Ecological ubiquity and phylogeny drive nestedness in phages–bacteria networks and shape the bacterial defensome

Chloé Feltin [1]*, Sylvain Piry [1], Benoit Moury [1], Lola Chateau [1], Karine Berthier [1], Jonathan M. Jacobs [2,3], Jules Butchacas [2,3], Linda Fenske [4], Lillian Ebeling-Koning [2,3], Theo H. M. Smits [5], Cindy E. Morris [1], Clara Torres-Barceló [6]*

[1] INRAE, Pathologie Végétale, 84140 Montfavet, France

[2] Department of Plant Pathology, The Ohio State University, Columbus, OH 43210, USA

[3] Infectious Disease Institute, The Ohio State University, Columbus, OH 43210, USA

[4] Bioinformatics and Systems Biology, Justus Liebig University Giessen, Giessen, Germany

[5] Environmental Genomics and Systems Biology Research Group, Institute of Natural Resource Sciences (IUNR), Zurich University of Applied Sciences ZHAW, Wädenswil, Switzerland

[6] PHIM Plant Health Institute, University of Montpellier, INRAE, CIRAD, Institut Agro, IRD, Montpellier, France

* Corresponding authors

E-mails: clara.torresbarcelo@inrae.fr and chloe.feltin@inrae.fr

Abstract

Identifying the ecological and evolutionary factors that shape phage–bacterial interactions is key to understanding their dynamics in microbial communities. Yet, such interactions remain poorly characterised in plant agroecosystems. Here, we investigate the ecological determinants of the interaction between a highly diverse set of 23 phages isolated from diseased apricot trees and 44 bacterial strains from the *Pseudomonas syringae* species complex collected either from diseased apricot trees, healthy plants or non-agricultural environment. Based on their ecological origin, we expected phages to preferentially infect bacterial strains from the same ecological context, forming modular host-range patterns. Contrary to these expectations, we discovered a significantly nested structure, suggesting generalised infection dynamics rather than local adaptation, primarily driven by the broad ecological dynamics of this pathosystem. Analysis of the bacterial genomes showed that both the profiles of anti-phage defence systems and the distribution profiles of prophages are strongly shaped by bacterial phylogeny. Furthermore, while the number of defence systems showed limited correlation with the breadth of bacterial sensitivity to phages, prophage abundance exhibited a strong, non-linear link with phage virulence. Together, these findings provide an ecological and evolutionary perspective on phage–bacterium infection networks and new insights into a better understanding of the role of phages in agricultural ecosystems.

Author Summary

Viruses that infect bacteria, known as phages, are part of microbial communities and influence the abundance, diversity, and traits of their hosts. In an agriculture-related context, they are commonly considered as potential biocontrol agents, but studying the bases of fundamental phage–bacterial interactions may help us better understand the plant microbiome and its

applications. Many factors influence these interactions, and identifying which ones matter most remains a challenge. In our study, we investigated how phages from diseased plants interact with bacteria collected from diseased and healthy plants, as well as from surrounding environments. We expected phages to mainly infect bacteria from similar environments, but instead observed that they often infected bacteria regardless of their source. This suggests that phage activity in this system has few barriers, reflecting the wide ecological distribution of their bacterial hosts. We further investigated how bacteria defend against phages by identifying both defence systems and prophages within their genomes and using this information to explore their contribution to bacterial resistance or sensitivity to phages. Together, our findings offer new insights into how phage–bacterium relationships evolve and function in plant ecosystems.

Introduction

Phages profoundly influence bacterial diversity and population dynamics (1). The impact of phages on bacterial ecology and evolution have been extensively examined in ecosystems like oceans and the human gut (2–7). However, in agriculture, research has predominantly focused on the use of phages as biocontrol agents, while their ecological roles, population dynamics, and interactions with bacterial communities in plant-associated ecosystems remain largely underexplored (8). A deeper understanding of these ecological aspects could provide fundamental insights to guide the development of phage-based biocontrol strategies and sustainable management of plant-associated microbial communities in agricultural systems.

Phage interactions with bacteria depends on a multitude of factors. The phage must recognise and infect a given bacterium, integrate into the host genome and/or replicate while avoiding or blocking bacterial defence systems (9,10). A range of approaches have been employed to

66 unravel the determinants involved in phage-bacterial interactions, including ecological and
67 evolutionary analyses of natural populations, and experimental evolution (11,12). Mechanistic
68 studies have also investigated the bacterial receptors (13,14) with more recent work focusing
69 on defence systems associated with susceptibility to phage (14–17). For example, a study
70 showed that *Vibrio lentus* strains that share identical phage receptors display varying levels of
71 susceptibility depending on the phage defence elements present in their genomes (18). As
72 with other aspects of phage-bacterial interactions, research on prophages and defence
73 systems in phytopathogenic bacteria remains limited and have yet to connect these elements
74 to host range data (19–25).

75 The complexities of phage-bacterial interactions can be visualised as ecological networks to
76 describe the underlying ecological and evolutionary drivers. These interaction networks are
77 rarely random and often exhibit distinct structural patterns, commonly modularity or
78 nestedness (26). Nestedness, a structure in which phages exhibit a gradient of host range
79 breadth and where bacteria display varying levels of susceptibility, is commonly observed
80 when studies involve limited genetic diversity of phages or bacteria (26–28). In contrast,
81 modularity, where phages and bacteria interact within modules but rarely across them, tends
82 to emerge at broader evolutionary or ecological scales, such as when interactions involve
83 diverse bacterial hosts across species or genera, or when examined over extended time
84 periods (11,29). On a large scale, phage-bacterium interaction networks often exhibit modular
85 structures with nested substructures within each module, patterns that are mostly observed
86 across broad geographic regions (30), and/or in a host set comprising different bacterial
87 species (31). A recent study has introduced new statistical methods that incorporate
88 quantitative host–parasite interactions to analyse the structures of ecological networks,
89 providing deeper insights into the processes shaping phage–bacterial interactions (32). Rather

90 than relying solely on binary data (presence or absence of infection), these approaches
91 consider the strength of interactions, reflecting how effective each phage is at infecting each
92 host, and vice versa. This gradient-based perspective enables a more nuanced understanding
93 of infection patterns, host susceptibility, and phage infectivity, revealing subtler ecological and
94 evolutionary dynamics that would be overlooked by binary analyses alone.

