2023). A
753 *Vibrio cholerae* viral satellite maximizes its spread and inhibits phage by
754 remodeling hijacked phage coat proteins into small capsids. *bioRxiv*,
755 2023.2003.2001.530633. 10.1101/2023.03.01.530633.
- 756 10. Moura de Sousa, J.A., and Rocha, E.P.C. (2022). To catch a hijacker: abundance,
757 evolution and genetic diversity of P4-like bacteriophage satellites. *Philos Trans R Soc*
758 *Lond B Biol Sci* 377, 20200475. 10.1098/rstb.2020.0475.

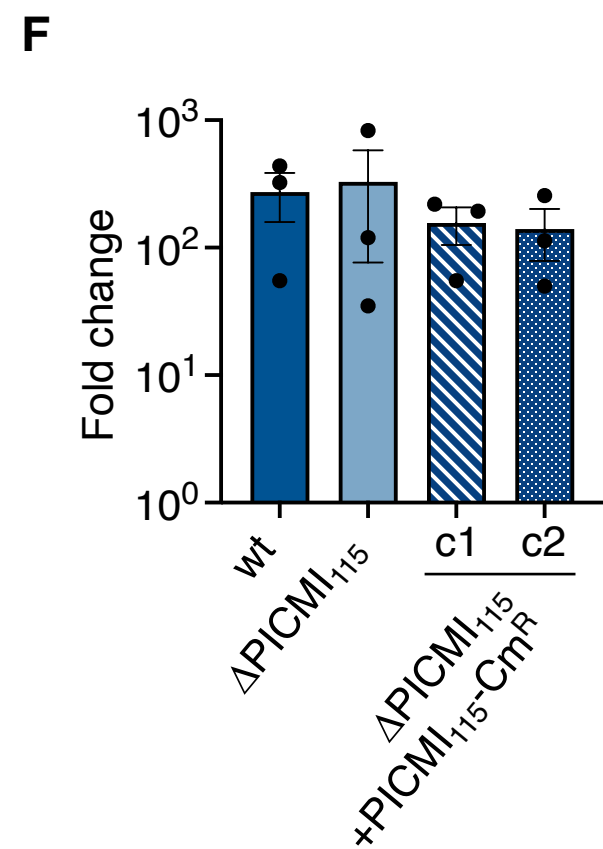
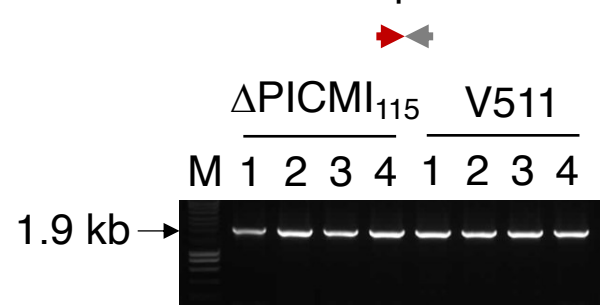
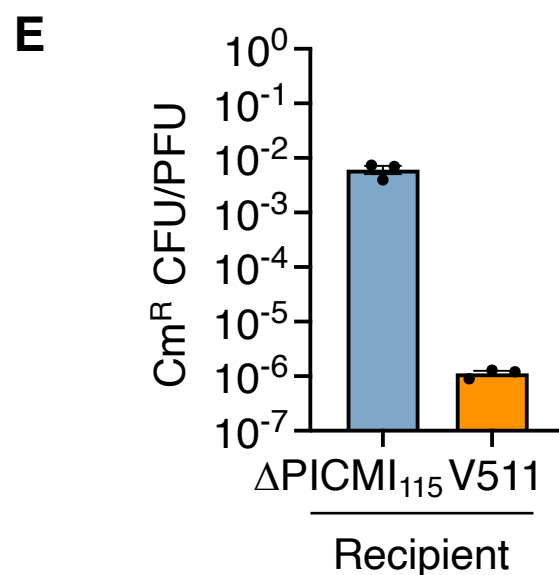
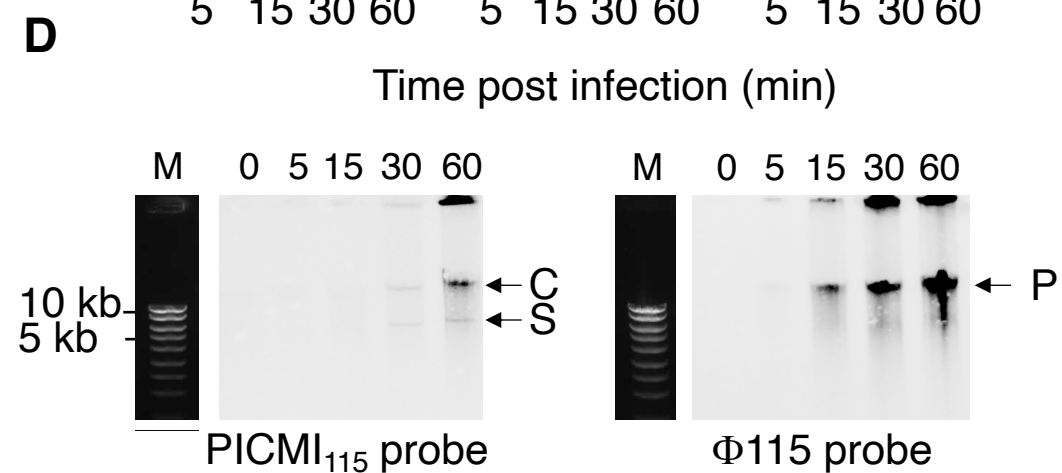
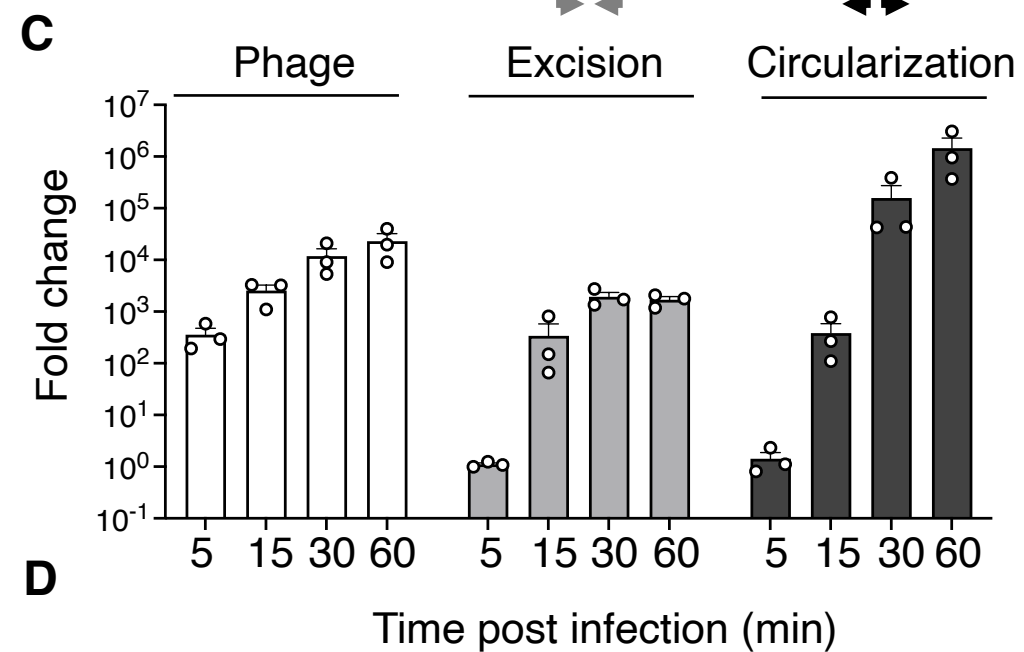
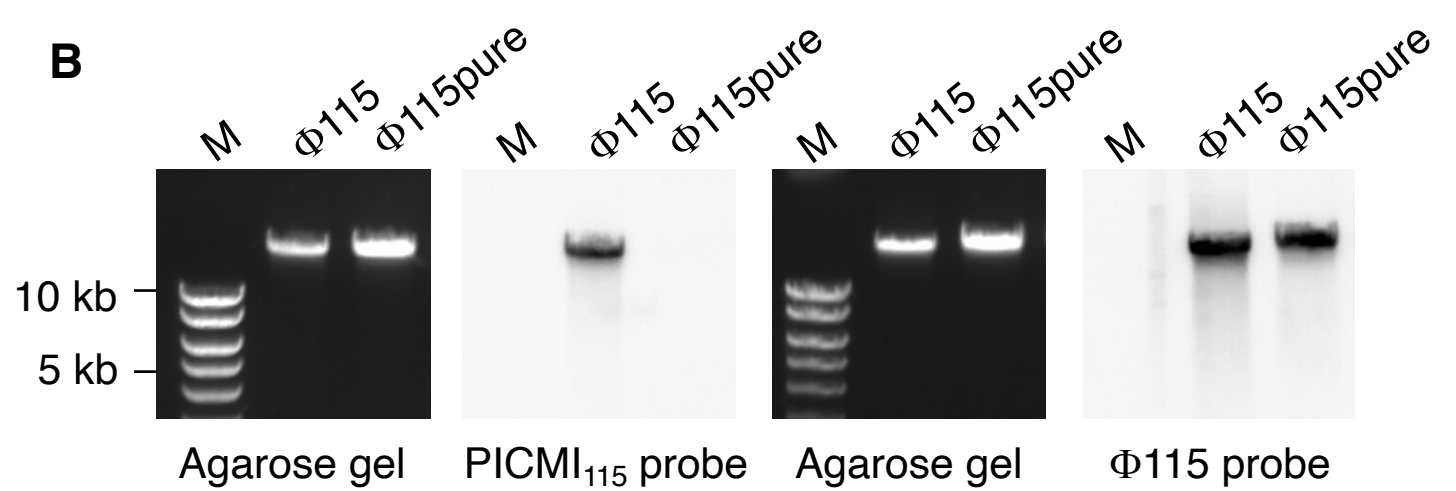
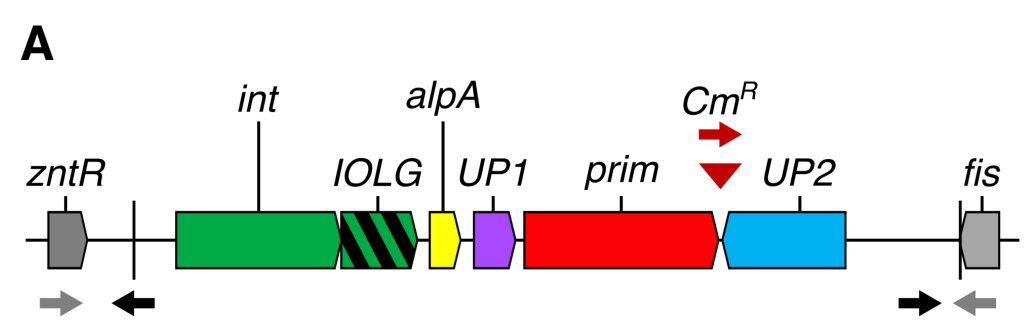
- 759 11. Martinez-Rubio, R., Quiles-Puchalt, N., Marti, M., Humphrey, S., Ram, G., Smyth, D.,
760 Chen, J., Novick, R.P., and Penades, J.R. (2017). Phage-inducible islands in the Gram-
761 positive cocci. *Isme J* *11*, 1029-1042. 10.1038/ismej.2016.163.
- 762 12. Alqurainy, N., Miguel-Romero, L., Moura de Sousa, J., Chen, J., Rocha, E.P.C., Fillol-
