

1 **Title: Molecular and evolutionary basis of O-antigenic polysaccharide**  
2 **driven phage sensitivity in environmental pseudomonads**

3

4 **Running title:** O-antigen-mediated phage and *Pseudomonas* interaction

5

6 Jordan Vacheron<sup>1#</sup>, Clara M. Heiman<sup>1</sup>, Julian R. Garneau<sup>1</sup>, Peter Kupferschmied<sup>1†</sup>, Ronnie de Jonge<sup>2</sup>,  
7 Daniel Garrido-Sanz<sup>1</sup>, Christoph Keel<sup>1#</sup>

8

9

10 <sup>1</sup>: Department of Fundamental Microbiology, University of Lausanne, Lausanne, Switzerland.

11 <sup>2</sup>: Plant-Microbe Interactions, Department of Biology, Science4Life, Utrecht University, Padualaan 8,  
12 Utrecht, 3584 CH, The Netherlands.

13

14 <sup>†</sup>: Present address: Peter Kupferschmied, Federal Office for Agriculture, Swiss Federal Plant Protection  
15 Service, Bern, Switzerland.

16

17 **#Correspondence:** Christoph Keel, [christoph.keel@unil.ch](mailto:christoph.keel@unil.ch), Jordan Vacheron,  
18 [jordan.vacheron@unil.ch](mailto:jordan.vacheron@unil.ch).

19

## 20 **Abstract**

21 *Pseudomonas protegens* CHA0, a bacterial strain able to suppress plant pathogens as well as efficiently  
22 kill lepidopteran pest insects, has been studied as biocontrol agent to prevent ensuing agricultural  
23 damage. However, the success of this method is dependent on the efficient plant colonization by the  
24 bacterial inoculant while it faces competition from the resident microbiota as well as predators such  
25 as bacteriophages. One of these naturally occurring phages,  $\Phi$ GP100, was found to drastically reduce  
26 the abundance of CHA0 once inoculated into plant microcosms, resulting in the loss of plant protection  
27 against a phytopathogen. Here, we investigated the molecular determinants implicated in the  
28 interaction between CHA0 and the phage  $\Phi$ GP100 using a high-density transposon-sequencing  
29 approach. We show that lipopolysaccharide cell surface decorations, specifically the longer OBC3-type  
30 O-antigenic polysaccharide (O-PS, O-antigen) of the two dominant O-PS of CHA0 is essential for the  
31 attachment and infection of  $\Phi$ GP100. Moreover, when exploring the distribution of the OBC3 cluster  
32 in bacterial genomes, we identified several parts of this gene cluster that are conserved in  
33 phylogenetically distant bacteria. Through heterologous complementation, we integrated an OBC3-  
34 type gene copy from a phylogenetically distant bacterium and were able to restore the phage  
35 sensitivity of a CHA0 mutant which lacked the ability to form long O-PS. Finally, we evidence that the  
36 OBC3 gene cluster of CHA0 displays a high genomic plasticity and likely underwent several horizontal  
37 acquisitions and genomic rearrangements. Collectively, this study underlines the complexity of phage-  
38 bacteria interaction and the multifunctional aspect of bacterial cell surface decorations.

39

## 40 **Importance**

41 The application of plant-beneficial microorganisms to protect crop plants is a promising alternative to  
42 the usage of chemicals. However, biocontrol research often faces difficulties to implement this  
43 approach due to the inconsistency of the bacterial inoculant to establish itself within the root  
44 microbiome. Beneficial bacterial inoculants can be decimated by the presence of their natural  
45 predators, notably bacteriophages (also-called phages). Thus, it is important to gain knowledge  
46 regarding the mechanisms behind phage-bacteria interaction to overcome this challenge. Here, we  
47 evidence that the major long O-antigenic polysaccharide (O-PS, O-antigen) of the widely used model  
48 plant-beneficial bacterium *Pseudomonas protegens* CHA0 is the receptor of its natural predator, phage  
49  $\Phi$ GP100. We examined the distribution of the gene cluster directing the synthesis of this O-PS and  
50 identified signatures of horizontal gene acquisitions. Altogether, our study highlights the importance  
51 of bacterial cell surface structure variation in the complex interplay between phages and their  
52 *Pseudomonas* hosts.

53

## 54 Introduction

55 The use of plant-beneficial microorganisms in agriculture is a promising alternative to pesticides  
56 for controlling plant pathogens and pests (1). Nonetheless, this method is dependent on the successful  
57 establishment of a microbial inoculant within the plant environment where it faces competition and  
58 predation from the resident plant microbiota.

59 *Pseudomonas protegens* bacteria, which are represented by the type and model strain CHA0 (2),  
60 are well-known plant root colonizers that can provide various plant-beneficial functions, which include,  
61 amongst others, suppression of diseases caused by phytopathogenic fungi and oomycetes, nutrient  
62 mobilization, and stimulation of plant growth and defenses (3). *P. protegens* bacteria are also capable  
63 of efficiently infecting and killing lepidopteran pest insects that cause serious damage to agricultural  
64 crops (1, 4–6). All these plant-beneficial traits make these bacteria promising candidates in assays  
65 aimed at improving plant health and performance in field conditions (7, 8). However, the transition  
66 from a controlled laboratory environment to the field constitutes a considerable leap into a system  
67 influenced by a plethora of biotic and abiotic factors. Notably, the resident root microbiota can be  
68 perceived as a biological barrier (9), which can interfere with the establishment of the bacterial  
69 inoculant. Indeed, inoculants will face competition from other members of the host microbiota for  
70 nutrients and space. Moreover, the root microbiota also harbors predators of bacteria, which will  
71 strongly influence the fate of the inoculant, and thereby may reduce its plant-beneficial activities.  
72 Amongst these predators, grazers such as protozoa and nematodes (10, 11) as well as bacteriophages  
73 (12) are known to shape the structure of the bacterial community within the root microbiota.

74 Bacteriophages, or briefly phages, are a type of virus that infects bacteria and uses this host as a  
75 cellular replication factory to maintain their viral population within a specific environment. Phages  
76 exhibit an impressive diversity, reflecting the variety of the bacterial microbiota and the environments  
77 they interact with (for a review see (13)). They play crucial roles in different ecosystems by controlling  
78 the abundance and diversity of bacterial populations as well as maintaining a dynamic gene flow within

79 bacterial communities through horizontal gene transfers (14–17). Moreover, the interest in phages as  
80 biocontrol agents is rising and provides promising results for the protection of plants against bacterial  
81 diseases (18, 19). However, naturally occurring phages may also target bacterial inoculants used to  
82 benefit agricultural crops. Indeed, a lytic bacteriophage that targets the plant-beneficial *P. protegens*  
83 strain CHA0 was isolated from the rhizosphere of cucumber plants (20, 21). The presence of this phage,  
84 named  $\Phi$ GP100, drastically reduced the abundance of CHA0 in the rhizosphere, resulting in the loss of  
85 the plant protection effect against of a root pathogenic oomycete (20).

86 Phages generally target specific cell surface structures for initial attachment to their host cell, such  
87 as exopolysaccharides, pilus or flagellum proteins as well as lipopolysaccharides (LPS) for Gram-  
88 negative bacteria (22–24). LPS are the principal components of the outer membrane of Gram-negative  
89 bacteria and generally are composed of three parts, the lipid A, the core oligosaccharide (core-OS),  
90 and the O-antigenic polysaccharide (O-PS, O-antigen) (25–27). The lipid A, which is anchored in the  
91 outer membrane, is bound to the core-OS made of different sugar residues. The lipid A-core-OS  
92 constitutes the conserved structure to which are ligated the highly variable O-PS, consisting of an  
93 assembly of repetitive glycans that are subject to various modifications. Due to their exposure at the  
94 cell surface, the O-PS components are of high importance in bacteria-host and intermicrobial  
95 interactions and a virulence factor in many pathogenic bacteria (25–27). O-PS significantly contribute  
96 to the insect pathogenicity and competitiveness of *P. protegens* strains (28, 29). *P. protegens* CHA0  
97 harbors four O-PS gene clusters specifying the formation of OSA, OBC1, OBC2 and OBC3, of which OSA  
98 and OBC3, a short and a long O-PS respectively, constitute the dominant LPS decorations of the strain  
99 (28).

100 To better understand the interaction between a root-colonizing beneficial bacterial inoculant (*P.*  
101 *protegens* CHA0) and its naturally occurring phage predator ( $\Phi$ GP100), we identified potential  
102 bacterial genetic determinants involved in the susceptibility towards the phage  $\Phi$ GP100, using a high-  
103 density transposon-sequencing (Tn-seq) approach. Through a reverse genetics approach, we provide

104 evidence that the OBC3-type O-PS, which constitutes the major long O-PS of *P. protegens* CHA0, is the  
105 receptor of the lytic phage  $\Phi$ GP100. We then examined the distribution of the OBC3 gene cluster in  
106 bacterial genomes and identified several parts of this gene cluster that are conserved in  
107 phylogenetically distant bacteria. We were able to restore the phage sensitivity of a CHA0 mutant  
108 depleted for the formation of the long O-PS by heterologous complementation, integrating a copy of  
109 a related OBC3-type gene from one of these phylogenetically distant bacteria. Finally, we evidence  
110 that the OBC3 gene cluster of CHA0 displays a high genomic plasticity as well as signatures of sequential  
111 acquisitions by horizontal gene transfer.

112

## 113 **Results & discussion**

### 114 **Identification of genes required for $\Phi$ GP100 phage resistance using a transposon** 115 **sequencing approach.**

116 To identify genes contributing to the sensitivity of *P. protegens* CHA0 towards the phage  $\Phi$ GP100,  
117 we generated a saturated transposon mutant library using the Tn5 transposon carried by the pRL27  
118 plasmid (**Table S1**, (30)). We obtained approximately 500'000 transposon insertions distributed  
119 throughout the genome of *P. protegens* CHA0 (i.e., around 70 transposon insertions every 1'000 bp)  
120 (**Fig. S1A**, **Table S2**). More than 93 % of the genes were affected by the insertion of the transposon  
121 with on average between 67 to 84 transposon hits per gene (**Table S2**). The Tn-seq results from the  
122 different conditions were highly reproducible with a Pearson correlation coefficient of 0.98 (**Fig. S2**).

123 To detect a wide range of genes involved in the sensitivity of CHA0 to the phage  $\Phi$ GP100, we  
124 exposed the Tn-mutant library to a gradually increasing concentration of phage particles. After an  
125 overnight exposure, we recovered the enriched Tn-mutants that had incorporated the Tn5 transposon  
126 in crucial genes responsible for the sensitivity of this strain towards the phage  $\Phi$ GP100, and proceeded  
127 with DNA extraction and sequencing to identify these genes (**Fig. 1A**, **Fig. S1B-D**). For the three  
128 different multiplicity of infection tested (MOI = 1; MOI = 10; MOI = 100), we obtained 3'467, 4'029 and

129 4'074 genes respectively, that were significantly less represented in the Tn-mutant library following  
130 the exposure to the phage  $\Phi$ GP100 (**Fig. 1B** and **Fig. S3**). The genes identified encompass essential  
131 genes for the development and growth of the bacterial host, as well as genes for which the disruption  
132 by the transposon might increase the sensitivity of the bacterium to the phage. Applying a gradually  
133 increasing concentration of phages allowed a deeper look into the phage infection process by focusing  
134 on bacterial genes potentially needed for the phage replication inside the cell. Indeed, by comparing  
135 the different phage concentrations, i.e., a MOI of 100 vs. 10 and a MOI of 100 vs. 1, we identified 30  
136 and 13 genes, respectively, related to the metabolism of the bacterial cell and the transcription  
137 machinery, including transcriptional and translational regulators (**Fig. S4, Dataset 1**). Conversely, a  
138 total of 385, 210 and 232 genes were detected as potentially involved in the bacterial sensitivity  
139 towards the phage at an MOI of 1, 10 and 100, respectively (**Fig. 1B** and **Fig. S3**). Globally, Tn-mutants  
140 that were the most enriched following phage infections were highly similar for all the different MOI  
141 applied (**Fig. S1**). Although most of the identified genes were assigned as unknown function according  
142 to the EggNOG functional annotation, 13 % were associated with the cell wall and membrane  
143 biosynthesis category (**Fig. 1B** and **Fig. S3**). Most of the genes from this EggNOG functional category  
144 participate in the different steps of the biosynthesis of LPS (**Dataset 1**). Indeed, the three most  
145 abundant Tn-mutant categories (**Fig. 1A, dataset 1**) possessed transposon insertions in the *lapA* gene,  
146 encoding an adhesin involved in biofilm attachment (31), in *galU*, required to form an operational LPS  
147 core-oligosaccharide in *Pseudomonas aeruginosa* (32, 33) or in the OBC3 gene cluster, specifying the  
148 formation of the major long O-PS in *P. protegens* CHA0 (28).