95 The *Pseudomonas syringae* species complex is particularly interesting for studying phage-
96 bacterial interactions in agricultural contexts due to its high genetic diversity and broad
97 ecological distribution (33). This phytopathogen colonises and causes disease across a wide
98 range of cultivated and wild plants and is also commonly found in water-related environments
99 such as snow, clouds, rain, rivers and epilithic biofilms (33–35). Several studies highlight the
100 dynamic exchanges occurring between agricultural and non-agricultural ecosystems
101 (33,35,36). A recent large-scale analysis of 590 *P. syringae* genomes revealed the diversity and
102 composition of its elaborate "defensome" integrated into the genome (including prophages,
103 and defence systems) (24). Yet, how this genomic arsenal translates into phage resistance
104 phenotypes remains unclear. Correlating phage–bacterial infection network with defensome
105 composition offers a valuable framework to infer potential links between genetic defences
106 and bacterial susceptibility to phages.

107 Historically, phage host range profiling was used to differentiate *P. syringae* strains (37–41).
108 More recent studies have emphasised the isolation and characterisation of novel *P. syringae*
109 phages and their potential as biocontrol agents (42,43). When reported, host-range assays
110 commonly aim to identify broad-host-range phages for practical applications (44). However,
111 ecological questions remain understudied, even though they offer valuable insights that
112 advance both fundamental and applied research. Investigating the roles of local adaptation,

host phylogeny and bacterial genomic defence arsenal (defensome) could help to clarify the mechanisms shaping these networks, thereby guiding more effective biocontrol strategies.

In apricot trees (*Prunus armeniaca*), *P. syringae* strains from diverse phylogroups (PG) are responsible for bacterial canker (45). The four main implicated phylogroups are: PG01, which includes plant pathogens and environmental strains; PG02, the most ubiquitous phylogroup, frequently isolated from diverse habitats (soil, water, plants, *ect*.); PG03, which comprises numerous pathovars pathogenic to woody plants but is rarely found in the environment; and PG07, which includes some pathogenic strains but is primarily associated with diverse environmental reservoirs (45,46). The high genetic diversity of these strains suggests that pathogenic success in apricot may be driven more by ecological fitness than by specific virulence factors alone (47). This fitness includes the ability of strains to survive, compete, and reproduce effectively within their environmental context, including interactions with the host plant and other microbes, such as phages. Notably, *P. syringae* lineages found in apricot trees are not host-specific but likely represent strains with general ecological fitness, enabling them to succeed in various environmental conditions (45). This raises questions about the ecological contexts in which these strains interact with phages, whether within apricot tree hosts or in external reservoirs. Exploring the phage infectivity patterns on these *P. syringae* strains can reveal how phylogenetic and ecological diversity shape phage-bacterium dynamics.

We focused on the causal agents of bacterial canker of apricot trees to investigate the structure and ecological drivers of phage–bacterium infection networks. Based on previous studies, we hypothesised that phages would be locally adapted to their bacterial hosts, leading to a modular infection network structure. We further proposed that this structure could be explained by the presence of prophages and diverse bacterial defence systems encoded in

136 host genomes. To test these hypotheses, we addressed the following questions: i) Are phages
137 from diseased apricot trees specifically adapted to resident *P. syringae* compared to closely-
138 related strains from other environmental reservoirs? ii) Do closely-related bacteria from
139 different environments differ in their adaptation to phage? iii) What bacterial genomic traits
140 underlie the observed infection network structure? We conducted a quantitative host-range
141 assay using 23 genetically-diverse *P. syringae* phages and 44 *P. syringae* strains. Whole-
142 genome sequencing of all bacterial strains enabled detailed annotation of prophages and
143 defence systems. This integrative approach provides mechanistic insights into how ecological
144 origin and bacterial genomic content shape phage–bacterial interactions, contributing to our
145 understanding of the roles that phages play in agricultural ecosystems.

146 **Results**

147 *Host range characterization of a diverse panel of phages on P. syringae bacteria*

148 We characterised the phage–bacterial infection network by testing 23 phages against 44
149 strains from the *P. syringae* species complex (hereafter referred to as *P. syringae* strains) in
150 liquid culture. This allowed us to go beyond the information provided by traditional plaque
151 assays – that reveal phage replication – by providing data on bacterial growth. Each interaction
152 was quantified using two complementary metrics: 1) a continuous variable, corresponding to
153 the percentage of bacterial growth inhibition by phages (0–100%) (S1A Table), and 2) a binary
154 variable, indicating whether the inhibition was statistically significant (1) or not (0), based on
155 triplicate measurements (S1B Table). Because these values can be interpreted from either the
156 phage or bacterial perspective, we provide clear terminology to distinguish them. From the
157 bacterial perspective, these values represent susceptibility (percentage inhibition) and

158 sensitivity (binary significance of inhibition). Conversely, from the phage perspective, these
159 values reflect virulence (percentage inhibition) and infectivity (binary success of infection).

160 Overall, the tested bacterial strains showed continuous distributions of susceptibility and
161 sensitivity values, reflecting the diversity of interactions of the panel with phages. On average,
162 bacterial susceptibility to phages was moderate, with a 24% reduction in growth, and strains
163 were sensitive to 12.3 out of the 23 tested phages (53.5%) (S1A Fig), indicating a broader
164 sensitivity than suggested by inhibition levels alone. The most susceptible bacterial strains
165 were DG1a.7, GAW0089, and DG12.6 (all from PG02), with average inhibition levels of 46%,
166 50%, and 42%, respectively (S1A Table). The most broadly sensitive strains, DG1a.7, 7C, and
167 CCE0067 (all from PG02), were significantly inhibited by 20, 19, and 19 phages, respectively
168 (87%, 83%, and 83% of tested phages respectively) (S1B Table). In contrast, bacterial strains
169 CC1583, DG8.15, and DGA9CS5-8 (from PG10, PG07 and PG10 respectively) showed the
170 highest resistance, with average inhibition percentages of only 3%, 6%, and 6%, respectively
171 (S1A Table). DG8.15 (from PG07) resisted to 17 phages (74% of tested phages) and seven other
172 strains (19B, 41D, CST0086, DGA9CS5-15, FMU-107, and USA0088 from PG01, PG02, PG02,
173 PG10, PG07 and PG02) were resistant to 16 phages (70%) (S1B Table).

174 From the phage perspective, average bacterial growth inhibition across the strain panel was
175 23%, reflecting moderate virulence. However, each phage was able to infect an average of
176 22.4 out of 42 tested *P. syringae* strains (53.3%) (S1B Fig), indicating a relatively broad
177 infectivity despite moderate inhibition levels. The most virulent phages were Ghual01 (65%
178 inhibition), Arace01 (57%), and Draal03 (44%) (S1A Table). The most infecting phages were
179 Nican01, Arace01, and Ghual01 infecting 39 (93% of tested bacteria), 37 (88%), and 34 (81%)
180 strains, respectively (S1B Table). Moreover, Arace01 and Ghual01 were the only phages

181 capable of infecting strains from other species in the *Pseudomonas* genus. Arace01 infected a
182 *P. rhizosphaerae* strain at 4.5×10^4 PFU/mL, while Ghual01 infected a *P. graminis* strain at 1.7
183 $\times 10^5$ PFU/mL. In contrast, Orimi01, Cygsa01, and Pavpe01 displayed the lowest virulence,
184 averaging 0%, 3%, and 4% respectively (S1A Table), while Orimi01, Cygsa01, and Touem01 had
185 the lowest infectivity, infecting only 2 (5% of tested bacteria), 5 (12%), and 11 (26%) strains,
186 respectively (S1B Table).