763 Salom, A., and Penades, J.R. (2023). A widespread family of phage-inducible
764 chromosomal islands only steals bacteriophage tails to spread in nature. *Cell Host*
765 *Microbe* *31*, 69-82 e65. 10.1016/j.chom.2022.12.001.
- 766 13. de Sousa, J.A.M., Fillol-Salom, A., Penades, J.R., and Rocha, E.P.C. (2023).
767 Identification and characterization of thousands of bacteriophage satellites across
768 bacteria. *Nucleic Acids Res.* 10.1093/nar/gkad123.
- 769 14. Barth, Z.K., Silvas, T.V., Angermeyer, A., and Seed, K.D. (2020). Genome replication
770 dynamics of a bacteriophage and its satellite reveal strategies for parasitism and viral
771 restriction. *Nucleic Acids Res* *48*, 249-263. 10.1093/nar/gkz1005.
- 772 15. O'Hara, B.J., Barth, Z.K., McKitterick, A.C., and Seed, K.D. (2017). A highly specific
773 phage defense system is a conserved feature of the *Vibrio cholerae* mobilome. *PLoS*
774 *Genet* *13*, e1006838. 10.1371/journal.pgen.1006838.
- 775 16. Quiles-Puchalt, N., Carpena, N., Alonso, J.C., Novick, R.P., Marina, A., and Penades,
776 J.R. (2014). Staphylococcal pathogenicity island DNA packaging system involving cos-