149

### 150 **The long O-PS of CHA0 is mandatory for the infection by $\Phi$ GP100.**

151 Following the results from the Tn-seq experiment, we decided to better characterize the role of  
152 the LPS of *P. protegens* CHA0 in the interaction with the phage  $\Phi$ GP100. The Tn-seq experiment  
153 highlighted the importance of a complete and operational core-OS for the phage infection process (**Fig.**  
154 **1C**). Upon  $\Phi$ GP100 phage exposure, Tn insertions were identified in several core LPS genes, notably in

155 *galU* and *algC* (**Fig. 1C**). The gene *galU* encodes an UTP-glucose-1P uridylyltransferase, which allows  
156 the production of UDP-galactose and UDP-glucose (**Fig. 1D**) that are essential sugar residues  
157 composing the core-OS in *P. aeruginosa* (32). Indeed, a *galU* mutant of *P. aeruginosa* was no longer  
158 able to bind O-PS as the core-OS appeared truncated (33). The gene *algC* encodes a  
159 phosphomannomutase / phosphoglucomutase, a key enzyme that catalyzes the transformation of  
160 glucose-6P into glucose-1P and mannose-6P to mannose-1P (**Fig. 1D**), which also are essential  
161 components of the core-OS (34). Indeed, an *algC* mutant of *P. aeruginosa* harbored a truncated core-  
162 OS unable to attach any O-PS (34), resulting in resistance towards two different phages (35). Other  
163 genes involved in the core-OS formation were identified by the Tn-seq experiment such as *wbpM* and  
164 *wbpL* (**Fig. 1C**). These two genes are located in the OSA cluster in the genomes of *P. aeruginosa* and *P.*  
165 *protegens* CHA0 and are conserved among various other *Pseudomonas* genomes (28). The  
166 glycosyltransferase WbpL initiates the synthesis of the O-PS by transferring a N-acetyl sugar to the lipid  
167 carrier in the cytoplasm of the bacterium (36–38). The deletion of *wbpL* in CHA0 and in other  
168 *Pseudomonas* results in strains unable to bind O-PS (28, 39). Finally, genes involved in the metabolism  
169 of amino acids and raw sugars were impacted, likely not only due to their essential role in bacterial  
170 growth but also as they supply the raw sugar subunits for the core-OS and the O-PS biosynthesis (**Fig.**  
171 **1D**).

172 *P. protegens* CHA0 harbors four O-PS gene clusters specifying the formation of OSA, OBC1, OBC2  
173 and OBC3, of which OSA and OBC3, a short and a long O-PS, respectively, constitute the dominant LPS  
174 decorations of the strain (28). In the present study, only Tn-insertions within the OBC3 gene cluster  
175 were enriched upon exposure to the phage  $\Phi$ GP100 (**Fig. 1A and 1C**). Then, 13 out of the 19 genes  
176 found in the OBC3 gene cluster showed a significantly higher number of transposon insertions  
177 compared to the control condition (**Fig. 1C**). To confirm the Tn-seq results, we tested deletion mutants  
178 impaired in the production of the different O-PS and confronted them with the phage (**Fig. 2A and Fig.**  
179 **S5**). A *wbpL* mutant of CHA0 which lacks both OSA and OBC3 (28), was fully resistant to the phage  
180  $\Phi$ GP100 (**Fig. 2A**). Neither the OBC1 nor the OBC2 mutants showed a different sensitivity pattern

181 towards the phage compared to the wild type CHA0 (**Fig. 2A**), supporting the Tn-seq data. However,  
182 the OBC3 mutant, which lacks a significant part of the OBC3 cluster (part I and part II of the cluster,  
183 **Fig. 1C**) became fully resistant to the phage (**Fig. 2A**), strongly suggesting that the  $\Phi$ GP100 phage  
184 targets the long O-PS displayed at the cell surface of CHA0. Phage resistance could also be achieved by  
185 deleting individual genes of the OBC3 cluster, including *fcl*, *fcl2*, PPRCHA0\_1957, PPRCHA0\_1962, and  
186 *galE* (**Fig. S5**). Although most of the genes belonging to the OBC3 gene cluster encode important  
187 enzymatic functions necessary for the production of the long O-PS, the gene *oafA* encoding an O-  
188 antigen acetylase (28, 40), was not detected as an important gene involved in the sensitivity of CHA0  
189 towards the phage  $\Phi$ GP100 by the Tn-seq analysis (**Fig. 1C**). Furthermore, a deletion mutant of *oafA*  
190 was still sensitive to the phage (**Fig. S5**), further corroborating the Tn-Sea data. Even though the  
191 acetylation of the long O-PS does not seem to play a role during the infection of *P. protegens* CHA0 by  
192 the phage  $\Phi$ GP100, it could still be important for the ecology of this bacterium, for example during the  
193 interaction with its eukaryote hosts such as plants and pest insects as suggested by a study on the  
194 experimental evolution of bacterial root colonization traits (40).

195 Interestingly, the plaques formed by the phage  $\Phi$ GP100 when deposited onto CHA0 during spot  
196 assays exhibited a clear center indicative of a complete bacterial lysis, surrounded by a turbid halo (**Fig.**  
197 **2A**). The deletion of the flippase-encoding gene *wzx*, which leads to the loss of the short OSA-type O-  
198 PS but maintains the long OBC3 O-PS, produced a mutant that did not display a clear center but the  
199 entire plaques remained turbid (**Fig. 2A**). The presence of such turbid zones is typical of phages  
200 exhibiting a depolymerase activity (41), that could be, in the case of the phage  $\Phi$ GP100, attributed to  
201 the degradation of the long OBC3 O-PS of CHA0. The absence of the clear center when the phage  
202 infects  $\Delta$ *wzx* was not related to either a lower phage adhesion on the cell surface (**Fig. 2B**) or a decrease  
203 of the phage infectivity (**Fig. 2C**). The depolymerase activity is frequently observed in phages of the  
204 *Podoviridae* family (42, 43). Although we previously thought that the  $\Phi$ GP100 phage belongs to the  
205 *Podoviridae* family considering its shape under electron microscope (20), the new genomic analysis we  
206 performed here assigned it to the *Zobellviridae* family (44), displaying, however, low similarity with

207 other closely related phages from this family (**Fig. 3A, Table S3**). Furthermore, we performed a new  
208 annotation of the phage genome (**Fig. 3B, Table S4**) to detect potential genes encoding depolymerases  
209 that could be responsible for the halo formation (**Fig. 2A**). We identified a gene predicted to encode a  
210 protein containing a pectate lyase domain (**Fig. 3B, Fig. S6**). The protein structure of this pectate lyase  
211 was predicted and displays structural similarities with a pectate lyase encoded in the genome of the  
212 phage P1 (**Fig. S6**) where the depolymerase activity had been demonstrated to be responsible for the  
213 formation of a halo when the phage infects *Acinetobacter* strains (45).

214 Based on these results, we propose the following model for the infection of CHA0 by the phage  
215  $\Phi$ GP100 (**Fig. 3C**). The phage recognizes the long O-PS (OBC3) at the surface of CHA0 (and other *P.*  
216 *protegens* group strains displaying OBC3, see below) as a primary receptor for the infection process.  
217 Then we suggest that the phage deploys the depolymerase activity to degrade the long O-PS to reach  
218 the bacterial outer membrane and injects its genetic material inside the bacterial cell (see plaques for  
219 the *wzx* mutant in **Fig. 2A**). However, the interaction between this depolymerase with the short O-PS  
220 (OSA type) remains to be investigated as well as the localization of the depolymerase on the phage  
221 (i.e., on the tail fibers, tail spike or the neck).

222

### 223 **Distribution of the OBC3 gene cluster and host range of the $\Phi$ GP100 phage.**

224 Next, we examined the distribution of the OBC3 gene cluster as well as the OSA gene cluster  
225 within bacterial genomes from the NCBI database to see whether the presence of the OBC3 gene  
226 cluster or parts of it can be linked to the phage sensitivity (**Fig. 4**). We split this gene cluster into three  
227 parts according to the arrangement of the genes within the cluster (**Fig. 1C**). We looked at the  
228 presence/absence of the three parts composing the OBC3 gene cluster by Blastp, applying a threshold  
229 of 60 % of amino acid identity on at least 70 % of the amino acid sequence length. In parallel, we  
230 determined the activity spectrum of the phage  $\Phi$ GP100 against different *Pseudomonas* strains as well  
231 as against bacteria where we detected the entire or parts of the OBC3 gene cluster.

232 Half of the strains belonging to the *P. protegens* species investigated in this study (CHA0, Pf1,  
233 PGNR1, BRIP and Cab57) possess the OSA gene cluster and the OBC3 gene cluster. Moreover, the OBC3  
234 gene cluster is located in the same genomic region of these bacterial genomes, i.e., between a gene  
235 encoding an aerotaxis receptor (*aer\_2*) and a gene encoding a CPBP family intramembrane  
236 metalloprotease (**Fig. 4**). Although these bacteria harbor the OSA gene cluster along with the OBC3  
237 gene cluster, BRIP and Cab57 exhibited a reduced sensitivity to the phage  $\Phi$ GP100 compared to CHA0  
238 (**Fig. 4** and **Fig. S7**). Since different strategies are deployed by bacteria to evade phage infection (46),  
239 the difference of phage sensitivity between CHA0 and BRIP / Cab57, despite harboring the same OBC3  
240 O-PS (28), may be based on the presence of intracellular anti-phage defense systems such as CRISPR-  
241 cas systems (46) that could limit the infection by the phage  $\Phi$ GP100 (46). CMR5c is also less sensitive  
242 to the phage compared to CHA0 (**Fig. 4** and **Fig. S7**). However, in this strain only the part III of the OBC3  
243 gene cluster was detected at the same chromosomal locus as in CHA0, while the two other parts are  
244 located in another genomic region and are not entirely complete (**Fig. 4**). On the contrary, we detected  
245 the parts II and III of the OBC3 cluster of *Pseudomonas* sp. R76 in the same genomic region as in CHA0  
246 (**Fig. 4**). Furthermore, several genes from the different parts of the OBC3 gene cluster were identified  
247 in genetically distant bacteria, belonging to the Alphaproteobacteria (*Rhizobium etli* CFN42 and  
248 *Rhizobium leucaenae* USDA9039) or the Betaproteobacteria (*Acidovorax avenae* subsp. *avenae* AA99,  
249 *Herbaspirillum rubrisubalbicans* M1 and *Azoarcus* sp. BH72) (**Fig. 4, Table S5**). All these distantly  
250 related bacteria as well as *Pseudomonas* sp. R76 were resistant to the phage  $\Phi$ GP100 (**Fig. 4**). The  
251 presence of genes from the OBC3 cluster or parts of it within the genome of distantly related bacteria  
252 raises the question of the origin of this gene cluster in the phage-sensitive *Pseudomonas* bacteria.

253

254 **Heterologous complementation of a CHA0 mutant lacking the long O-PS restores phage**  
255 **sensitivity.**

256 The OBC3 gene cluster of *P. protegens* CHA0 harbors several genes of the part I with relatively  
257 high sequence identities (67-75 %) to genes of the O-PS biosynthesis locus of *Rhizobium etli* CFN42

258 **(Fig. 4 and Fig. 5)**. Three of these genes encode enzymes (a GDP-mannose 4,6-dehydratase and a GDP-  
259 L-fucose synthetase) and a glycosyltransferase (the fucosyltransferase WreE) that are sufficient for the  
260 synthesis of GDP-L-fucose from GDP-D-mannose and the covalent binding of the sugar moiety to O-PS  
261 repeat units in *R. etli* CFN42 (**Fig. 5**; (47)). The homologous genes in CHA0 were at least 67 % identical  
262 to the *Rhizobium* genes and seemed to also be organized as an operon, indicating that CHA0 is capable  
263 of synthesizing GDP-L-fucose (**Fig. 5**). As previously reported, the deletion of *fcl* (encoding a putative  
264 GDP-L-fucose synthetase) in CHA0 resulted in the loss of long OBC3-type O-PS (**Fig. 5**; (28)), while re-  
265 introduction of *fcl* into the genome of  $\Delta fcl$  rescued the biosynthesis of the long O-PS (**Fig. 5**).  
266 Furthermore, *cis* complementation of the same mutant strain with *fcl* from *R. etli* CFN42 similarly  
267 rescued the production of long O-PS, although this resulting O-PS appeared to have lower molecular  
268 weights compared to the wild type. Nevertheless, it can be assumed that the *fcl* gene from CHA0  
269 encodes a GDP-L-fucose synthetase as in *R. etli*, since glycosyltransferases generally display strict  
270 substrate specificity (48). The long O-PS generated by the  $\Delta fcl$  mutant of CHA0 expressing the *fcl* gene  
271 of *R. etli* CFN42 restored the phage susceptibility of the strain (**Fig. 5**), indicating that acquiring genes  
272 involved in the O-PS biosynthesis from a phylogenetically distant bacterium can be functional and  
273 exploited by the recipient strain.