187 These findings reveal substantial variability in phage virulence and infectivity, as well as
188 bacterial susceptibility and sensitivity, with over one-third (39%) of combinations resulting in
189 significant growth inhibition. While some phages exhibit a broad host range, extending even
190 beyond *P. syringae*, others are highly specialised, only infecting a few strains, reflecting a
191 complex interplay between phage infectivity and bacterial resistance.

192 *Nested structure of the phage-bacterial interaction network and association with*
193 *bacterial traits*

194 The structure of the phage-bacterial infection network was analysed using multiple algorithms
195 to assess its nestedness and modularity. Given the diversity of both phages and bacterial
196 hosts, spanning 13 phage genera and several phylogroups of *P. syringae*, and the ecological
197 origins of phages (diseased apricot trees) and bacteria (diseased apricot trees, healthy plants
198 and other environmental reservoirs), we hypothesised a modular structure type. However,
199 the network representing bacterial growth inhibition percentage by phages displayed a strong
200 nested structure, as indicated by scores of 0.48 and 0.44, on a scale varying from 0 to 1, with
201 the *WINE* and *wNODF* algorithms, respectively (Table 1). Nestedness was highly significant (p
202 < 0.01) against nearly all null models, except null model B, which was shown to be excessively
203 conservative (32). Modularity analyses differed depending on the algorithm (Table 1). With

204 the *spinglass* algorithm, selected for its good performance in terms of false positive rate (32),
205 the modularity score was low (0.01, on a scale varying from -0.5 to 1) and was not significant,
206 whatever the null model. The four alternative modularity algorithms produced slightly higher
207 scores (0.09 < modularity < 0.25) and showed significant modularity ($p < 0.01$) with both null
208 models C1 and R1. These four algorithms show lower performances than *spinglass* in terms of
209 false positive rates, but combined with null models C1 and R1, they show generally acceptable
210 performances in terms of false positive and false negative rates (32). These results suggest
211 globally a slight degree of modular organisation within the network. The strong and significant
212 nestedness and the weak modularity were further supported by the analysis of the binary data
213 representing the significance of bacterial growth inhibition (S2 Table).

214

215 **Table 1. Analysis of nestedness and modularity of the bacterial growth inhibition network**
216 **combining 44 strains of the *P. syringae* species complex and 23 phages**.

Analysis	Algorithm	Score (max. = 1.0)	Null model						
			B	N	C1	R1	S	C2	R2
Nestedness	wNODF	0.44	0.44	**0**	**0**	**0**	**0**	**0**	**0**
	WINE	0.48	1	**0**	**0.01**	**0**	**0**	**0**	**0**
Modularity	spinglass	0.01	1	1	1	1	0.07	0.29	0.18
	edge betweenness	0.09	**0**	**0.02**	**0**	**0**	1	1	1
	fast greedy	0.24	**0**	**0**	**0**	**0**	1	0.81	**0.02**
	leading eigenvector	0.14	**0**	**0**	**0**	**0**	1	1	0.94
	beckett	0.25	**0**	**0**	**0**	**0**	1	**0.01**	1

217 Null models are described in Moury *et al.* (32). For each null model and each algorithm, the
218 frequency of simulations, among 100, showing a strictly higher nestedness (or modularity)

219 score than that of the actual matrix is indicated. Significantly nested or modular networks (p
220 ≤ 0.05) are indicated in red.

221

222 To assess whether the nested structure of phage–bacterial networks is shaped by bacterial
223 evolutionary history, we first tested for a phylogenetic signal in bacterial sensitivity summed
224 by bacteria (Fig 1). The analysis revealed a strong phylogenetic signal ($\lambda = 1.0$; LR ($\lambda = 0$) =
225 48.26; $p < 0.01$), indicating that closely-related bacterial strains tend to be infected by a similar
226 number of phages. To explore this pattern further, we evaluated the effect of phylogroup
227 identity using both a Generalised Linear Model (GLM) with the bacterial sensitivity summed
228 by bacteria (Fig 1C) and a PErmutational Multivariate ANalysis Of VAriance (PERMANOVA) on
229 bacterial susceptibility profiles (S2A Fig). Both analyses confirmed a significant influence of
230 phylogroup on bacterial susceptibility, explaining almost 10% of the variance (GLM: $p < 0.01$;
231 PERMANOVA: $R^2 = 0.09$; $p < 0.01$). In contrast, ecological origin had no significant effect in
232 either test (GLM: $p = 0.87$; PERMANOVA: $R^2 = 0.05$; $p = 0.29$). Specifically, PG02 strains
233 emerged as the most sensitive to phages, being infected on average by 14.6 phages and
234 showing the highest susceptibility, with 20.7% growth inhibition. PG01 strains showed
235 intermediate sensitivity being infected by 10.4 phages and an average of 14.6% inhibition,
236 followed closely by PG10 strains with 15.3% inhibition. In contrast, PG07 strains were the most
237 resistant overall, infected on average by only 8.4 phages and displaying the lowest
238 susceptibility with 8.73% inhibition.

239

240 **Fig 1. Networks of percentages of growth inhibition of 44 *P. syringae* species complex strains**
241 **(rows) by 23 phages (columns).**

242 (A) Taxonomic organisation of the network presented on a maximum likelihood phylogenetic
243 tree based on bacterial core-genomes and phage genera. A node with a maximum likelihood
244 support value below 0.90 is marked with a red dot. (B) The same network after row and
245 column permutations showing its nested pattern, with phylogroups and ecological origins of
246 *P. syringae* strains displayed on the right side of panels A and B. (C) Bacterial sensitivity to
247 phages by bacterial phylogroup and ecological origins. Letters indicate Tukey post-hoc test
248 results from a generalised linear model (GLM) with a Poisson distribution; ns indicates non-
249 significant differences ($p > 0.05$). (D) Phylogenetic Host Range Index (PHRI) across phage
250 genera.