777 site packaging and phage-encoded HNH endonucleases. *Proc Natl Acad Sci U S A* *111*,
778 6016-6021. 10.1073/pnas.1320538111.
- 779 17. Christie, G.E., and Dokland, T. (2012). Pirates of the Caudovirales. *Virology* *434*, 210-
780 221. 10.1016/j.virol.2012.10.028.
- 781 18. Diana, C., Deho, G., Geisselsoder, J., Tinelli, L., and Goldstein, R. (1978). Viral
782 interference at the level of capsid size determination by satellite phage P4. *J Mol Biol*
783 *126*, 433-445. 10.1016/0022-2836(78)90050-5.
- 784 19. Ram, G., Chen, J., Kumar, K., Ross, H.F., Ubeda, C., Damle, P.K., Lane, K.D., Penades,
785 J.R., Christie, G.E., and Novick, R.P. (2012). Staphylococcal pathogenicity island
786 interference with helper phage reproduction is a paradigm of molecular parasitism. *Proc*
787 *Natl Acad Sci U S A* *109*, 16300-16305. 10.1073/pnas.1204615109.
- 788 20. Rousset, F., Depardieu, F., Miele, S., Dowding, J., Laval, A.L., Lieberman, E., Garry,
789 D., Rocha, E.P.C., Bernheim, A., and Bikard, D. (2022). Phages and their satellites
790 encode hotspots of antiviral systems. *Cell Host Microbe* *30*, 740-753 e745.
791 10.1016/j.chom.2022.02.018.