274

275 **The OBC3 gene cluster of CHA0 possesses a high genetic plasticity and displays signatures of**  
276 **horizontal gene transfer acquisitions.**

277 The conserved genomic location in *Pseudomonas* genomes where the OBC3 gene cluster is  
278 integrated could be a propitious region for acquisition of LPS gene clusters. Indeed, another O-PS gene  
279 cluster is inserted in the same conserved genomic region in strain *Pseudomonas* sp. AU20219 (**Fig. 4**).  
280 Moreover, a transposable element (TE) is inserted upstream the putative O-antigen acetylase (*oafA*),  
281 within the part III of the OBC3 gene cluster of *P. protegens* CHA0 (**Fig. 6A**). This TE is assigned as an  
282 insertion sequence (IS) belonging to the IS3 family according to the ISFinder database (49) and flanked  
283 by repetitive palindromic sequences (**Fig. 6A**). Interestingly, similar ISs (i.e., belonging to the IS3 family)

284 were detected within the part III of the OBC3 gene cluster in other bacterial genomes, notably of *P.*  
285 *protegens* subgroup strains (**Fig. 4**). The presence of these ISs suggests that the function associated to  
286 these TEs is conserved. ISs are well known drivers of bacterial genome evolution through their  
287 insertions at crucial genetic locations (50, 51). They can modulate gene expression by inserting  
288 themselves in a promoter region or in the middle of a gene (52). For instance, the presence of an IS  
289 within the gene cluster involved in the lipid A formation of *Acinetobacter baumannii* provoked the  
290 complete loss of LPS production and *de facto* colistin resistance (53). Thus, as in the OBC3 gene cluster  
291 of CHA0, the IS is located upstream of the *oafA* gene, it may affect the transcription of the latter. We  
292 also identified nine genomic portions flanked by inverted repeat sequences (IRs) that are exclusively  
293 detected within the OBC3 gene cluster (**Fig. 6A, Table S6**). The presence of IRs has been proposed to  
294 reflect genomic instability (54). These genetic elements can form special DNA conformations such as  
295 hairpin-like or cruciform-like DNA structures that can facilitate genome rearrangements and mutations  
296 (55). IRs were detected within a LPS gene cluster of a *Xanthomonas* strain and were proposed to either  
297 have a role in gene regulation or in recombination during horizontal gene transfer (56). Here, it could  
298 be hypothesized that the presence of IRs within the OBC3 gene cluster permitted the genomic  
299 rearrangement of the different parts of this cluster and might have facilitated their mobilization.

300 Parametric methods can be used to infer horizontal gene transfers by looking at genomic regions  
301 that differ from the genomic average in terms, for example, of GC content (57). CHA0 has a relatively  
302 high average GC content of 63.4 % while the GC content of the OBC3 gene cluster is 51 %. Moreover,  
303 the OBC3 gene cluster is also predicted to be a genomic island by IslandViewer (58) on the basis of  
304 codon usage bias (SIGI-HMM method, (59), **Fig. S8**). The low GC content along with a different codon  
305 usage are sufficient to infer that the presence of the OBC3 gene cluster within the genome of CHA0  
306 originates from horizontal gene transfer. Remarkably, most of the O-PS gene clusters investigated in  
307 this study present lower GC contents than the average of the genome in which they are located (**Fig.**  
308 **S9**). The same observation was made within other bacterial models (60). Indeed, a low GC content  
309 could be detected in all LPS clusters of *Vibrio* strains (61). The exchange and recombination of these

310 LPS clusters led to new bacterial serogroups benefitting the adaptation of these strains to different  
311 environments (61). Low GC contents were also reported in the LPS clusters of several strains of the  
312 plant-pathogen *Xanthomonas* (62). The genetic mobilization of LPS gene clusters was also observed in  
313 *Herbaspirillum*, where clinical strains had acquired an increased number of genetic islands including  
314 LPS biosynthesis gene clusters compared to environmental strains as a possible adaptation to the  
315 human host (63). The genetic diversity of LPS in *P. aeruginosa* has also been assessed and unveiled a  
316 high heterogeneity within the O-PS biosynthesis gene clusters that is driven by environmental selective  
317 pressures such as bacteriophages (64).

318

### 319 **Evolutionary origins of the OBC3 gene cluster**

320 Phylogenetic methods are also used to evidence horizontal gene transfer based on the difference  
321 of tree topology between genes and species trees. Here, we compared the phylogenetic tree based on  
322 the deduced amino acid sequences of the OBC3 gene cluster to the species tree to identify any  
323 incongruence that would allow the inference of horizontal gene transfer. We generated different trees  
324 based on the organization of this gene cluster as mentioned above (part I, part II and part III). Only  
325 amino acid sequences that are conserved in relevant genomes were concatenated according to the  
326 part of the OBC3 cluster they belong to, to obtain the most informative picture of the evolutionary  
327 history of this gene cluster (**Fig. 6B**). Different tree topologies were observed compared to the species  
328 tree for the parts I and II of the OBC3 gene cluster (**Fig. 6B** and **Fig. S10**). *Pseudomonas* sp. CMR5c,  
329 which belongs to the *P. protegens* subgroup, was found to be distant to Cab57 and CHA0 in the tree  
330 built with the amino acid sequences of the part I, and to be clearly distant from the other *Pseudomonas*  
331 for the part II. It is likely that *Pseudomonas* sp. CMR5c could have acquired these two parts  
332 independently, contrarily to other *P. protegens* strains (i.e., not by a vertical transmission from a  
333 common ancestor). Nevertheless, it is evident from the tree topology that the part II of the OBC3  
334 cluster of CMR5c was not acquired from another *Pseudomonas* strain (**Fig. 6B** and **Fig. S10**). Moreover,  
335 the different genomic locations of the parts I and II in the genome of CMR5c (**Fig. 3**) support the

336 hypothesis of their independent acquisition compared to other *P. protegens* strains possessing the  
337 OBC3 gene cluster. On the contrary, the tree obtained from the amino acid sequences of the part III is  
338 consistent with the species phylogeny. The observation that this part of the OBC3 gene cluster is  
339 located in the same genomic region in the *Pseudomonas* genomes suggests that the acquisition of the  
340 part III may be older than that of the other two parts.

341

## 342 **Conclusion**

343 This study provides evidence that the lytic phage  $\Phi$ GP100 uses the long glycan chains of the OBC3 O-  
344 PS as a receptor for CHA0 infection. The fact (i) that several OBC3-like gene clusters could be identified  
345 in various *Pseudomonas* strains and phylogenetically more distant bacteria such as *Rhizobium* species,  
346 (ii) that the GC content differs strongly from the average GC content of the different genomes  
347 investigated in this study, and (iii) that mobile genetic markers such as transposases, IS, IR sequences  
348 are abundant in the OBC3 genomic region, indicates that this O-PS gene cluster or parts of it have been  
349 exchanged via horizontal gene transfer. This potential genetic trade could have resulted in the high  
350 genetic and phenotypic diversity observed in the investigated bacterial LPS profiles and more generally  
351 could be a fundamental factor influencing the interactions of individuals constituting the plant  
352 microbiome with phages, and presumably also with the plant host. Our results underline the  
353 complexity and the multifunctional aspect of bacterial cell surface decorations during bacterial  
354 interactions and emphasize a Cornelian dilemma of being at the mercy of predators or losing potential  
355 LPS-associated beneficial interactions. Our study provides a further illustration of how phages can be  
356 considered as a major driving force of bacterial ecology and evolution.

357

## 358 **Materials and Methods**

### 359 **Bacterial strains, phage, plasmids, media, and culture conditions**

360 All strains and plasmids used in this study are listed in **Table S1, S5 and S7**. Bacterial strains were  
361 routinely cultured on nutrient agar (NA), in nutrient yeast broth (NYB) or in lysogeny broth (LB)  
362 supplemented with the appropriate antibiotics when needed (ampicillin, 100  $\mu\text{g mL}^{-1}$ ;  
363 chloramphenicol, 10  $\mu\text{g mL}^{-1}$ ; kanamycin, 25  $\mu\text{g mL}^{-1}$ ; gentamycin, 10  $\mu\text{g mL}^{-1}$ ). *Pseudomonas* strains  
364 were grown at 25 °C if not mentioned otherwise, while *E. coli* was cultured at 37 °C. Bacteriophage  
365  $\Phi\text{GP100}$  lysate stock were prepared as described previously (20). Briefly, the phage was added to NYB  
366 medium containing CHA0 (optical density at 600 nm ( $\text{OD}_{600\text{nm}}$ ) = 1) and incubated overnight at 25 °C.  
367 The resulting lysate was centrifuged at 2000  $g$  for 5 min and the supernatant was filtered (Millipore  
368 filters, pore size 0.22  $\mu\text{m}$ ). Phages were precipitated using 10 % of polyethylene glycol (PEG 8000) and  
369 centrifuged (15,000  $g$ , 30 min). The pellet containing phages was resuspended in SM buffer (100 mM  
370 NaCl, 8 mM  $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$ , 50 mM Tris-Cl, pH 7.5). Titration of this purified phage stock suspension (500  
371 mL) was performed by the double layer agar technique and led to  $10^9$  PFU  $\text{mL}^{-1}$  (Plaque Forming Unit  
372 per milliliter). The phage stock suspension was tested for bacterial contamination by plating 200  $\mu\text{L}$  on  
373 NA plate. Half of the phage stock suspension was kept at 4 °C for up to 4 months while the other one  
374 was aliquoted and stored in glycerol (50 % vol/vol final) at -80 °C for long-term storage.

375

### 376 **Transposon mutant library construction**

377 A high density Tn5 mutant library of *P. protegens* CHA0 was prepared following a protocol described  
378 previously for the strain *P. protegens* Pf-5 (29). Briefly, competent cells of CHA0 were electroporated  
379 with the plasmid pRL27 containing the mini-Tn5 transposon (30) and immediately rescued with 1 mL  
380 of super optimal broth with catabolite repression medium and incubated at 35 °C. After 2.5 h, the  
381 obtained bacterial suspension was serially diluted and plated onto ten NA plates supplemented with  
382 25  $\mu\text{g mL}^{-1}$  of kanamycin and incubated at 25 °C. Following 24 h of incubation, approximately 600,000

383 Km-resistant colonies were recovered and placed into a sterile 0.8 % NaCl solution. The bacterial  
384 suspension was homogenized by vortexing and centrifuged to concentrate the full Tn5-library into a  
385 final volume of 5 mL of which aliquots of 1 mL were stored with glycerol (50 % vol/vol final) at -80 °C  
386 for subsequent use.

387

## 388 **Selection of phage-resistant Tn5-mutants and sequencing**

389 Aliquots of 150 µL of the Tn5 mutant library were added to 8 mL of NYB medium and incubated for 8  
390 h at 25 °C. The OD<sub>600nm</sub> of this culture was adjusted to 0.01 ( $2 \times 10^6$  cells mL<sup>-1</sup>) in four tubes containing  
391 10 mL of fresh NYB. Three of these tubes were inoculated with ΦGP100 phage stock solution at  
392 different concentrations;  $10^6$ ,  $10^7$  and  $10^8$  PFU mL<sup>-1</sup> (i.e., at MOI of 1, 10 and 100, respectively). The  
393 last tube was inoculated with SM buffer as control. The four tubes were then incubated at 25 °C for 10  
394 h. After incubation, the cells were centrifuged and washed three times with sterile water to remove  
395 any traces of DNA from lysed cells. Genomic DNA was then extracted from the cell pellets using the  
396 MagAttract® HMW DNA kit (Qiagen). The library preparations and paired-end sequencing (HiSeq2500,  
397 Illumina) were performed as detailed previously (29). Three independent biological replicates were  
398 performed.