251

252 The broad taxonomic diversity of the phages studied here, covering 13 genera, includes both
253 specialists and generalists. To characterise phage host range, we calculated the Phylogenetic
254 Host Range Index (PHRI), which incorporates the number, the evenness, and the phylogenetic
255 diversity of host bacteria (48). The PHRI values ranged from 0 (Orimi01, a specialist phage) to
256 0.38 (Arace01, a generalist phage), forming a gradient that reflects the nested structure of the
257 phage-bacteria interaction network (Fig 1D). Based on PHRI classification, 9% of phages were
258 categorised as specialists, 65% as generalists, and 26% as having an intermediate host range
259 (Fig 1D). At the genetic level, significant differences in virulence were observed between
260 phage genera (GLMM; $p < 0.01$). The generalist genus *Aravirus* exhibited the highest virulence
261 (52.8%), while the specialist genera *Cygnusvirus* and *Orionvirus* displayed markedly lower
262 virulence levels, averaging 1.1% and 4.2%, respectively (S2B Fig). Additionally, phages with
263 similar host infection profiles clustered by taxonomy, with those sharing >70% nucleotide
264 identity (same genus) explaining 63.5% of the variation in virulence (PERMANOVA: $R^2 = 0.64$,

265 $p < 0.01$), indicating genus-level host specificity (Fig 1A). For example, the two phages from
266 the *Lyravirus sulafat* species and the *Uliginvirus* genus exhibited highly similar infection host
267 profiles.

268 These results demonstrate that the phage–bacterial interaction networks are predominantly
269 nested, with minimal modularity suggesting broad host range overlap among phages rather
270 than local adaptation based on ecological origin. Bacterial susceptibility is phylogenetically
271 constrained, with PG02 strains being highly sensitive and PG07 strains resistant. Additionally,
272 although more than half of the phages were generalists, genus-level similarity drives host-
273 range patterns.

274 *Defence systems: profile and association with ecology and phylogeny*

275 The 44 strains of *P. syringae* analysed in this study were examined for their accessory genome
276 content, with a particular focus on their repertoire of defence systems to counteract mobile
277 genetic elements such as phages. Across these strains, a total of 48 defence system types and
278 86 subtypes were identified (Fig 2A). The most abundant defence subtypes included Septu (n
279 = 81), type IV Restriction-Modification (RM) (n = 37), and Gabija (n = 31) (Fig 2A). In contrast,
280 31 subtypes were sporadically detected across the bacterial genomes, including rare defence
281 systems such as three Abi subtypes (AbiH, AbiJ, and AbiO), two Asv subtypes (AsvI and AsvII),
282 and three DTR subtypes (DTR6, DTR8, and DTR3) (Fig 2A). Defence systems are present at an
283 average rate of 9.93 per strain, with two strains (AF0015 and P3.01.09C09, both from PG01)
284 harbouring the highest number (15 systems), and four strains (DG1A_7 from PG02, CSZ0293,
285 SZ0119, and SZB0008 from PG10) the lowest number (5 systems) (Fig 2A). Each defence
286 system can be classified into different functional categories to better characterise its role.

287 Specifically, 38.4% of defence system subtypes are associated with abortive infection, 24.4%
288 with nucleic acid degradation, and 37.2% remain unclassified (Fig 2A).

289

290 **Fig 2. Defence system distribution in bacterial genomes and relationship with bacterial**
291 **sensitivity to phages.**

292 (A) Matrix showing the number of defence systems present in the 44 *P. syringae* strains,
293 organised on a maximum likelihood phylogenetic tree based on their core-genomes. A node
294 with a maximum likelihood support value below 0.90 is marked with a red dot. The phylogroup
295 affiliation of each bacterial strain is indicated on the left using colours, and defence systems
296 are represented by subtypes with colours on the right indicating their respective functional
297 category. (B) Number of defence systems per bacterial strain according to their phylogroups
298 and ecological origins. ns: not significant ($p > 0.05$), according to post-hoc Tukey tests from a
299 Generalised Linear Model with a Poisson distribution. (C) Bacterial sensitivity to phages as a
300 function of the number of defence systems, with linear regressions performed within each
301 phylogroup.

302

303 To assess whether the distribution of defence systems across bacterial genomes was
304 influenced by phylogeny, we tested for a phylogenetic signal in the total number of encoded
305 systems per genome. A strong signal indicated that closely related strains tend to harbour
306 similar numbers of defence systems ($\lambda = 1.0$; LR test ($\lambda = 0$): 44.18; $p < 0.01$) (S3 Fig), as
307 described in a previous study (24). Phylogroup significantly explained variation in defence
308 system profiles (PERMANOVA, $R^2 = 0.14$, $p < 0.01$), with a weaker effect from ecological origin
309 ($R^2 = 0.06$, $p < 0.01$), while growth rate of the bacteria had no significant impact ($R^2 = 0.02$, p

310 = 0.51) (Fig 2A). Post-hoc comparisons based on this analysis revealed that strains from PG10
311 displayed significantly distinct defence system profiles compared to those from PG02 and
312 PG07, while strains isolated from healthy plants exhibited distinct defence system profiles
313 compared to other ecological origins. However, none of these factors significantly affected
314 the total number of defence systems, although the effect of phylogroup approached
315 significance (GLM: $p = 0.06$) (Fig 2B). This phylogroup trend is supported by a larger dataset of
316 590 *P. syringae* genome assemblies, where PG01 strains encode the highest number of
317 defence systems, while PG07 and PG10 harbour some of the lowest counts (24).

318 While closely-related strains tend to possess similar numbers of defence systems, we further
319 investigated whether they also share the same subtypes of defence systems. To this end, we
320 calculated the percentage of shared defence system subtypes between each pair of bacterial
321 strains, where 100% indicates complete identity in the sets of subtypes. However, no
322 significant correlation was observed between the percentage of shared defence system
323 subtypes and the genome-wide Average Nucleotide Identity (ANI) between strains (linear
324 regression; adjusted $R^2 = 0.03$; $p < 0.01$) (S4 Fig). This suggests that, although phylogeny and
325 ecological origins influence the profile of defence systems, their inheritance does not follow a
326 strict vertical transmission. Furthermore, only 1.14% of the identified defence systems were
327 linked to prophages within bacterial genomes, indicating that the vast majority are located
328 outside prophage regions. This low prevalence indicates that while defence systems may
329 spread horizontally among bacterial populations, their dissemination is likely mediated by
330 mobile genetic elements other than prophages.

331 Defence systems may impose varying energetic costs on bacterial cells, potentially influencing
332 their allocation of resources. In strains of *P. syringae*, defence systems related to nucleic acid

333 degradation comprised, on average, 54% of the total defence systems, while abortive
334 infection systems, which are considered more metabolically costly (49), made up an average
335 of 25% (S3 Fig). We then revealed that the proportion of defence system categories were
336 significantly shaped by phylogroup, explaining 16% of the variance (PERMANOVA; $R^2 = 0.16$; p
337 $= 0.03$), and growth rate ($R^2 = 0.07$; $p = 0.04$), but not ecological origin ($R^2 = 0.08$; $p = 0.11$).
338 Pairwise comparisons based on this analysis showed that strains from PG07 had significantly
339 different defence category profiles compared to those from PG01 and PG02.