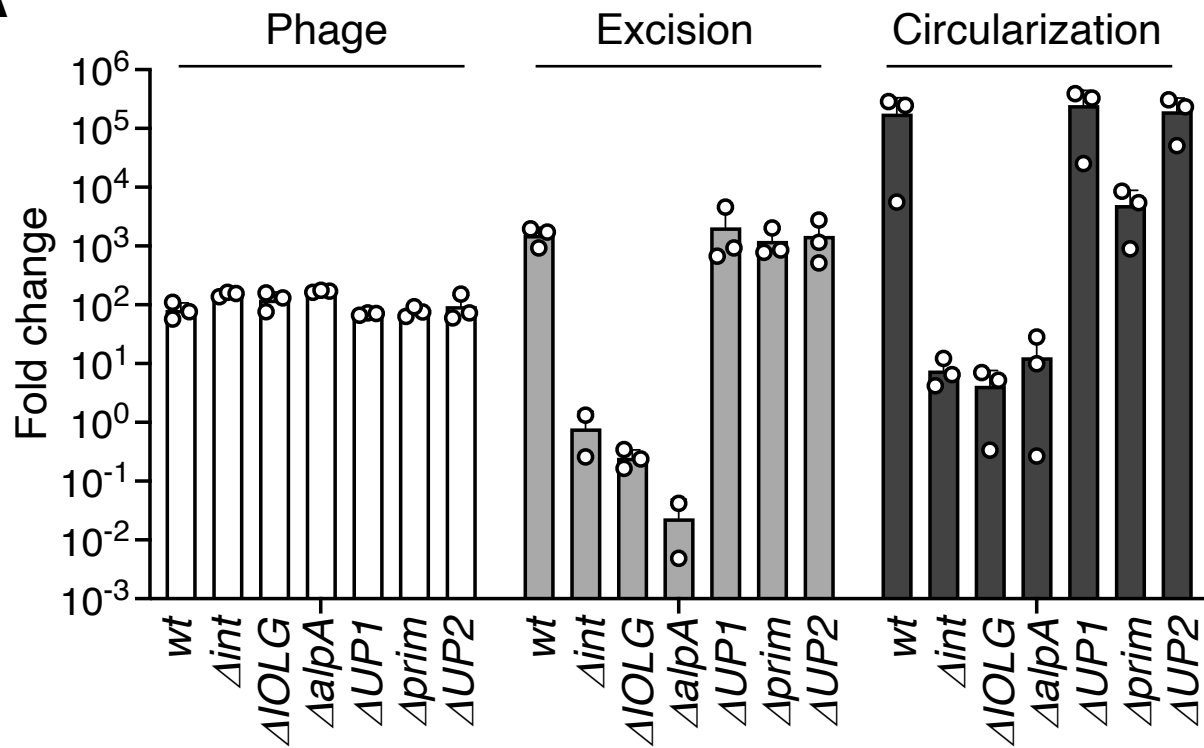
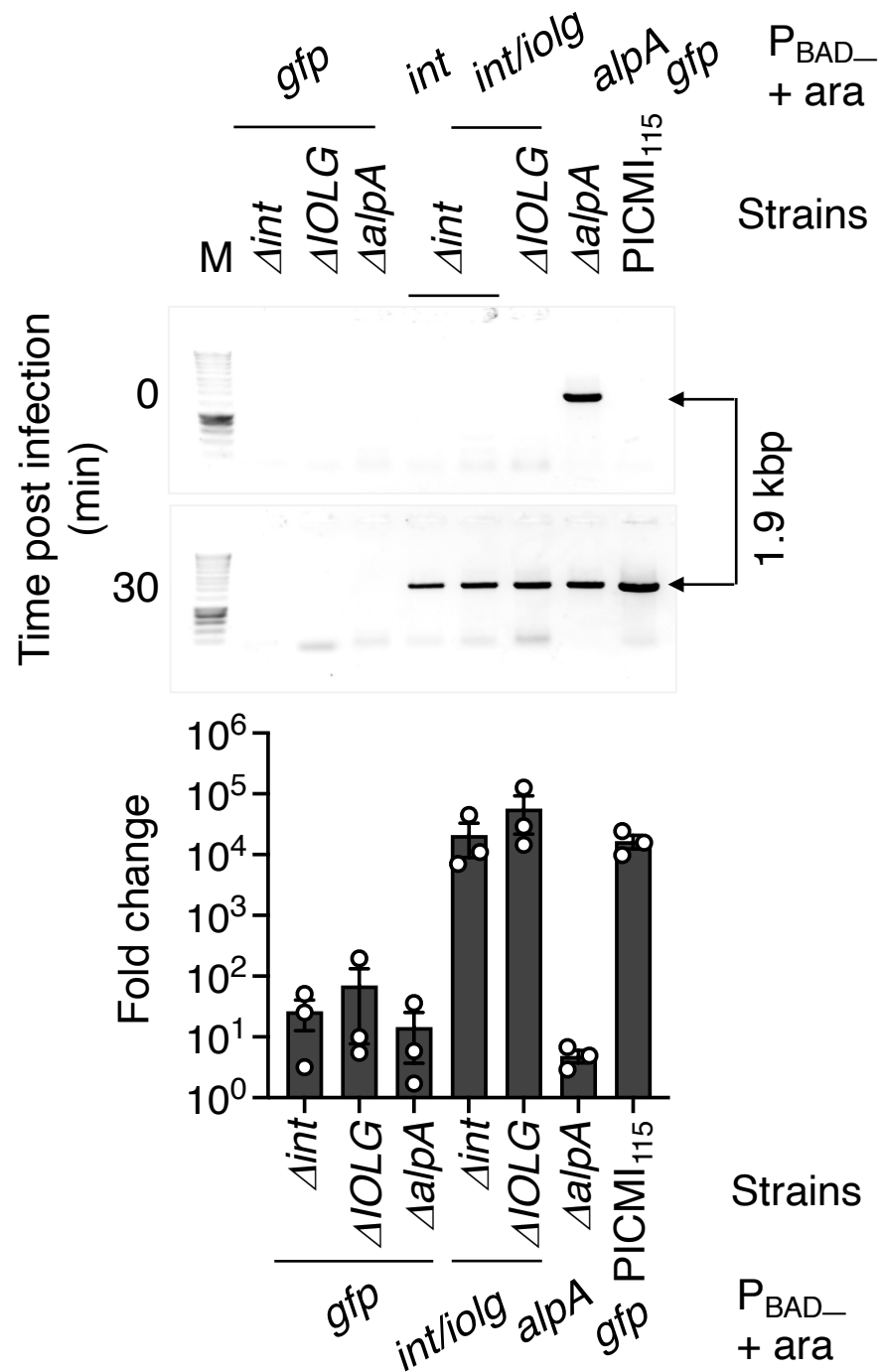
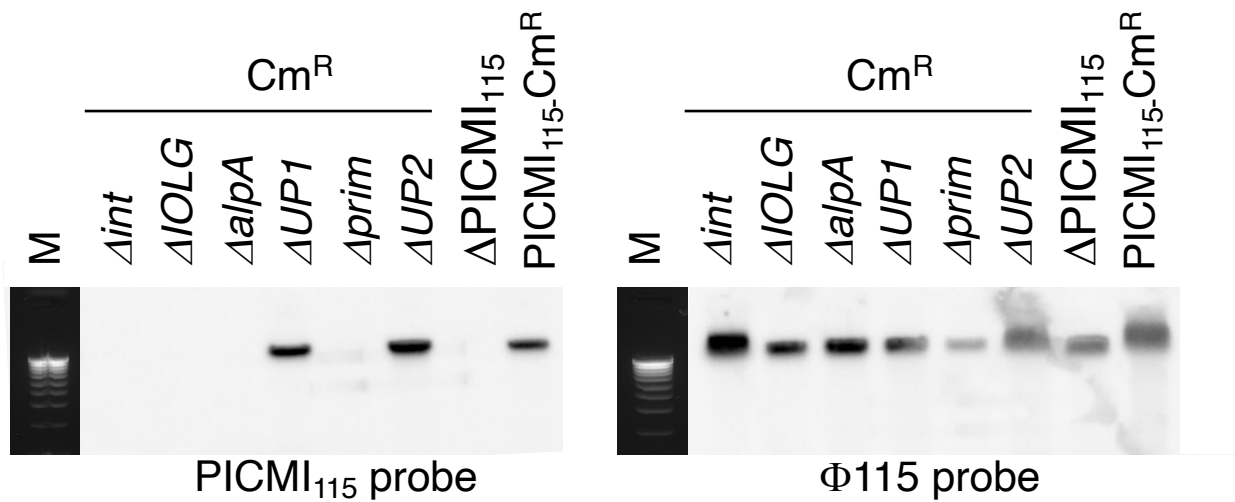
- 792 21. Fillol-Salom, A., Rostol, J.T., Ojiogu, A.D., Chen, J., Douce, G., Humphrey, S., and
793 Penades, J.R. (2022). Bacteriophages benefit from mobilizing pathogenicity islands
794 encoding immune systems against competitors. *Cell* 185, 3248-3262 e3220.
795 10.1016/j.cell.2022.07.014.
- 796 22. Cahier, K., Piel, D., Barcia-Cruz, R., Goudenege, D., Wegner, K.M., Monot, M.,
797 Romalde, J.L., and Le Roux, F. (2023). Environmental vibrio phage-bacteria interaction
798 networks reflect the genetic structure of host populations. *Environ Microbiol.*
799 10.1111/1462-2920.16366.
- 800 23. Piel, D., Bruto, M., Labreuche, Y., Blanquart, F., Goudenege, D., Barcia-Cruz, R.,
801 Chenivresse, S., Le Panse, S., James, A., Dubert, J., et al. (2022). Phage-host coevolution
802 in natural populations. *Nat Microbiol.* 10.1038/s41564-022-01157-1.