399

## 400 **Sequence processing and statistical analysis**

401 The reads were processed and analyzed as was previously described (29). Reads were preprocessed  
402 using cutadapt (v.2.3, (65)) and reaper (v.15-065, (66)) and mapped to the genome of *Pseudomonas*  
403 *protegens* CHA0 (LS999205.1) using bwa (v.0.7.17, (67)). FeatureCounts (v.1.6.0) was used to  
404 summarize the number of read counts per gene locus and TRANSIT (v.2.5.2) was used for statistical  
405 analysis. Insert site counts were normalized using the TTR method (68). The comparative analysis for  
406 determining the conditional essentiality of genes was performed using the resampling method (10,000  
407 permutations, (68)).

408

## 409 **Genomic analysis**

410 Functional annotation of genes into clusters of orthologous groups (COGs) from the *P. protegens* CHA0  
411 genome was performed using eggNOG-mapper (v2) (69). The synteny and the detection of sequence  
412 homology of the OBC3 gene cluster was obtained by cblaster (70, 71) with the homology threshold set  
413 at 60 % of amino acid sequence identity on at least 70 % of the sequence length. Inverted repeats in  
414 the OBC3 gene cluster of CHA0 were detected by GeneQuest (DNASTAR's software, v15.3). The GC  
415 contents were calculated with a window range of 500 bp using Artemis (72). The functional annotation  
416 of the phage  $\Phi$ GP100 was updated using PROKKA (73) with a priority annotation using the PHASTER  
417 protein database (version Dec 22, 2020; (74)). The similarity between the different phage genomes  
418 was calculated and visualized using Circoletto (75). The protein structure of the pectate lyase was  
419 predicted using AlphaFold2 (76). Protein structure comparison was performed using Dali server (77)  
420 and protein domain were detected by interrogating the Conserved Domain Database from NCBI (78).

421

## 422 **Phylogenies**

423 The maximum-likelihood (ML) phylogenetic tree of the 34 bacterial genomes (**Table S5**) was built based  
424 on single-copy proteins, as previously described (79). Briefly, proteomes were analyzed with  
425 OrthoFinder v2.2.6 (80) using diamond v0.9.21.122 (81) searches. Orthologous sequences of 478  
426 single-copy proteins present in all the genomes were further aligned with Clustal Omega (82) and  
427 concatenated. Gblocks v0.91 (83) was used to remove highly divergent regions and poorly aligned  
428 columns. The resulting alignment of concatenated sequences was imported into RAXML-NG v1.0.2 (84)  
429 to build the ML phylogeny. The LG model of amino acid evolution (85), gamma-distributed substitution  
430 rates, and empirical frequencies of amino acids were used. Fast bootstrap was applied, with  
431 subsequent search for the best-scoring tree (86) and autoMRE (87). For the phylogenetic  
432 reconstruction of the different parts of the OBC3 cluster, amino acid sequences were retrieved and

433 aligned using Clustal Omega. Maximum likelihood trees were obtained from these aligned sequences  
434 using the LG substitution model (85) with SPR topology search. The robustness of each tree was  
435 assessed with 100 bootstrap replicates. The taxonomy of the phage  $\Phi$ GP100 were assigned according  
436 to the proteomic tree generated on the online platform ViPTree (88).

437

### 438 **Mutant construction**

439 Different mutants of *P. protegens* CHA0 were constructed (**Table S7**) to confirm the involvement of  
440 specific genes in phage resistance following the results of the Tn-seq experiment. These mutants were  
441 obtained using the suicide vector pEMG and the I-SceI system (89) with a protocol adapted for *P.*  
442 *protegens* CHA0 (29, 89), with plasmids and primers listed in the **Tables S1 and S8**, respectively.

443

### 444 **Heterologous expression of *fcl* from *Rhizobium etli* CFN42 and** 445 **lipopolysaccharide visualization**

446 For complementation of the  $\Delta fcl$  mutant of CHA0 *in cis* with the *fcl* gene of *Rhizobium etli* CFN42 or *in*  
447 *trans* with the *fcl* gene of CHA0, the respective *fcl* genes were cloned under the control of the  $P_{tac/lacIq}$   
448 promoter in the mini-Tn7 delivery vector pME8300 (90), generating the pME11021 and pME11022,  
449 respectively (**Table S1**). The pME8300 derivatives and the Tn7 transposition helper plasmid pUX-BF13  
450 were co-electroporated into competent cells of the CHA0  $\Delta fcl$  mutant to create the complemented  
451 strains CHA5211 and CHA5212, having  $P_{tac/lacIq}$ -*fcl*(CFN42) or  $P_{tac/lacIq}$ -*fcl*(CHA0), respectively, integrated  
452 at the unique chromosomal Tn7 attachment site (**Table S7**). Expression of *fcl* *in cis* or *in trans* was then  
453 induced by adding isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.1 mM to  
454 the cultures. Extraction, separation and visualization of LPS were performed as previously described  
455 (28).

456

### 457 **Phage sensitivity assay**

458 Infection of different *Pseudomonas* strains and isogenic mutants of CHA0 by the phage  $\Phi$ GP100 was  
459 assessed by a double agar layer assay. Aliquots of 4 mL of LB soft agar (5 g L<sup>-1</sup> Oxoid bacteriological  
460 agar), supplemented with IPTG if relevant, were mixed with 100  $\mu$ L of cell suspension of a given strain  
461 that was prepared from an overnight culture grown in 10 mL of LB (supplemented with antibiotics and  
462 IPTG, if relevant) at 25 °C with shaking at 180 rpm. The mixture then was poured onto a sterile NA  
463 plate. The phage stock suspension (containing 10<sup>9</sup> PFU mL<sup>-1</sup>) was serially diluted in a 96-well plate and  
464 10  $\mu$ L of each dilution were spotted onto the solidified double layer. Plates were evaluated for  
465 occurrence of plaques in the bacterial growth after overnight incubation at room temperature.

466

### 467 **Phage burst size and adsorption rate characterization**

468 The adsorption rate and the number of new phages released from infected cells, i.e., the burst size of  
469 the phage  $\Phi$ GP100, were determined following the infection of the different CHA0 mutants as  
470 compared to the parental strain. A bacterial suspension of 900  $\mu$ L containing  $2.5 \times 10^8$  bacterial cells  
471 at exponential growth stage ( $OD_{600nm} = 1$ ) was supplemented with 100  $\mu$ L containing  $2.5 \times 10^7$  PFU, i.e.,  
472 at a MOI of 0.1, and incubated at room temperature for 5 min. 100  $\mu$ L of this suspension was collected  
473 in a new tube containing 5  $\mu$ L of chloroform to determine the adsorption rate. The rest of the  
474 suspension was diluted to obtain a suspension containing 10<sup>5</sup> bacterial cells in 10 mL of fresh NYB. The  
475 burst size was monitored by sampling 100  $\mu$ L at 60 min then every 10 min until 150 min. Final samples  
476 were collected at 180 min. The collected samples were placed into a new tube containing 5  $\mu$ L of  
477 chloroform, incubated for 10 min, and then caps of the tubes were opened 10 min before samples  
478 being diluted into a 96-well plate. The enumeration of phages was then performed by spotting 10  $\mu$ L  
479 of each dilution onto the solidified double layer containing CHA0 wild type.

480

## 481 **Acknowledgements**

482 This study was supported by the Swiss National Centre of Competence in Research (NCCR)  
483 Microbiomes (no. 51NF40\_180575) and 310030\_184666 from the Swiss National Science Foundation  
484 (SNSF). Open access funding was provided by the University of Lausanne. We thank Dale Noel from  
485 Marquette University for providing us with *Rhizobium etli* CFN42 and Laure Weisskopf from the  
486 University of Fribourg for providing us with *Pseudomonas* sp. R76.

487

## 488 **Data availability**

489 The generated Tn-seq Datasets were deposited on EBI platform, sample accession number:  
490 SAMEA7616892 - SAMEA7616904 under the BioProject PRJEB61467.

491

## 492 References

- 493 1. Kupferschmied P, Maurhofer M, Keel C. 2013. Promise for plant pest control: root-  
494 associated pseudomonads with insecticidal activities. *Front Plant Sci* 4:287
- 495 2. Ramette A, Moënne-Loccoz Y, Défago G. 2003. Prevalence of fluorescent pseudomonads  
496 producing antifungal phloroglucinols and/or hydrogen cyanide in soils naturally  
497 suppressive or conducive to tobacco black root rot. *FEMS Microbiol Ecol* 44:35–43.
- 498 3. Haas D, Défago G. 2005. Biological control of soil-borne pathogens by fluorescent  
499 pseudomonads. *Nat Rev Microbiol* 3:307–319.
- 500 4. Péchy-Tarr M, Bruck DJ, Maurhofer M, Fischer E, Vogne C, Henkels MD, Donahue KM,  
501 Grunder J, Loper JE, Keel C. 2008. Molecular analysis of a novel gene cluster encoding an  
502 insect toxin in plant-associated strains of *Pseudomonas fluorescens*. *Environ Microbiol*  
503 10:2368–2386.
- 504 5. Flury P, Aellen N, Ruffner B, Péchy-Tarr M, Fataar S, Metla Z, Dominguez-Ferreras A,  
505 Bloemberg G, Frey J, Goesmann A, Raaijmakers JM, Duffy B, Höfte M, Blom J, Smits THM,  
506 Keel C, Maurhofer M. 2016. Insect pathogenicity in plant-beneficial pseudomonads:  
507 phylogenetic distribution and comparative genomics. *ISME J* 10:2527–2542.
- 508 6. Vacheron J, Péchy-Tarr M, Brochet S, Heiman CM, Stojiljkovic M, Maurhofer M, Keel C.  
509 2019. T6SS contributes to gut microbiome invasion and killing of an herbivorous pest  
510 insect by plant-beneficial *Pseudomonas protegens*. *ISME J* 13:1318–1329.
- 511 7. Imperiali N, Dennert F, Schneider J, Laessle T, Velatta C, Fesselet M, Wyler M, Mascher  
512 F, Mavrodi O, Mavrodi D, Maurhofer M, Keel C. 2017. Relationships between root  
513 pathogen resistance, abundance and expression of *Pseudomonas* antimicrobial genes,  
514 and soil properties in representative Swiss agricultural soils. *Front Plant Sci* 8:427.
- 515 8. Jaffuel G, Imperiali N, Shelby K, Campos-Herrera R, Geisert R, Maurhofer M, Loper J, Keel  
516 C, Turlings TCJ, Hibbard BE. 2019. Protecting maize from rootworm damage with the  
517 combined application of arbuscular mycorrhizal fungi, *Pseudomonas* bacteria and  
518 entomopathogenic nematodes. *Sci Rep* 9:3127.
- 519 9. Vannier N, Agler M, Hacquard S. 2019. Microbiota-mediated disease resistance in plants.  
520 *PLoS Pathog* 15:e1007740.
- 521 10. Jousset A, Rochat L, Péchy-Tarr M, Keel C, Scheu S, Bonkowski M. 2009. Predators  
522 promote defence of rhizosphere bacterial populations by selective feeding on non-toxic  
523 cheaters. *ISME J* 3:666–674.