340 Our results revealed that the defence arsenal profiles in *P. syringae* are shaped by phylogeny
341 and ecological origin, but their specific composition is not strictly inherited, likely due to
342 horizontal transfer. Overall, these findings underscore the complexity of bacterial defence
343 strategies and indicate that multiple evolutionary and ecological factors influence phage
344 resistance.

345 *Prophages: profile and association with ecology and phylogeny*

346 We identified 88 prophages among the 44 genomes of *P. syringae*. On average, each strain
347 carried two prophages, with one strain (FMU-107) harbouring five prophages and 18 strains
348 (41%) containing a single cryptic prophage (Fig 3A). Among these prophages, 36% were
349 classified as high quality (complete or nearly complete according to CheckV software), 39% as
350 medium quality, and 25% as low quality. Prophage sizes ranged from 12.0 kb to 89.8 kb, with
351 an average size of 32.2 kb. The size distribution exhibited three peaks at approximately 18 kb,
352 44 kb, and 57 kb. To classify these prophages taxonomically, we used VIRIDIC, which identified
353 65 novel phage species and 42 novel genera (S3 Table), while proteomic tree analysis with
354 Viptree revealed six distinct prophage clusters, corresponding approximately to family-level
355 taxonomic divisions (Fig 3A) (50). One of these, the CL4 family-like prophage, was found in all

356 44 *P. syringae* strains, which span four distinct phylogroups (Fig 3A). CL4 family-like prophage
357 is a tailocin, specifically an R-type syringacin, a bacteriocin derived from a bacteriophage
358 (51,52). These tailocins are widely distributed across the *P. syringae* species complex and are
359 thought to play a key role in inter-bacterial competition by enabling strains to kill closely-
360 related competitors (53,54). Moreover, we found that the taxonomy of the CL4 family-like
361 prophage mirrors the phylogeny of their bacterial hosts across 68 strains from 10 phylogroups
362 (including the 44 strains analysed in this study), with a significant cophylogenetic congruence
363 signal (Procrustes Approach to Cophylogeny (PACo): residual sum of squares = 0.15; $p < 0.01$)
364 (S5 Fig).

365

366 **Fig 3. Diversity and distribution of prophages in bacterial genomes and relationship with**
367 **bacterial sensitivity to phages.**

368 (A) Distribution of prophages in the genomes of the 44 *P. syringae* strains, categorised by
369 genus and family-like cluster (VipTree). Bacterial strains are organised based on a maximum
370 likelihood phylogenetic tree based on their core-genomes, with their phylogroups and
371 ecological origins indicated on the right. A node with a maximum likelihood support value
372 below 0.90 is marked with a red dot. (B) Abundance of prophages per bacterial strain
373 according to their phylogroups and ecological origins. ns: not significant ($p > 0.05$), according
374 to post-hoc Tukey tests from a Generalised Linear Model with a Poisson distribution. (C)
375 Bacterial sensitivity to phages as a function of the number of prophages, coloured by
376 phylogroups.

377

378 The abundance of prophages (*i.e.* the number of prophages per bacterial genome) and the
379 prophage distribution profiles in bacterial genomes are strongly shaped by phylogeny. A
380 significant phylogenetic signal was detected in prophage abundance ($\lambda = 1.0$; LR ($\lambda = 0$) = 46.48;
381 $p < 0.01$), indicating that closely related strains tend to harbour similar numbers of prophages
382 (Fig 3B). Phylogroup significantly explained 16% of variation in prophage distribution profiles
383 (PERMANOVA, $R^2 = 0.16$; $p < 0.01$), with a weaker effect from ecological origin ($R^2 = 0.07$; $p <$
384 0.01), while growth rate had no significant impact ($R^2 = 0.03$; $p = 0.09$) (Fig 3A). All distinct
385 phylogroups, as well as ecological origins, exhibited significantly different prophage
386 distribution profiles, supporting the idea that lineage and ecological context influence the
387 genomic landscape of prophages. However, none of these factors significantly affected the
388 abundance of prophages (GLM: phylogroup, $p = 0.27$; ecological origin, $p = 0.92$; growth rate,
389 $p = 0.30$), despite some observed variation. Indeed, PG07 strains showed the highest prophage
390 abundance, nearly double that of PG10 strains (Fig 3B). However, this pattern contrasts with
391 findings from broader dataset (590 *P. syringae* genome assemblies), in which PG02 strains
392 encoded among the highest numbers of prophages, PG07 among the lowest, and PG01 and
393 PG10 showed intermediate levels (24). In contrast, differences across ecological origins were
394 minimal (Fig 3B).

395 In conclusion, the comprehensive analysis of prophages within the *P. syringae* species complex
396 revealed a diverse array of complete and incomplete prophages integrated into the genomes.
397 The prophage distribution profiles are strongly explained by bacterial phylogeny and
398 ecological origins of the host strains.

399 *Contribution of bacterial defence systems and prophages to bacterial sensitivity to*
400 *phages*

401 Since bacterial defence systems target phages (49), a negative correlation is expected
402 between their number and bacterial sensitivity to phages. To test this, we examined the
403 correlation between bacterial sensitivity and the number of encoded defence systems, while
404 accounting for phylogenetic background. For both PG01 and PG02, no significant correlation
405 was detected (linear regression for PG01: adjusted $R^2 = 0.12$; $p = 0.18$; for PG02: adjusted $R^2 =$
406 -0.07; $p = 0.94$) (Fig 2C). In contrast, a negative correlation was detected in PG10, supporting
407 our expectation (linear regression for PG10: adjusted $R^2 = 0.34$; $p = 0.05$) (Fig 2C). In contrast,
408 PG07 exhibited an unexpected positive correlation, suggesting that strains with a higher
409 number of defence systems were more sensitive to phages (linear regression for PG07:
410 adjusted $R^2 = 0.92$; $p < 0.01$) (Fig 2C).

411 While prophages can contribute to bacterial immunity against phage infection (55), our results
412 revealed no significant correlation between prophage abundance and bacterial sensitivity to
413 phages. This suggests that the detected prophages do not play a defining role in resistance,
414 either overall (linear regression; adjusted $R^2 = -0.02$; $p = 0.81$) or within individual phylogroups
415 (PG01: adjusted $R^2 = 0.01$, $p = 0.33$; PG02: adjusted $R^2 = -0.02$, $p = 0.43$; PG07: adjusted $R^2 =$
416 0.01, $p = 0.38$; PG10: adjusted $R^2 = -0.12$, $p = 0.94$) (Fig 3C). Moreover, no correlation was
417 observed between prophage abundance and the number of defence systems encoded in
418 bacterial genomes, suggesting that these systems do not provide protection against prophage
419 integration or maintenance (linear regression; adjusted $R^2 = -0.02$; $p = 0.64$).