- 803 24. Le Roux, F., and Blokesch, M. (2018). Eco-evolutionary Dynamics Linked to
804 Horizontal Gene Transfer in Vibrios. *Annu Rev Microbiol* 72, 89-110.
805 10.1146/annurev-micro-090817-062148.
- 806 25. Le Roux, F., Wegner, K.M., and Polz, M.F. (2016). Oysters and Vibrios as a Model for
807 Disease Dynamics in Wild Animals. *Trends Microbiol.* 10.1016/j.tim.2016.03.006.
- 808 26. Marbouty, M., Baudry, L., Cournac, A., and Koszul, R. (2017). Scaffolding bacterial
809 genomes and probing host-virus interactions in gut microbiome by proximity ligation
810 (chromosome capture) assay. *Sci Adv* 3, e1602105. 10.1126/sciadv.1602105.
- 811 27. Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragozy, T.,
812 Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009).
813 Comprehensive mapping of long-range interactions reveals folding principles of the
814 human genome. *Science* 326, 289-293. 10.1126/science.1181369.
- 815 28. Hussain, F.A., Dubert, J., Elsherbini, J., Murphy, M., VanInsberghe, D., Arevalo, P.,
816 Kauffman, K., Rodino-Janeiro, B.K., Gavin, H., Gomez, A., et al. (2021). Rapid
817 evolutionary turnover of mobile genetic elements drives bacterial resistance to phages.
818 *Science* 374, 488-492. 10.1126/science.abb1083.
- 819 29. Cali, S., Spoldi, E., Piazzolla, D., Dodd, I.B., Forti, F., Deho, G., and Ghisotti, D.
820 (2004). Bacteriophage P4 Vis protein is needed for prophage excision. *Virology* 322,
821 82-92. 10.1016/j.virol.2004.01.016.
- 822 30. Mirdita, M., Schutze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M.
823 (2022). ColabFold: making protein folding accessible to all. *Nat Methods* 19, 679-682.