- 524 11. Jiang Y, Liu M, Zhang J, Chen Y, Chen X, Chen L, Li H, Zhang X-X, Sun B. 2017. Nematode  
525 grazing promotes bacterial community dynamics in soil at the aggregate level. *ISME J*  
526 11:2705–2717.
- 527 12. Koskella B, Taylor TB. 2018. Multifaceted impacts of bacteriophages in the plant  
528 microbiome. *Annual Rev of Phytopathol* 56:361–380.
- 529 13. Dion MB, Oechslin F, Moineau S. 2020. Phage diversity, genomics and phylogeny. *Nat*  
530 *Rev Microbiol* 18:125–138.
- 531 14. van Elsas JD, Turner S, Bailey MJ. 2003. Horizontal gene transfer in the phytosphere. *The*  
532 *New Phytol* 157:525–537.
- 533 15. Bruto M, Prigent-Combaret C, Luis P, Moënne-Loccoz Y, Muller D. 2014. Frequent,  
534 independent transfers of a catabolic gene from bacteria to contrasted filamentous  
535 eukaryotes. *Proc R Soc B* 281:20140848.
- 536 16. Pratama AA, van Elsas JD. 2018. The ‘neglected’ soil virome – potential role and impact.  
537 *Trends Microbiol* 26:649–662.
- 538 17. Pratama AA, Terpstra J, de Oliveria ALM, Salles JF. 2020. The role of rhizosphere  
539 bacteriophages in plant health. *Trends Microbiol* 28:709–718.
- 540 18. Wang X, Wei Z, Yang K, Wang J, Jousset A, Xu Y, Shen Q, Friman V-P. 2019. Phage  
541 combination therapies for bacterial wilt disease in tomato. *Nat Biotechnol* 37:1513–  
542 1520.
- 543 19. Holtappels D, Fortuna K, Lavigne R, Wagemans J. 2021. The future of phage biocontrol in  
544 integrated plant protection for sustainable crop production. *Curr Opin Biotechnol* 68:60–  
545 71.
- 546 20. Keel C, Ucurum Z, Michaux P, Adrian M, Haas D. 2002. Deleterious impact of a virulent  
547 bacteriophage on survival and biocontrol activity of *Pseudomonas fluorescens* strain  
548 CHA0 in natural soil. *Mol Plant Microbe Interac* 15:567–576.
- 549 21. Vacheron J, Kupferschmied P, Resch G, Keel C. 2018. Genome sequence of the  
550 *Pseudomonas protegens* phage ΦGP100. *Genome Announ* 6:e00261-18.
- 551 22. Chaturongakul S, Ounjai P. 2014. Phage–host interplay: examples from tailed phages and  
552 Gram-negative bacterial pathogens. *Front Microbiol* 5:442
- 553 23. Shen Y, Loessner MJ. 2021. Beyond antibacterials – exploring bacteriophages as  
554 antivirulence agents. *Curr Opin Biotechnol* 68:166–173.

- 555 24. Chevallereau A, Pons BJ, van Houte S, Westra ER. 2021. Interactions between bacterial  
556 and phage communities in natural environments. *Nat Rev Microbiol* 20: 49-62
- 557 25. King JD, Kocíncová D, Westman EL, Lam JS. 2009. Review: Lipopolysaccharide  
558 biosynthesis in *Pseudomonas aeruginosa*. *Innate Immun* 15:261–312.
- 559 26. Okuda S, Sherman DJ, Silhavy TJ, Ruiz N, Kahne D. 2016. Lipopolysaccharide transport  
560 and assembly at the outer membrane: the PEZ model. 6. *Nat Rev Microbiol* 14:337–345.
- 561 27. Simpson BW, Trent MS. 2019. Pushing the envelope: LPS modifications and their  
562 consequences. *Nat Rev Microbiol* 17:403–416.
- 563 28. Kupferschmied P, Chai T, Flury P, Blom J, Smits THM, Maurhofer M, Keel C. 2016. Specific  
564 surface glycan decorations enable antimicrobial peptide resistance in plant-beneficial  
565 pseudomonads with insect-pathogenic properties. *Environ Microbiol* 18:4265–4281.
- 566 29. Heiman CM, Maurhofer M, Calderon S, Dupasquier M, Marquis J, Keel C, Vacheron J.  
567 2022. Pivotal role of O-antigenic polysaccharide display in the sensitivity against phage  
568 tail-like particles in environmental *Pseudomonas* kin competition. *ISME J* 16:1683-1693.
- 569 30. Larsen R, Wilson M, Guss A, Metcalf W. 2002. Genetic analysis of pigment biosynthesis  
570 in *Xanthobacter autotrophicus* Py2 using a new, highly efficient transposon mutagenesis  
571 system that is functional in a wide variety of bacteria. *Arch Microbiol* 178:193–201.
- 572 31. Smith TJ, Font ME, Kelly CM, Sondermann H, O’Toole GA. 2018. An N-terminal retention  
573 module anchors the giant adhesin LapA of *Pseudomonas fluorescens* at the cell surface:  
574 a novel subfamily of type I secretion systems. *J Bacteriol*:e00734-17.
- 575 32. Dean CR, Goldberg JB. 2002. *Pseudomonas aeruginosa galU* is required for a complete  
576 lipopolysaccharide core and repairs a secondary mutation in a PA103 (serogroup O11)  
577 *wbpM* mutant. *FEMS Microbiol Lett* 210:277–283.
- 578 33. Choudhury B, Carlson RW, Goldberg JB. 2005. The structure of the lipopolysaccharide  
579 from a *galU* mutant of *Pseudomonas aeruginosa* serogroup-O11. *Carbohydr Res*  
580 340:2761–2772.
- 581 34. Coyne MJ, Russell KS, Coyle CL, Goldberg JB. 1994. The *Pseudomonas aeruginosa algC*  
582 gene encodes phosphoglucomutase, required for the synthesis of a complete  
583 lipopolysaccharide core. *J Bacteriol* 176:3500–3507.
- 584 35. Ong SP, Azam AH, Sasahara T, Miyanaga K, Tanji Y. 2020. Characterization of  
585 *Pseudomonas* lytic phages and their application as a cocktail with antibiotics in  
586 controlling *Pseudomonas aeruginosa*. *J Biosci Bioengi* 129:693–699.

- 587 36. Burrows LL, Charter DF, Lam JS. 1996. Molecular characterization of the *Pseudomonas*  
588 *aeruginosa* serotype O5 (PAO1) B-band lipopolysaccharide gene cluster. *Mol Microbiol*  
589 22:481–495.
- 590 37. Reeves PP, Wang L. 2002. Genomic organization of LPS-specific loci. *Curr Top Microbiol*  
591 *Immunol* 264:109–135.
- 592 38. Wang L, Wang Q, Reeves PR. 2010. The variation of O-antigens in Gram-negative  
593 Bacteria, p. 123–152. *In* Wang, X, Quinn, PJ (eds.), *Endotoxins: Structure, Function and*  
594 *Recognition*. Springer Netherlands, Dordrecht.
- 595 39. Kutschera A, Schombel U, Wröbel M, Gisch N, Ranf S. 2019. Loss of *wbpL* disrupts O-  
596 polysaccharide synthesis and impairs virulence of plant-associated *Pseudomonas* strains.  
597 *Mol Plant Pathol* 20:1535–1549.
- 598 40. Li E, Zhang H, Jiang H, Pieterse CMJ, Jousset A, Bakker PAHM, de Jonge R. 2021.  
599 Experimental-evolution-driven identification of *Arabidopsis* rhizosphere competence  
600 genes in *Pseudomonas protegens*. *mBio* 12:e00927-21.
- 601 41. Pires DP, Oliveira H, Melo LDR, Sillankorva S, Azeredo J. 2016. Bacteriophage-encoded  
602 depolymerases: their diversity and biotechnological applications. *Appl Microbiol*  
603 *Biotechnol* 100:2141–2151.
- 604 42. Broecker NK, Barbirz S. 2017. Not a barrier but a key: How bacteriophages exploit host's  
605 O-antigen as an essential receptor to initiate infection. *Mol Microbiol* 105:353–357.
- 606 43. Knecht LE, Veljkovic M, Fieseler L. 2020. Diversity and function of phage encoded  
607 depolymerases. *Front Microbiol* 10:2949
- 608 44. Krupovic M, Turner D, Morozova V, Dyall-Smith M, Oksanen HM, Edwards R, Dutilh BE,  
609 Lehman SM, Reyes A, Baquero DP, Sullivan MB, Uchiyama J, Nakavuma J, Barylski J,  
610 Young MJ, Du S, Alfenas-Zerbini P, Kushkina A, Kropinski AM, Kurtböke I, Brister JR, Lood  
611 C, Sarkar BL, Yigang T, Liu Y, Huang L, Wittmann J, Chanishvili N, van Zyl LJ, Rumnieks J,  
612 Mochizuki T, Jalasvuori M, Aziz RK, Łobocka M, Stedman KM, Shkoporov AN, Gillis A, Peng  
613 X, Enault F, Knezevic P, Lavigne R, Rhee S-K, Cvirkaite-Krupovic V, Moraru C, Moreno Switt  
614 AI, Poranen MM, Millard A, Prangishvili D, Adriaenssens EM. 2021. Bacterial viruses  
615 subcommittee and archaeal viruses subcommittee of the ICTV: update of taxonomy  
616 changes in 2021. *Arch Virol* 166:3239–3244.
- 617 45. Oliveira H, Costa AR, Konstantinides N, Ferreira A, Akturk E, Sillankorva S, Nemeč A,  
618 Shneider M, Dötsch A, Azeredo J. 2017. Ability of phages to infect *Acinetobacter*  
619 *calcoaceticus*-*Acinetobacter baumannii* complex species through acquisition of different  
620 pectate lyase depolymerase domains. *Environ Microbiol* 19:5060–5077.

- 621 46. Bernheim A, Sorek R. 2020. The pan-immune system of bacteria: antiviral defence as a  
622 community resource. *Nat Rev Microbiol* 18:113–119.
- 623 47. Ojeda KJ, Simonds L, Noel KD. 2013. Roles of predicted glycosyltransferases in the  
624 biosynthesis of the *Rhizobium etli* CE3 O-antigen. *J Bacteriol* 195:1949–1958.
- 625 48. Rini JM, Esko JD. 2015. Glycosyltransferases and Glycan-Processing Enzymes, p. . *In* Varki,  
626 A, Cummings, RD, Esko, JD, Stanley, P, Hart, GW, Aebi, M, Darvill, AG, Kinoshita, T, Packer,  
627 NH, Prestegard, JH, Schnaar, RL, Seeberger, PH (eds.), *Essentials of glycobiology*, 3rd ed.  
628 Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).
- 629 49. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. 2006. ISfinder: the reference  
630 centre for bacterial insertion sequences. *Nucleic Acids Res* 34:D32-36.
- 631 50. Golubov A. 2021. Chapter 5 - Genome instability in bacteria: Causes and  
632 consequencesaaAdopted from chapter “Genome instability in bacteria: causes and  
633 consequences”, first edition, written by Ashley B. Williams., p. 73–90. *In* Kovalchuk, I,  
634 Kovalchuk, O (eds.), *Genome Stability (Second Edition)*. Academic Press, Boston.
- 635 51. Kaleta P, O’Callaghan J, Fitzgerald GF, Beresford TP, Ross RP. 2010. Crucial role for  
636 insertion sequence elements in *Lactobacillus helveticus* evolution as revealed by  
637 interstrain genomic comparison. *Appl Environ Microbiol* 76:212–220.
- 638 52. Boutoille D, Corvec S, Caroff N, Giraudeau C, Espaze E, Caillon J, Plésiat P, Reynaud A.  
639 2004. Detection of an IS21 insertion sequence in the *mexR* gene of *Pseudomonas*  
640 *aeruginosa* increasing  $\beta$ -lactam resistance. *FEMS Microbiol Lett* 230:143–146.
- 641 53. Moffatt JH, Harper M, Adler B, Nation RL, Li J, Boyce JD. 2011. Insertion sequence ISAb11  
642 is involved in colistin resistance and loss of lipopolysaccharide in *Acinetobacter*  
643 *baumannii*. *Antimicrob Agents Chemother* 55:3022–3024.
- 644 54. Voineagu I, Narayanan V, Lobachev KS, Mirkin SM. 2008. Replication stalling at unstable  
645 inverted repeats: Interplay between DNA hairpins and fork stabilizing proteins. *Proc Natl*  
646 *Acad Sci U S A* 105:9936–9941.
- 647 55. Bi X, Liu LF. 1996. DNA rearrangement mediated by inverted repeats. *Proc Natl Acad Sci*  
648 *U S A* 93:819–823.
- 649 56. Patil PB, Sonti RV. 2004. Variation suggestive of horizontal gene transfer at a  
650 lipopolysaccharide (LPS) biosynthetic locus in *Xanthomonas oryzae* pv. *oryzae*, the  
651 bacterial leaf blight pathogen of rice. *BMC Microbiol* 4:40.
- 652 57. Lawrence JG, Ochman H. 2002. Reconciling the many faces of lateral gene transfer.  
653 *Trends Microbiol* 10:1–4.