420 *Influence of multiple bacterial and phage traits on bacterial growth inhibition*

421 Having identified several bacterial and phage factors that individually influenced the bacterial
422 growth inhibition, we next sought to integrate these explanatory variables into a unified
423 framework to gain a more comprehensive understanding of phage–bacteria interaction
424 dynamics. The primary model included prophage abundance and the total number of defence
425 systems per strain as bacterial traits associated with phage resistance, alongside phage- and
426 bacterium-related traits examined in this study. We also tested an alternative model
427 incorporating PCA-derived variables related to bacterial defensome, but it did not improve
428 model performance (Akaike Information Criterion (AIC) = 16.70) compared to the initial model
429 (AIC = 16.69). Therefore, we retained the first model for its greater interpretability.

430 Consistent with the above findings, bacterial phylogeny and phage taxonomy emerged as the
431 strongest explanatory variables for inhibition percentage, with an importance weight of 1.0
432 across the 256 models (Table 2). In addition, a variable indicating whether the phage had been
433 isolated using one of the tested bacterial strains was significantly associated with inhibition
434 level. Specifically, phages isolated from strains within the tested host range exhibited a slightly
435 lower inhibition capacity (mean = 15.1%) compared to those isolated from external strains
436 (mean = 17.7%).

437

438 **Table 2. Relative importance of explanatory variables influencing bacterial growth inhibition**
439 **by phages.**

Explanatory variable	*Sum of AIC weights*	*N Models*
Phage genera ϕ	**1.00**	54

Bacterial phylogroups	**1.00**	53
Number of prophages in bacterial genomes	**0.72**	29
Phages isolated with bacteria from the data set ϕ	**0.67**	31
Number of defence systems in bacterial genomes	0.38	26
Bacteria from the data set having isolated phages	0.37	24
Bacterial growth rate	0.29	23
Bacterial ecological origins	0.12	21

Sum of AIC weights was computed from model averaging across the top-ranked models whose cumulative weights reached ≥ 0.95, selected from a full set of 256 candidate models generated by multi-model inference. The relative importance of each variable corresponds to the sum of AIC weights across all retained models in which it appears. Bold values indicate variables with high relative importance (sum of weights ≥ 0.60), selected as significant predictors in the analysis. N Models: Number of retained models in which each variable appears. Variables related to phages are marked with the symbol ϕ; bacterial variables are unmarked.

The model identified the total number of prophages in bacterial genomes as a significant predictor of inhibition of bacterial growth (sum of weights = 0.72; Table 2). Inhibition levels varied nonlinearly with prophage abundance: inhibition levels were 15.6%, 14.7%, 17.5%, 28.4%, and 10.9% for strains carrying 1 to 5 prophages, respectively. While prophage abundance did not predict bacterial sensitivity, it appeared to influence bacterial susceptibility, suggesting that prophage-rich bacteria may be more susceptible to highly-effective phages. In contrast, the total number of defence systems did not significantly explain

455 bacterial growth inhibition (sum of weights = 0.38; Table 2), despite its correlation with the
456 bacterial sensitivity for some phylogroups.

457 Overall, these findings highlight that the primary determinants of the bacterial growth
458 inhibition by phage infection within the *P. syringae* species complex are bacterial phylogeny
459 and phage taxonomy rather than local ecological adaptation. The integration of multiple
460 explanatory variables within a unified modelling framework further revealed that prophage
461 abundance, but not the total number of defence systems, significantly contributes to
462 variations in inhibition levels. This suggests that while prophages may not act as strict barriers
463 to infection, their presence can modulate the efficacy of phage infection. Together, these
464 results emphasize the complex and multifactorial nature of bacterial susceptibility to phages,
465 shaped by both evolutionary history and prophage-mediated genomic landscapes.

466

467 **Discussion**

468 Our study uncovered a strong nested structure in the phage–bacterial infection network
469 involving genetically-diverse *P. syringae* strains and phages. This discovery is in strong,
470 contrast with the modular structure frequently reported (11,29). This pronounced nestedness
471 likely reflects the ubiquitous nature of the highly diverse *P. syringae* complex, mirrored by a
472 phage community ranging from specialists to generalists in host range. While bacterial
473 phylogeny had a marked influence on the nestedness of the infection network, we found no
474 evidence of local phages-bacteria co-adaptation. Defence systems and prophages contributed
475 to phage resistance, but their distribution was largely shaped by bacterial phylogeny and
476 ecological origin, rather than phage pressure alone. Together, these findings highlight how

477 evolutionary and ecological forces shape phage–bacterial infection networks in natural
478 environments.

479 *Ecological origin and phylogenetic influences on bacterial sensitivity to phages*

480 The observed nested infection network is unexpected when we consider the high genetic
481 diversity of the phages (13 genera, 17 species) and bacteria (4 phylogroups and 7 clades,
482 corresponding to genus and species levels, respectively) used in this work as well as the varied
483 ecological origins of the bacterial strains. Such a diversity of organisms could be expected to
484 reveal local adaptation, where phages perform better on sympatric than allopatric bacterial
485 hosts, in line with what has been demonstrated previously (56,11). However, our findings
486 suggest that phage–host interactions in this pathosystem transcend ecological boundaries.
487 Phages isolated from apricot tree soil infected bacterial strains from both diseased orchards
488 and aquatic environments with comparable efficiency, and bacterial origin did not broadly
489 influence bacterial sensitivity/susceptibility to phages. This absence of modularity, combined
490 with a lack of ecological signal in host range, points to broader ecological connectivity and
491 population flow within the *P. syringae* species complex. These results align with previous work
492 (36) showing that *P. syringae* population structures are shaped more by fluctuating ecological
493 dynamics than by geographic isolation, likely due to documented exchanges between
494 agricultural and aquatic environments (57). Overall, the nested structure observed here
495 implies that interaction dynamics in this system are maintained by widespread generalism and
496 ongoing ecological mixing, rather than niche-specific adaptation.

497 Bacterial phylogeny plays a central role in shaping both the composition of defence systems,
498 the distribution of prophages, and the sensitivity of *P. syringae* strains to phages. In our study,
499 PG02 strains, known for their ubiquitous presence across habitats and for containing strains

500 capable of infecting a broad range of plant species (46,35), were also the most sensitive to
501 phages, carried a moderate number of defence systems, and encoded one of the highest
502 prophage loads (24). In contrast, PG07 strains, which are predominantly associated with
503 diverse environmental reservoirs (46), were the most resistant to phages despite encoding
504 relatively few defence systems and prophages. PG01 strains, mainly isolated from diseased
505 plants (46), showed intermediate bacterial sensitivity to phages, possessed the greatest
506 number of defence systems, and encoded a moderate prophage load (24). Finally, PG10
507 strains, exclusively environmental, showed moderate bacterial sensitivity to phages, low
508 numbers of defence systems, and a moderate prophage count (24). These findings suggest a
509 potential trade-off between ecological ubiquity and bacterial sensitivity to phages, a
510 hypothesis that requires direct experimental investigation. In a similar vein, a recent study
511 reported that effector gene counts in *P. syringae*, which are key to pathogenicity, positively
512 correlate with phage defences and show partial conservation within phylogroups (24).