824 10.1038/s41592-022-01488-1.

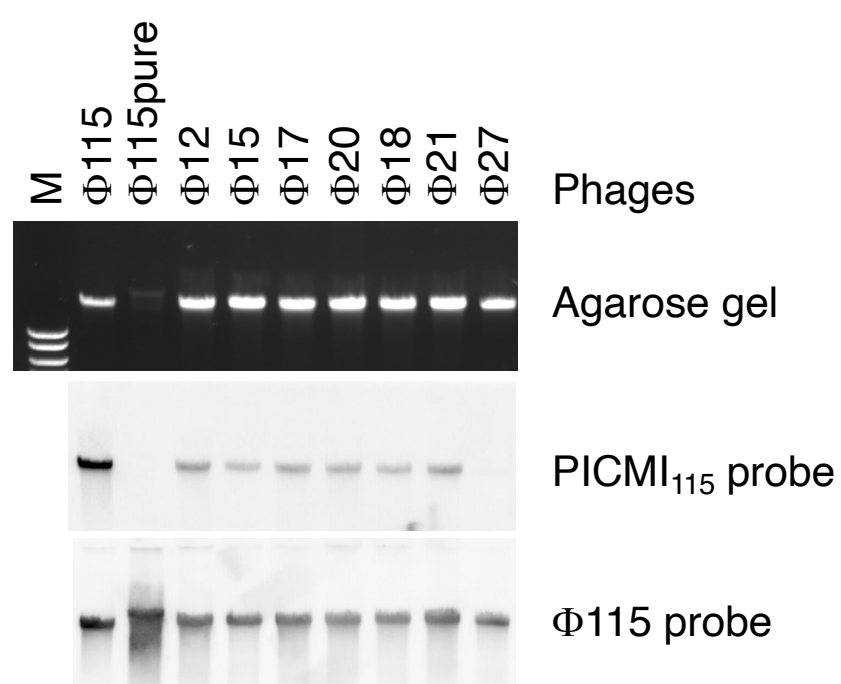
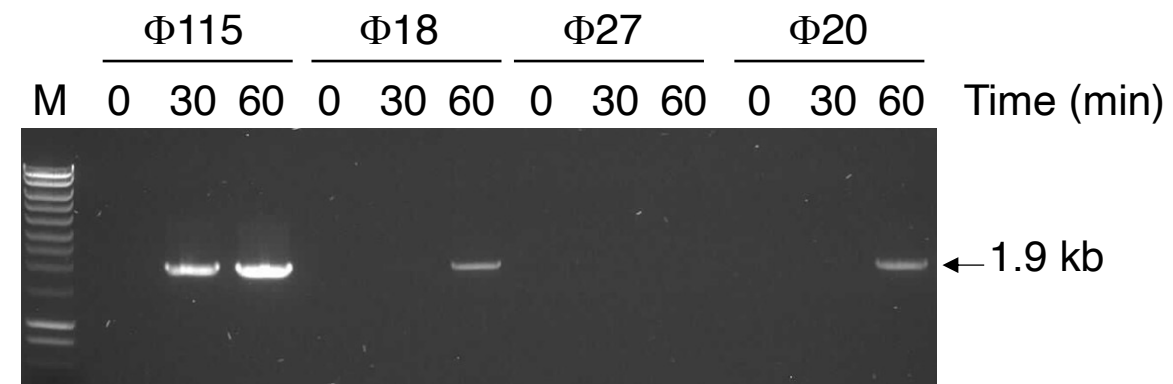
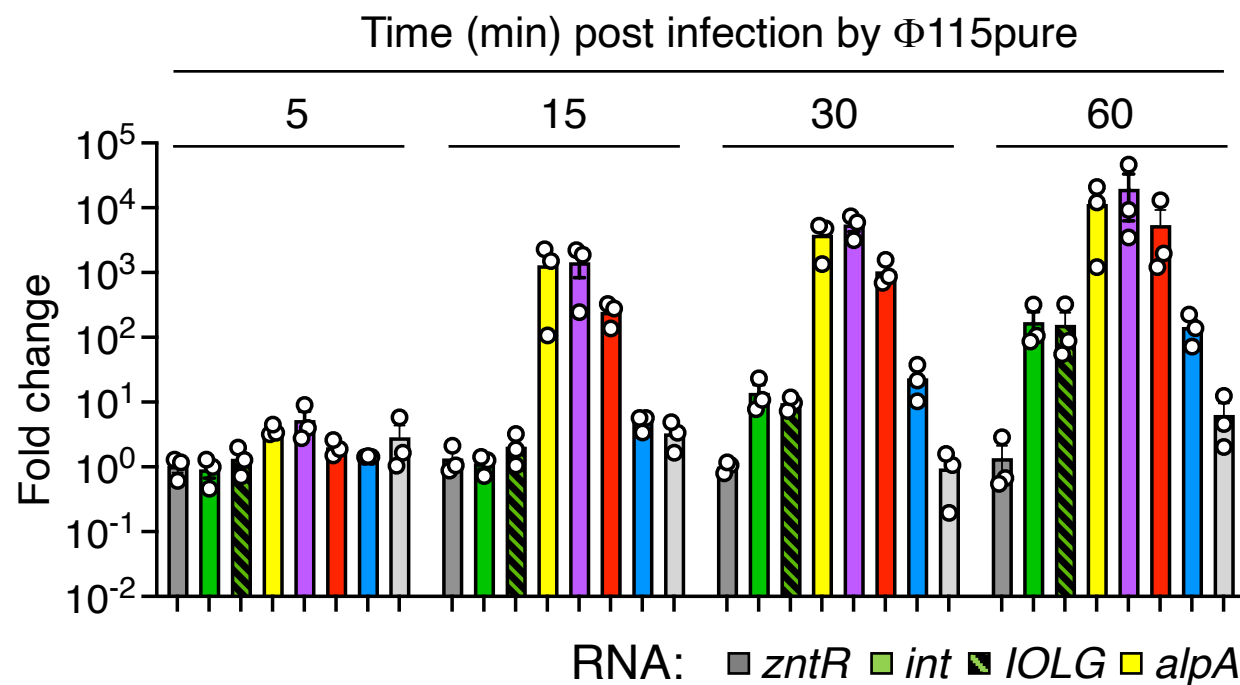
- 825 31. Néron, B., Denise, R., Coluzzi, C., Touchon, M., Rocha, E.P.C., and Abby, S.S. (2023).
826 MacSyFinder v2: Improved modelling and search engine to identify molecular systems
827 in genomes. *Peer Community Journal* 3. 10.24072/pcjournal.250.
828 <https://peercommunityjournal.org/articles/10.24072/pcjournal.250/>.
- 829 32. Kauffman, K.M., Chang, W.K., Brown, J.M., Hussain, F.A., Yang, J., Polz, M.F., and
830 Kelly, L. (2022). Resolving the structure of phage-bacteria interactions in the context
831 of natural diversity. *Nat Commun* 13, 372. 10.1038/s41467-021-27583-z.
- 832 33. Tesson, F., Herve, A., Mordret, E., Touchon, M., d'Humieres, C., Cury, J., and
833 Bernheim, A. (2022). Systematic and quantitative view of the antiviral arsenal of
834 prokaryotes. *Nat Commun* 13, 2561. 10.1038/s41467-022-30269-9.
- 835 34. Moraru, C., Varsani, A., and Kropinski, A.M. (2020). VIRIDIC-A Novel Tool to
836 Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses. *Viruses* 12.
837 10.3390/v12111268.
- 838 35. Le Roux, F., Binesse, J., Saulnier, D., and Mazel, D. (2007). Construction of a *Vibrio*
839 *splendidus* mutant lacking the metalloprotease gene *vsm* by use of a novel
840 counterselectable suicide vector. *Appl Environ Microbiol* 73, 777-784.
- 841 36. Kauffman, K.M., and Polz, M.F. (2018). Streamlining standard bacteriophage methods
842 for higher throughput. *MethodsX* 5, 159-172. 10.1016/j.mex.2018.01.007.