- 654 58. Bertelli C, Laird MR, Williams KP, Simon Fraser University Research Computing Group,  
655 Lau BY, Hoad G, Winsor GL, Brinkman FS. 2017. IslandViewer 4: expanded prediction of  
656 genomic islands for larger-scale datasets. *Nucleic Acids Res* 45:W30–W35.
- 657 59. Waack S, Keller O, Asper R, Brodag T, Damm C, Fricke WF, Surovcik K, Meinicke P, Merkl  
658 R. 2006. Score-based prediction of genomic islands in prokaryotic genomes using hidden  
659 Markov models. *BMC Bioinformatics* 7:142.
- 660 60. Lerouge I, Vanderleyden J. 2002. O-antigen structural variation: mechanisms and  
661 possible roles in animal/plant–microbe interactions. *FEMS Microbiol Rev* 26:17–47.
- 662 61. Chen Y, Bystricky P, Adeyeye J, Panigrahi P, Ali A, Johnson JA, Bush C, Morris J, Stine O.  
663 2007. The capsule polysaccharide structure and biogenesis for non-O1 *Vibrio cholerae*  
664 NRT36S: genes are embedded in the LPS region. *BMC Microbiol* 7:20.
- 665 62. Lu H, Patil P, Sluys M-AV, White FF, Ryan RP, Dow JM, Rabinowicz P, Salzberg SL, Leach  
666 JE, Sonti R, Brendel V, Bogdanove AJ. 2008. Acquisition and evolution of plant  
667 pathogenesis–associated gene clusters and candidate determinants of tissue-specificity  
668 in *Xanthomonas*. *PLoS One* 3:e3828.
- 669 63. Faoro H, Oliveira WK, Weiss VA, Tadra-Sfeir MZ, Cardoso RL, Balsanelli E, Brusamarello-  
670 Santos LCC, Camilios-Neto D, Cruz LM, Raittz RT, Marques ACQ, LiPuma J, Fadel-Picheth  
671 CMT, Souza EM, Pedrosa FO. 2019. Genome comparison between clinical and  
672 environmental strains of *Herbaspirillum seropedicae* reveals a potential new emerging  
673 bacterium adapted to human hosts. *BMC Genomics* 20:630.
- 674 64. Lam JS, Taylor VL, Islam ST, Hao Y, Kocíncová D. 2011. Genetic and functional diversity of  
675 *Pseudomonas aeruginosa* lipopolysaccharide. *Front Microbiol* 2:118.
- 676 65. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing  
677 reads. *EMBnet.journal* 17:10–12.
- 678 66. Davis MPA, van Dongen S, Abreu-Goodger C, Bartonicek N, Enright AJ. 2013. Kraken: a  
679 set of tools for quality control and analysis of high-throughput sequence data. *Methods*  
680 63:41–49.
- 681 67. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler  
682 transform. *Bioinformatics* 25:1754–1760.
- 683 68. DeJesus MA, Ambadipudi C, Baker R, Sasseti C, Ioerger TR. 2015. TRANSIT - A software  
684 tool for Himar1 TnSeq analysis. *PLoS Comput Biol* 11:e1004401.
- 685 69. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende  
686 DR, Letunic I, Rattei T, Jensen LJ, von Mering C, Bork P. 2019. eggNOG 5.0: a hierarchical,

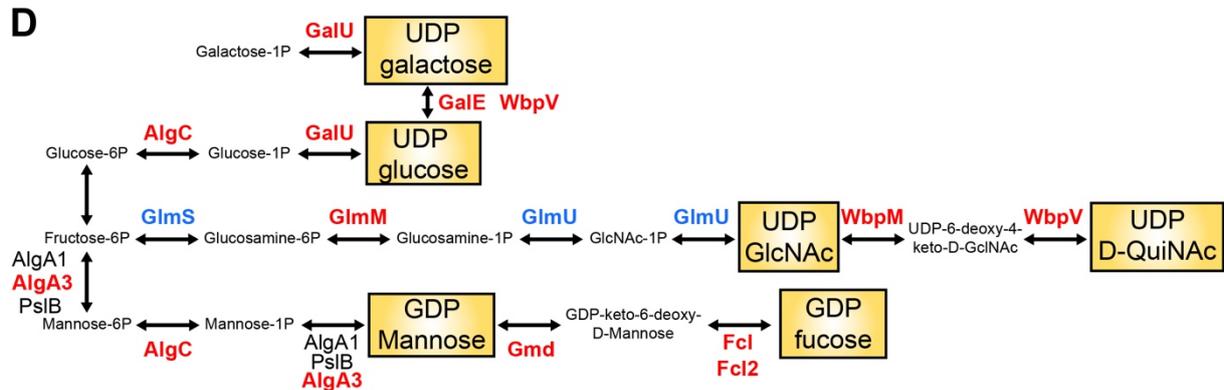
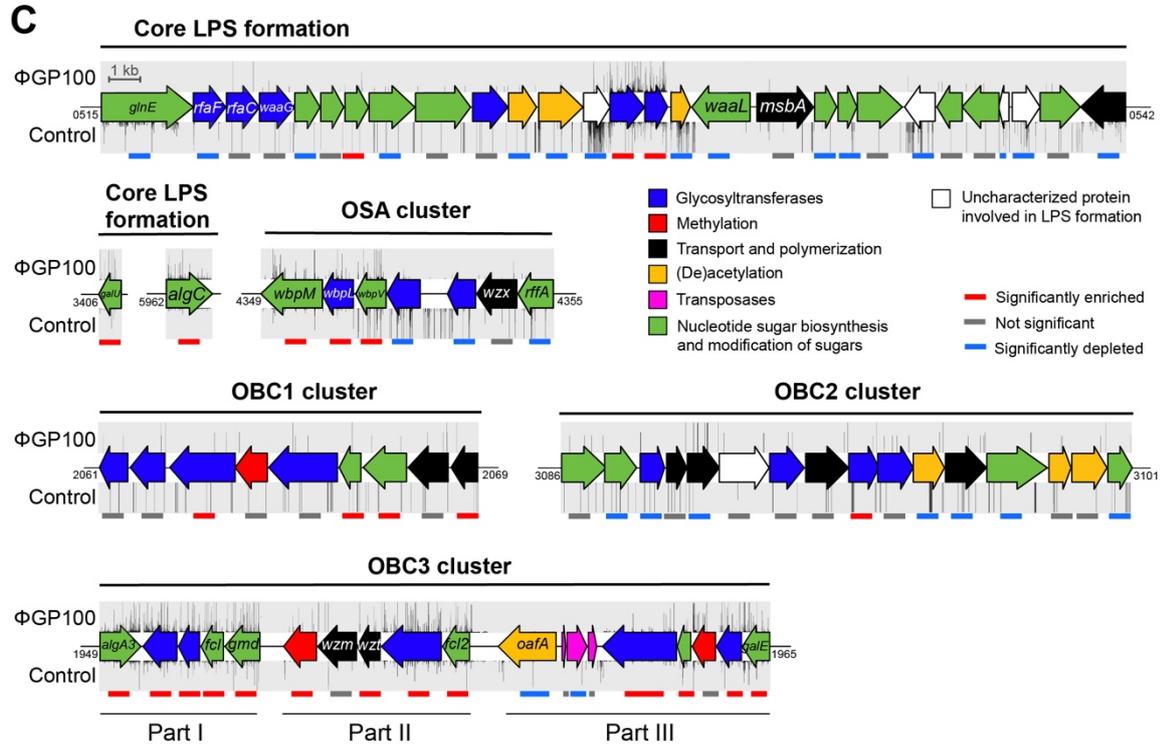
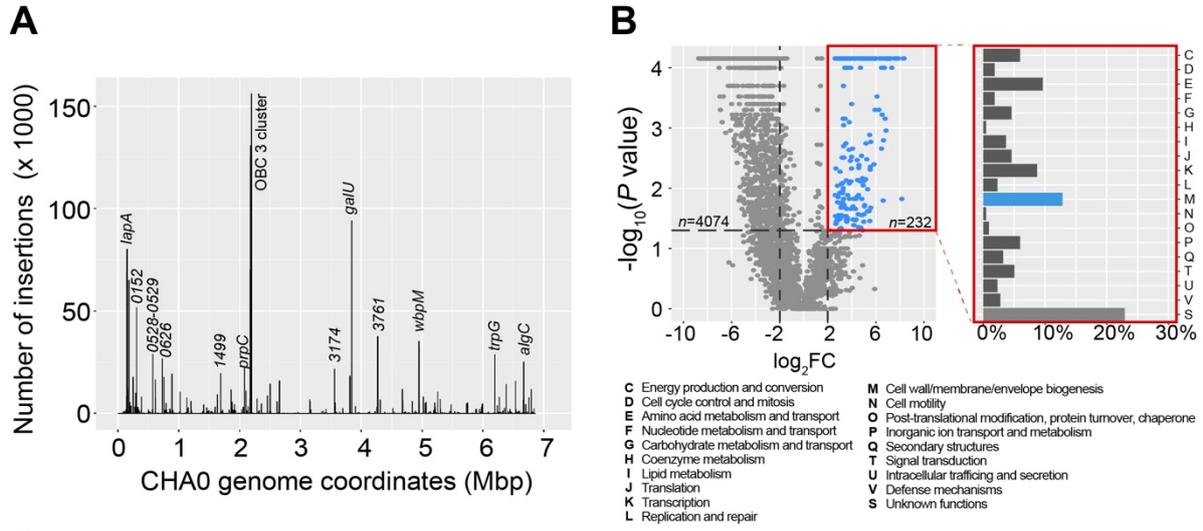
- 687 functionally and phylogenetically annotated orthology resource based on 5090  
688 organisms and 2502 viruses. *Nucleic Acids Res* 47:D309–D314.
- 689 70. Medema MH, Takano E, Breitling R. 2013. Detecting sequence homology at the gene  
690 cluster level with MultiGeneBlast. *Mol Biol Evol* 30:1218–1223.
- 691 71. Gilchrist CLM, Booth TJ, van Wersch B, van Grieken L, Medema MH, Chooi Y-H. 2021.  
692 cblaster: a remote search tool for rapid identification and visualization of homologous  
693 gene clusters. *Bioinformatics Adv* 1:vbab016.
- 694 72. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. 2012. Artemis: an integrated  
695 platform for visualization and analysis of high-throughput sequence-based experimental  
696 data. *Bioinformatics* 28:464–469.
- 697 73. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*  
698 30:2068–2069.
- 699 74. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better,  
700 faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–21.
- 701 75. Darzentas N. 2010. Circoletto: visualizing sequence similarity with Circos. *Bioinformatics*  
702 26:2620–2621.
- 703 76. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K,  
704 Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A,  
705 Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E,  
706 Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals  
707 O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure  
708 prediction with AlphaFold. 7873. *Nature* 596:583–589.
- 709 77. Holm L. 2022. Dali server: structural unification of protein families. *Nucleic Acids Res*  
710 50:W210–W215.
- 711 78. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI,  
712 Marchler GH, Song JS, Thanki N, Yamashita RA, Yang M, Zhang D, Zheng C, Lanczycki CJ,  
713 Marchler-Bauer A. 2020. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic*  
714 *Acids Res* 48:D265–D268.
- 715 79. Garrido-Sanz D, Redondo-Nieto M, Martin M, Rivilla R. 2021. Comparative genomics of  
716 the *Pseudomonas corrugata* subgroup reveals high species diversity and allows the  
717 description of *Pseudomonas ogarae* sp. nov. *Microb Genom* 7:000593.
- 718 80. Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative  
719 genomics. *Genome Biol* 20:238.

- 720 81. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND.  
721 Nat Methods 12:59–60.
- 722 82. Sievers F, Higgins DG. 2014. Clustal Omega, accurate alignment of very large numbers of  
723 sequences. Methods Mol Biol 1079:105–116.
- 724 83. Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use  
725 in phylogenetic analysis. Mol Biol Evol 17:540–552.
- 726 84. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAXML-NG: a fast, scalable  
727 and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics  
728 35:4453–4455.
- 729 85. Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. Mol Biol  
730 Evol 25:1307–1320.
- 731 86. Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAXML  
732 web servers. Syst Biol 57:758–771.
- 733 87. Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. 2010. How  
734 many bootstrap replicates are necessary? J Comput Biol 17:337–354.
- 735 88. Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H, Goto S. 2017. ViPTree: the viral  
736 proteomic tree server. Bioinformatics 33:2379–2380.
- 737 89. Martínez-García E, de Lorenzo V. 2011. Engineering multiple genomic deletions in Gram-  
738 negative bacteria: analysis of the multi-resistant antibiotic profile of *Pseudomonas*  
739 *putida* KT2440. Environ Microbiol 13:2702–2716.
- 740 90. Kupferschmied P, Péchy-Tarr M, Imperiali N, Maurhofer M, Keel C. 2014. Domain  
741 shuffling in a sensor protein contributed to the evolution of insect pathogenicity in plant-  
742 beneficial *Pseudomonas protegens*. PLoS Pathog 10:e1003964.
- 743 91. Vacheron J, Heiman CM, Keel C. 2021. Live cell dynamics of production, explosive release  
744 and killing activity of phage tail-like weapons for *Pseudomonas* kin exclusion. Commun  
745 Biol 4:1–14.