513 While phylogroups show consistent trends in defence system profiles, closely-related strains
514 often do not share identical defence systems, suggesting acquisition through horizontal gene
515 transfers (49,58). However, strains within the same phylogroup tend to retain consistent
516 features in their defensome, including the number and functional categories of defence
517 systems. This pattern suggests that each *P. syringae* phylogroup may follow a common anti-
518 phage strategy, while relying on distinct sets of defence systems to implement it. For strains
519 belonging to phylogroup PG10, we observed a negative correlation between the number of
520 defence systems and sensitivity to phages, suggesting a protective effect from accumulating
521 diverse systems in their genome (17). Conversely, strains in PG07, although generally the most
522 resistant to phages, showed a positive correlation between bacterial sensitivity to phages and
523 defence systems. This counter-intuitive pattern could reflect historical predation by phages,

524 where strains subjected to intense phage attack would have accumulated defence systems
525 (59). Alternatively, the apparent lack of protection may arise because these systems target
526 phages or mobile genetic elements (MGEs) other than those tested here. However, it is
527 important to interpret these trends cautiously, as they are based on relatively small sample
528 sizes (n=6 for PG07; n=10 for PG10), and our understanding remains partial given the continual
529 discovery of novel defence systems (60).

530 *Limitations and alternative approaches for understanding phage-bacterium*
531 *interactions*

532 To comprehensively assess the impact of defence systems on phage infection, measuring
533 phage adsorption rates on bacteria would provide valuable insights (14,17). Successful
534 adsorption without subsequent lysis may indicate the action of post-adsorption defence
535 mechanisms, whereas failure of adsorption suggests that surface receptor compatibility,
536 rather than intracellular defences, governs the interaction.

537 Analysis of bacterial membrane receptors recognised by phages is therefore essential in
538 interpreting host range patterns (61,12). Such factors have enabled accurate prediction of
539 phage-bacteria interactions at the strain level in *E. coli* (14). However, receptors for phages
540 infecting *P. syringae* remain poorly characterised compared to other bacterial models, with
541 only three distinct receptors identified in the *P. syringae* phage receptor database compared
542 to over 50 in *E. coli* (62). Increasing research on phage–phytopathogen interactions will help
543 to close this knowledge gap and improve our understanding of the molecular basis underlying
544 the observed interaction patterns.

545 A complementary strategy to uncover bacterial genetic determinants of sensitivity to phages
546 and resistance phenotypes involves genotype–phenotype association analyses, including

547 approaches based on gene presence–absence matrices, phylogenetic trait correlation
548 methods and genome-wide association studies (GWAS). This latter method has recently been
549 applied to 80 *P. aeruginosa* strains, identifying host genes that influence susceptibility to two
550 phages, particularly those related to flagella and LPS/O-antigens associated with
551 lipopolysaccharide biosynthesis (63). While our dataset of 44 *P. syringae* genomes may
552 provide a sufficient basis for GWAS, challenges remain. The high genomic diversity of the *P.*
553 *syringae* species complex and the large number of phages tested (n = 23), each potentially
554 interacting with different receptors or resistance pathways, add considerable complexity.
555 Future studies may address this by focusing on specific phage-host combinations or by
556 applying comparative genomics approaches that can help disentangle the genetic basis of
557 resistance across diverse strains.

558 *Applied perspectives and future directions*

559 From an applied perspective, we identified several generalist phages capable of infecting a
560 wide range of *P. syringae* species complex strains. Understanding the molecular basis of their
561 generalism could provide important phage evolutionary clues, *e.g.,* whether generalism
562 involves flexible receptor-binding proteins, the ability to evade diverse bacterial defence
563 systems, or other mechanisms. Many *P. syringae* strains are non-pathogenic members of the
564 plant microbiome, highlighting the need for targeted phage cocktails that selectively control
565 pathogenic strains while preserving beneficial microbes. Such specificity offers a more
566 sustainable and ecologically-balanced approach to biocontrol compared to using broad-
567 spectrum generalist phages (64). This exemplifies how fundamental ecological understanding
568 of phage–bacterium infection networks can directly inform and enhance the development of
569 effective biocontrol strategies.

570 To further test our hypothesis that ecological connectivity shapes the nested structure of
571 phage–*P. syringae* interactions, future studies should isolate phages from diverse non-
572 agricultural environmental, such as water and healthy plants. These newly isolated phages
573 should then be tested against the same bacterial strains used in our study, or against new
574 strains from corresponding habitats. If similar nested network structures emerge, it would
575 strengthen the idea that phage and *P. syringae* populations continuously migrate between
576 agricultural and non-agricultural environments, forming an ecologically-connected meta-
577 community. Moving forward, several key research priorities should guide future
578 investigations: verification of the role of specific defence systems through genetic
579 manipulation like gene deletion and heterologous expression; systematic investigation of
580 receptor usage by our studied phages to reveal molecular determinants of host range; and
581 strategic expansion of the bacterial strain collection to improve phylogenetic and ecological
582 coverage, with particular focus on understudied phylogroups. These complementary
583 approaches would collectively strengthen our understanding of the molecular and ecological
584 factors driving the phage-bacteria interactions observed in this system. Extending such
585 analyses to other ecologically and phylogenetically-distinct plant-associated bacteria, such as
586 *Xanthomonas*, *Ralstonia*, or *Agrobacterium*, could reveal whether the patterns observed in *P.*
587 *syringae*-phages infection network are generalisable or reflect unique evolutionary dynamics
588 of this highly diverse species complex.

589

590 ## Materials and Methods

591 *Phages panel*

592 The 23 phages infecting strains of the *P. syringae* species complex used in this study were
593 isolated from soil samples collected in apricot orchards affected by bacterial canker and fully
594 described in a previous study (65). Briefly, this phage collection exhibits high genetic diversity,
595 encompassing 13 genera and 17 species, representing the three most common morphological
596 types (myovirus, podovirus, and siphovirus) and comprising 20 virulent and 3 temperate
597 phages. The collection included two characterised DNA packaging mechanisms (direct
598 terminal repeat (DTR) and headful) alongside two newly identified but unverified mechanisms,
599 while the remaining packaging strategies remain unknown. The genomic GC content of the
600 phages ranged from 46% to 58%. All phage stocks were standardised to a concentration of 10^7
601 PFU/mL and stored in SM buffer (100 mM NaCl, 8 mM $MgSO_4 \cdot 7H_2O$, 50 mM Tris-HCl, pH 7.4)
602 at 4°C.