- 843 37. Moreau, P., Cournac, A., Palumbo, G.A., Marbouty, M., Mortaza, S., Thierry, A., Cairo,
844 S., Lavigne, M., Koszul, R., and Neuveut, C. (2018). Tridimensional infiltration of DNA
845 viruses into the host genome shows preferential contact with active chromatin. *Nat*
846 *Commun* 9, 4268. 10.1038/s41467-018-06739-4.
- 847 38. Matthey-Doret, C., Baudry, L., Breuer, A., Montagne, R., Guiglielmoni, N., Scolari, V.,
848 Jean, E., Campeas, A., Chanut, P.H., Oriol, E., et al. (2020). Computer vision for pattern
849 detection in chromosome contact maps. *Nat Commun* 11, 5795. 10.1038/s41467-020-
850 19562-7.
- 851 39. Demarre, G., Guerout, A.M., Matsumoto-Mashimo, C., Rowe-Magnus, D.A., Marliere,
852 P., and Mazel, D. (2005). A new family of mobilizable suicide plasmids based on broad
853 host range R388 plasmid (IncW) and RP4 plasmid (IncPalpha) conjugative machineries
854 and their cognate *Escherichia coli* host strains. *Res Microbiol* 156, 245-255.
- 855 40. Val, M.E., Skovgaard, O., Ducos-Galand, M., Bland, M.J., and Mazel, D. (2012).
856 Genome engineering in *Vibrio cholerae*: a feasible approach to address biological
857 issues. *PLoS genetics* 8, e1002472. 10.1371/journal.pgen.1002472.

- 858 41. Le Roux, F., Davis, B.M., and Waldor, M.K. (2011). Conserved small RNAs govern
859 replication and incompatibility of a diverse new plasmid family from marine bacteria.
860 *Nucleic Acids Res* 39, 1004-1013. 10.1093/nar/gkq852.
- 861 42. Hyman, P., and Abedon, S.T. (2009). Practical methods for determining phage growth
862 parameters. *Methods Mol Biol* 501, 175-202. 10.1007/978-1-60327-164-6_18.
- 863 43. Deorowicz, S., Debudaj-Grabysz, A., and Gudys, A. (2016). FAMSA: Fast and accurate
864 multiple sequence alignment of huge protein families. *Sci Rep* 6, 33964.
865 10.1038/srep33964.
- 866 44. Steinegger, M., and Soding, J. (2018). Clustering huge protein sequence sets in linear
867 time. *Nat Commun* 9, 2542. 10.1038/s41467-018-04964-5.
- 868 45. Cantu, V.A., Salamon, P., Seguritan, V., Redfield, J., Salamon, D., Edwards, R.A., and
869 Segall, A.M. (2020). PhANNs, a fast and accurate tool and web server to classify phage
870 structural proteins. *PLoS Comput Biol* 16, e1007845. 10.1371/journal.pcbi.1007845.
- 871 46. Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, G.A.,
872 Bileschi, M.L., Bork, P., Bridge, A., Colwell, L., et al. (2023). InterPro in 2022. *Nucleic*
873 *Acids Res* 51, D418-D427. 10.1093/nar/gkac993.
- 874 47. van Kempen, M., Kim, S.S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C.L.M.,
875 Soding, J., and Steinegger, M. (2023). Fast and accurate protein structure search with
876 Foldseek. *Nat Biotechnol*. 10.1038/s41587-023-01773-0.
- 877 48. Perrin, A., and Rocha, E.P.C. (2021). PanACoTA: a modular tool for massive microbial
878 comparative genomics. *NAR Genom Bioinform* 3, lqaa106. 10.1093/nargab/lqaa106.
- 879 49. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von
880 Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods
881 for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol* 37, 1530-1534.
882 10.1093/molbev/msaa015.
- 883 50. McNair, K., Zhou, C., Dinsdale, E.A., Souza, B., and Edwards, R.A. (2019).
884 PHANOTATE: a novel approach to gene identification in phage genomes.
885 *Bioinformatics* 35, 4537-4542. 10.1093/bioinformatics/btz265.
- 886 51. Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P.A. (2019). Assembly of long, error-
887 prone reads using repeat graphs. *Nat Biotechnol* 37, 540-546. 10.1038/s41587-019-
888 0072-8.
- 889 52. Vaser, R., and Šikić, M. (2021). Raven: a de novo genome assembler for long reads.
890 bioRxiv, 2020.2008.2007.242461. 10.1101/2020.08.07.242461.

891 53. Chen, Y., Nie, F., Xie, S.Q., Zheng, Y.F., Dai, Q., Bray, T., Wang, Y.X., Xing, J.F.,
892 Huang, Z.J., Wang, D.P., et al. (2021). Efficient assembly of nanopore reads via highly
893 accurate and intact error correction. *Nat Commun* 12, 60. 10.1038/s41467-020-20236-
894 7.
895



A**C****B**

A**B****C****D**