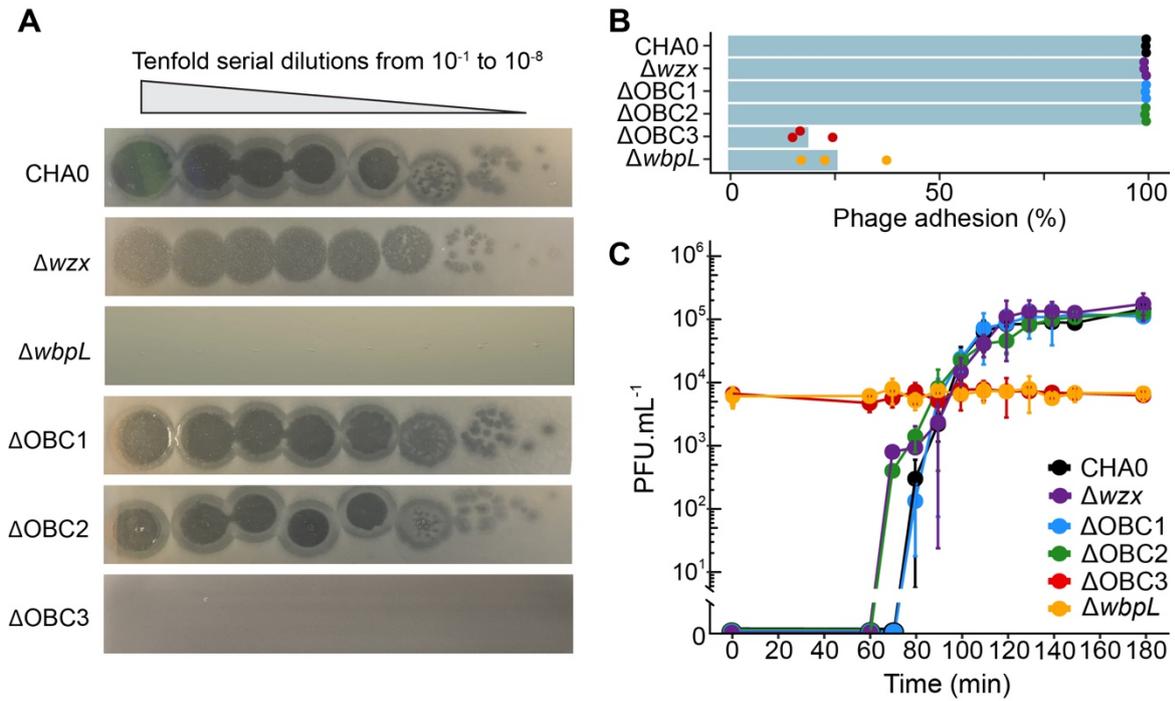
746

747

748 **Figures legends**

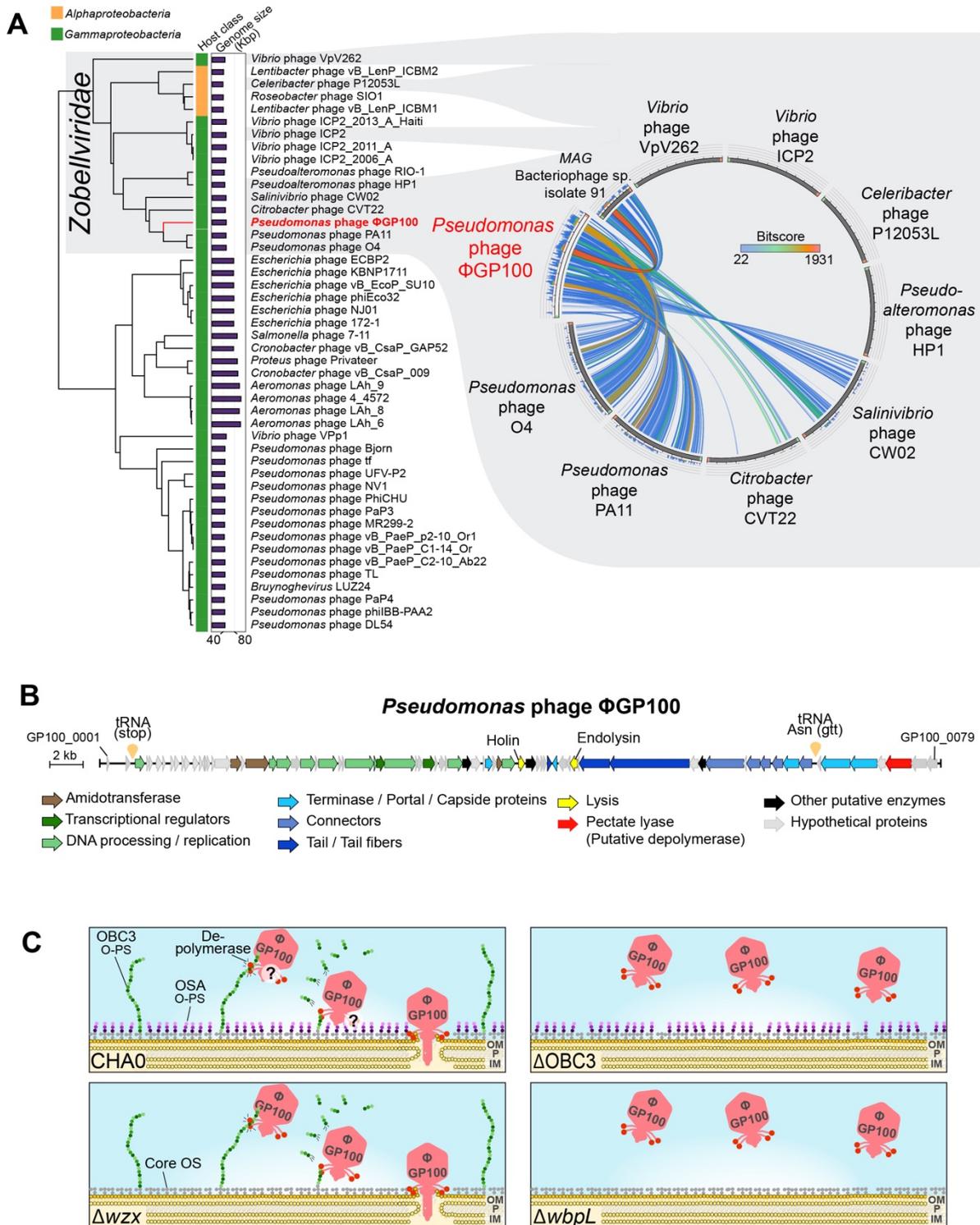


750 **Fig. 1: Tn-seq analysis specifies lipopolysaccharides (LPS) components of *Pseudomonas protegens***  
751 **CHA0 as the main receptors of the phage  $\Phi$ GP100.** (A) Distribution of the transposon insertions across  
752 the genome of *P. protegens* CHA0 in presence of the phage  $\Phi$ GP100 at a multiplicity of infection (MOI)  
753 of 100. The results for all MOI levels assessed in this study are available in **Figure S1**. (B) Volcano plot  
754 highlighting the enriched 232 candidate genes involved in phage  $\Phi$ GP100 resistance and their cluster  
755 of orthologous groups (COGs) assignation. (C) Visualization of the transposon enrichment in the  
756 different gene clusters involved in LPS formation of CHA0 upon exposure to  $\Phi$ GP100. The density of  
757 the transposon insertion is indicated above (condition with the phage) or below (control condition)  
758 the genes. The colored lines below the gene clusters refer to the Tn-seq results: gray bars refer to a  
759 non-significant involvement of the corresponding gene in the phage sensitivity ( $-2 < \log_2FC < 2$  and  $p$   
760  $> 0.05$ ); red and blue lines refer to a significantly enriched ( $\log_2FC > 2$  and  $p < 0.05$ ) or depleted ( $\log_2FC$   
761  $< -2$  and  $p < 0.05$ ) number of transposon insertions, respectively, in the corresponding genes. Numbers  
762 at the start or the end of the respective clusters correspond to the gene locus tags for *P. protegens*  
763 CHA0 (prefix PPRCHA0\_...). (D) Biosynthetic pathways of the nucleotide sugars involved in the LPS  
764 formation important for the sensitivity of CHA0 to  $\Phi$ GP100. Enzymes written in red or blue display,  
765 respectively, a significantly higher or lower number of Tn insertions in their respective coding genes  
766 when CHA0 is exposed to  $\Phi$ GP100. Enzymes written in black were not significantly highlighted in the  
767 Tn-seq analysis.  
768  
769



770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781

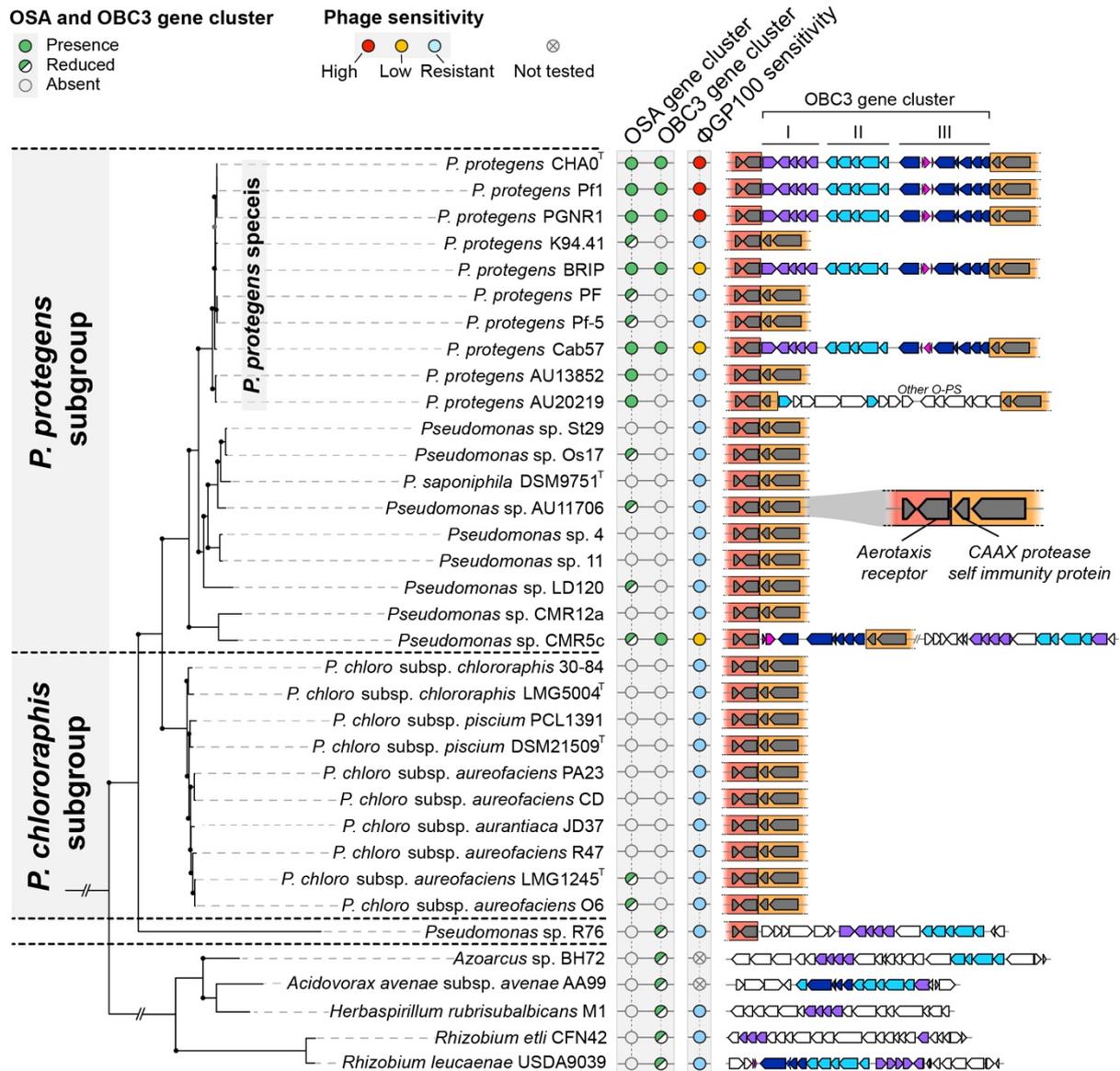
**Fig. 2: The long O-PS of CHA0 is mandatory for phage infection.** (A) Spot assays of  $\Phi$ GP100 onto a bacterial lawn of wild type CHA0 or mutants impaired in the production of different O-PS. The wild type CHA0 produces two major O-PS, the short OSA and the long OBC3 (28). The  $\Delta OBC3$  mutant lacks the OBC3 O-PS, whereas the  $\Delta wzx$  mutant does not possess an OSA O-PS. The  $\Delta wbpL$  mutant lacks both O-PS. OBC1 and OBC2 make no apparent contribution to O-PS formation (28). (B) The number of new phages released following infection of CHA0 wild type and the different O-PS mutants of CHA0. (C) Phage adsorption onto the cell surface of CHA0 wild type and the different O-PS mutants of CHA0.



782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792

**Fig. 3: Characterization of the phage ΦGP100 and potential mode of action towards CHA0. (A)** Taxonomy assignment of the phage ΦGP100 into the *Zobellviridae* family. The proteomic tree was generated on the ViPTree online platform (88). The host category as well as the sizes of the different phage genomes are displayed with colored vertical and horizontal bars, respectively. The genomic similarity between the phage ΦGP100 and other closely related phages from the *Zobellviridae* family is visualized alongside the proteomic tree. Colors reflect the similarity levels related to the values of the maximum hit scores (Bitscore). The list of genomes used in this figure panel are available in **Table S3**. **(B)** Update of the phage annotation with the identification of the pectate lyase that is potentially involved in depolymerization of the long OBC3-type O-PS of CHA0. Genes are color-coded according

793 to their protein functions. The detailed functional annotations of the  $\Phi$ GP100 phage is provided in  
794 **Table S4. (C)** Schematic representation of the proposed infection process of the phage  $\Phi$ GP100  
795 according to the sensitivity profiles obtained for the different O-PS mutants.  $\Phi$ GP100 targets the long  
796 OBC3-type O-PS at the surface of CHA0 cells and displays a depolymerization activity that potentially  
797 degrades this long O-PS (see plaques for the *wzx* mutant in **Figure 2A**). The localization of the  
798 depolymerase on the phage as well as its interaction with the short O-PS (OSA type) still need to be  
799 investigated, which is the reason why they are represented with interrogation points. OM: outer  
800 membrane; P: peptidoglycan; IM: inner membrane.  
801  
802



803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

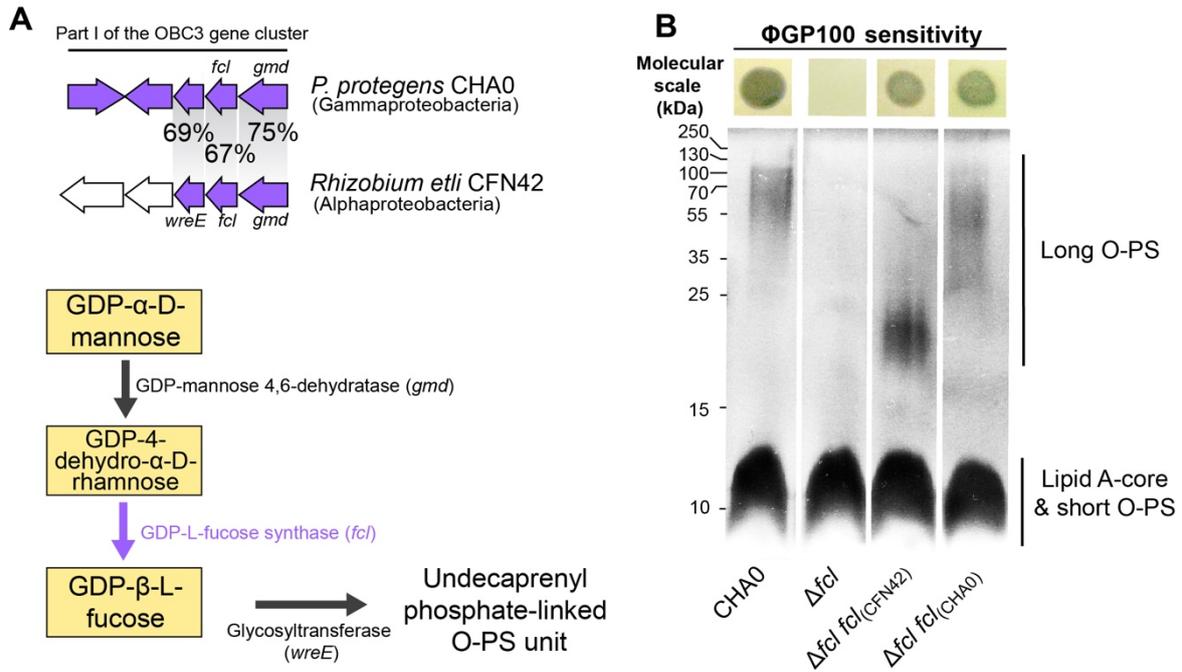
819

820

821

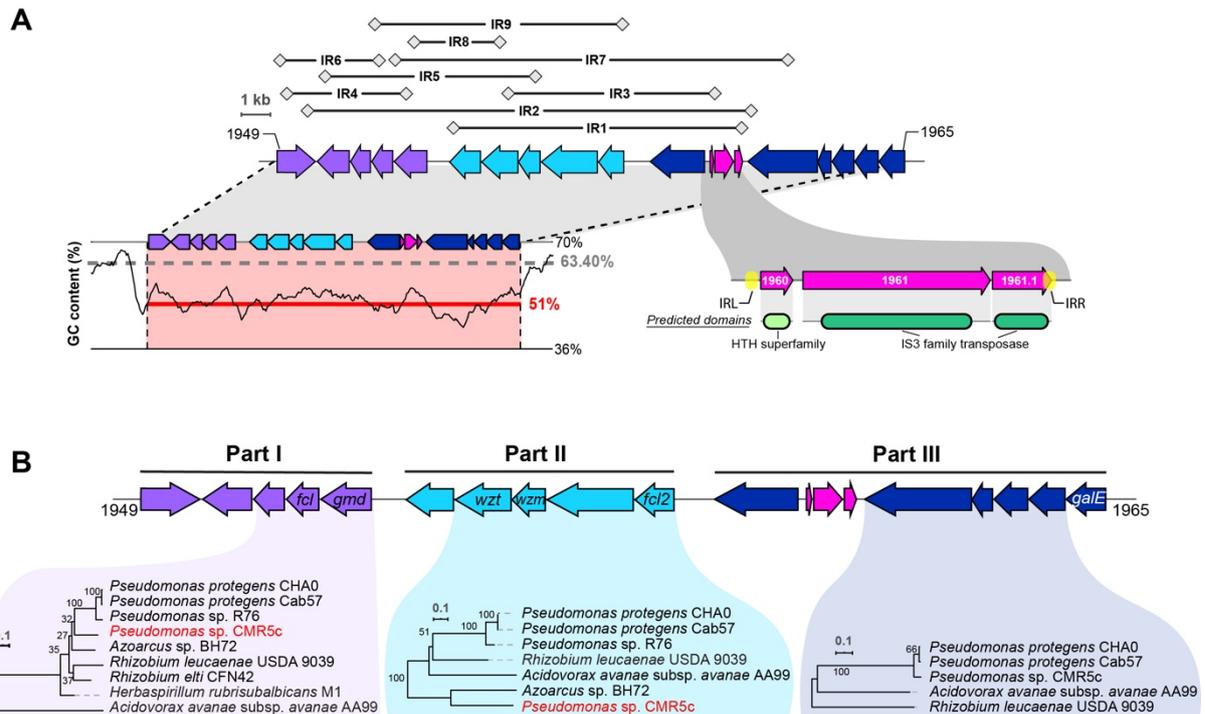
822

**Fig. 4: Distribution of the OSA-type and OBC3-type gene clusters among bacterial genomes and phage sensitivities.** The presence of the OBC3 and the OSA gene clusters was assessed by searching for orthologs of the OSA and OBC3 gene clusters of *P. protegens* CHA0 whose ensuing amino-acid sequences shared a minimum of 60 % of amino acid identity on at least 70 % of the amino acid sequence length. Entire or reduced versions of these two O-PS gene clusters were detected in closely related *Pseudomonas* (belonging to the *P. protegens* species *sensu-stricto* (91)) or phylogenetically distant bacterial genomes. The distribution of these two O-PS as well as the sensitivity of the different strains tested are represented alongside the phylogenetic tree. The synteny of the detected OBC3 gene clusters (entire or reduced version are represented on the right part of the figure). The OBC3 gene cluster is divided into three different parts, colored with purple, dark and light blue to better visualize which parts are found in the different genomes. Parts or the entire OBC3 gene cluster were found between two specific conserved regions inside the *Pseudomonas* genomes that are underlined in dark and light orange. Genes represented in pink correspond to mobile genetic elements including transposases. The maximum-likelihood phylogenetic tree of the 34 bacterial genomes was built based on the concatenation of single-copy proteins (see Material and methods section). Bootstrap values above 95% are represented with black dots.



823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837

**Fig. 5: The heterologous complementation of the long O-PS synthesis restores phage sensitivity.** (A) Nucleotide homology between genes of the part I of the OBC3 gene cluster of *Pseudomonas protegens* CHA0 and the O-PS biosynthesis locus detected in *Rhizobium etli* CFN42. Three genes of the OBC3 gene cluster show amino acid sequence identities of at least 67 % with genes that are necessary for the biosynthesis of L-fucose-containing O-PS (47). GDP-L-fucose is synthesized by conversion of GDP-D-mannose by a GDP-mannose dehydrogenase and a GDP-L-fucose synthase and can subsequently be used to assemble O-antigen units. (B) SDS-PAGE with LPS extracted from CHA0, its  $\Delta fcl$  mutant and *cis* and *trans* complemented strains. Long O-PS were lost in the  $\Delta fcl$  mutant and (partly) rescued by complementation with *fcl* from *R. etli* CFN42 and CHA0, respectively. Each strain was tested for susceptibility to infection by phage  $\Phi$ GP100 by performing a double layer assay. The corresponding pictures are shown above the gel. Clear zones indicate bacterial lysis upon phage infection.



838  
839

840 **Fig. 6: The OBC3 gene cluster of *Pseudomonas protegens* CHA0 displays a high genetic plasticity and**  
 841 **signatures of horizontal gene transfer acquisitions. (A)** Nine inverted repeats (IR, represented above  
 842 the OBC3 gene cluster), which are unique at genome scale, were detected. The nucleotide sequences  
 843 of these IRs are available in **Tables S6**. The GC content of the OBC3 gene cluster is represented at the  
 844 bottom left of the gene cluster. The dark line illustrates the GC skew calculated using a 500-bp window  
 845 frame. The gray dashed lines represent the average GC content of the whole genome, while the red  
 846 lines correspond to the average GC content of the OBC3 gene cluster. A detailed view of the mobile  
 847 genetic elements present within the OBC3 gene cluster is found at the bottom right of the gene cluster.  
 848 This transposon is flanked by two inverted repeats (IRL: 5' TGA ATC GCT CCG GGT TTC GTA GGC ACC  
 849 TCT TTG CCT TAG AAT GAG GCC AA 3' and IRR: 5' AAG CGT TAC GCA ATG AGT TTG CAA GGT GTC TAG  
 850 AGA GTC CGG GGC GAT TT 3') and comprises two genes (PPRCHA0\_1960 and PPRCHA0\_1961)  
 851 predicted to encode a protein containing a HTH superfamily domain and a IS3 family transposase,  
 852 respectively. The third coding sequence (PPRCHA0\_1961.1) corresponds to a manual annotation of a  
 853 CDS for a protein with a partial transposase domain. **(B)** Maximum likelihood trees were obtained from  
 854 the concatenated amino acid sequences of the corresponding genes conserved in all genomes, which  
 855 are highlighted in the different parts. The robustness of each tree was assessed with 100 bootstraps  
 856 replicates. The *Pseudomonas* strain highlighted in red displays incongruences in the tree topology  
 857 compared to the species phylogenetic tree shown in the **Fig. 4**.

858