603 *Bacterial strains panel*

604 The 44 *P. syringae* species complex strains were obtained from the Plant Pathology Unit
605 collection at INRAE (Montfavet, France) and were not isolated as part of this project (S4 Table).
606 The strains belonged to seven clades within four phylogroups (PG01, PG02, PG07, PG10),
607 corresponding to the species and genus-like levels, respectively, according to Berge *et al.* (46).
608 The strains were sampled from three different environments: diseased apricot trees
609 (exhibiting bacterial canker symptoms) (45), healthy plant tissues, and non-agricultural
610 environments (S6 Table). Bacterial strains were revived from -80°C storage and cultured on
611 King's B agar (66). Cultures were maintained at 4°C for a maximum of one week.

612 *Sequencing and analysis of whole genomes of bacterial strains*

613 The bacterial cultures were grown on King's B (KB) agar medium for 48 hours at 25°C. A
614 bacterial suspension was then prepared to achieve an approximate concentration of 2×10^9
615 CFU in 4 mL of demineralised water. The suspension was subsequently centrifuged at 3,215 *g*
616 for 10 minutes at 20°C to pellet the bacterial cells, after which the supernatant was carefully
617 discarded. Genomic DNA was extracted using the Qiagen Genomic-tips kit (Qiagen, Valencia,
618 CA, USA) following the manufacturer's protocol.

619 Of the 44 *P. syringae* strains, 39 were sequenced using 20 µL of genomic DNA, with sequencing
620 performed by Plasmidsaurus Inc (San Francisco, USA) using Oxford Nanopore Technology
621 (ONT) and a custom analysis and annotation pipeline (https://plasmidsaurus.com/) (S7 Table).
622 The long-read sequencing library was prepared without amplification using v14 library
623 preparation chemistry, and sequencing was conducted with R10.4.1 flow cells. The genomes
624 of five additional strains were obtained from the RefSeq library at NCBI: CC1559
625 (GCF_000452685.1), USA0007 (GCF_000452545.1), 41A (GCF_000935775.1), MAFF302273^{PT}
626 (also known as JD08) (GCF_000145925.1), and CC1583 (GCF_000452665.1).

627 The core genome analysis was performed using a private database within the EDGAR 3.5
628 platform (67). The core genome was defined as the set of orthologous genes present in all 332
629 genomes of the *P. syringae* species complex, including the 44 strains analysed in this study. A
630 total of 1,346 core genes were identified per genome, resulting in a total of 446,872 aligned
631 genes across 332 genomes. Each set of orthologous genes was aligned individually at the
632 protein level using MUSCLE (67). The resulting alignment were concatenated into a single
633 alignment, resulting in a total of 512,223 amino acid residues per genome and 170,058,036
634 residues across all genomes. Phylogenetic relationships were inferred using the FastTree

635 algorithm within EDGAR, which constructs maximum-likelihood phylogenetic trees based on
636 the core genome alignment (68). All phylogenetic trees of bacterial strains shown in the figures
637 were rooted at the midpoint of the longest path and the nodes with a maximum likelihood
638 support value below 0.90 were marked with red dots.

639 *Identification of defence systems and prophages in bacterial genomes*

640 The defence systems of the 44 *P. syringae* strains were identified using *DefenseFinder* version
641 1.2.2 (69) based on assembled nucleotide sequences from the bacterial genomes. Each
642 defence system type was categorised into one of three functional groups: nucleic acid
643 degradation, abortive infection (Abi), or undetermined (ND).

644 The prophages within the 44 *P. syringae* strains were identified using *geNomad* version 1.5.0
645 (70) based on assembled nucleotide sequences of bacterial genomes. Among the 97 identified
646 prophage sequences, only those classified as viral by *geNomad* and *CheckV* (71) were retained.
647 Specifically, sequences labelled as "No Terminal Repeats" in the topology column and those
648 with a "viral_genes" value of 0 were excluded. This filtering resulted in a final dataset of 88
649 verified viral prophages, accounting for 91% of the total identified prophages across the 44
650 strains. Additionally, a genetic analysis of the prophages was performed using *VIRIDIC* (72) to
651 classify them into genera and species (S3 Table). *VipTree* was used to generate prophage
652 clusters based on a proteomic tree analysis using distance-based methods (BIONJ from
653 tBLASTx scores) (73). *geNomad* and *VipTree* were then employed to identify prophages
654 belonging to the CL4 cluster and to construct a proteomic tree encompassing 68 CL4
655 prophages (24 reference strains + 44 analysed strains) (74).

656 To assess the congruence between the taxonomy of CL4 prophages and the phylogeny of their
657 bacterial hosts, we performed a PACo (Procrustes Approach to Cophylogeny) analysis.

658 Phylogenetic distance matrices were generated from the host and prophage trees using the
659 *cophenetic* function (75). These matrices were then centred and corrected for negative
660 eigenvalues using the "Cailliez" correction method. Congruence was tested using the *PACo*
661 function from the *paco* package with 1,000 permutations to assess statistical significance. The
662 co-phylogeny visualisation was performed with *phytools* package and *cophylo* function (76).

663 *Phage host range analysis*

664 *a. Experimental setup*

665 To analyse the host range of phages infecting *P. syringae*, bacterial growth kinetics in the
666 presence of phages were measured using 96-well plates (77). The dataset comprised 44
667 bacterial strains tested in biological triplicate without phage as a positive control, and the
668 same 44 strains in biological triplicate exposed individually to each of the 23 phages (S6 Fig).
669 Each phage was tested against each biological triplicate of bacterial strain, resulting in 3,036
670 growth kinetic measurements (23 phages × 44 bacterial strains × 3 replicates) (S6 Fig).
671 Negative controls (culture medium) were used to check for contamination, with plates
672 containing more than two contaminated wells out of four discarded (S6 Fig and S7A Fig). The
673 combination of the two internal controls were added to each plate: the *P. syringae* strain
674 P1.01.01B03 (PG01a, isolated from apricot buds in France in 2014) and phage Ghuch01,
675 chosen because it completely inhibited the growth of strain P1.01.01B03 under the
676 experimental conditions of this study (45,65) (S6 Fig). A plate was considered valid if the two
677 internal controls had all three replicates above the first quartile threshold of bacterial growth
678 and growth inhibition (S7B-C Fig).

679 Bacteria were incubated in 5 mL of liquid KB medium at 25°C and 200 RPM for 4 hours. Each
680 replicate was standardised to an OD_{600nm} = 0.01 (~10^7 CFU/mL) in phosphate buffer and diluted

681 in liquid KB medium to a final concentration of 10^5 CFU/mL. Each 96-well spectrophotometric
682 plate (NUNC, Thermo Scientific™ 267334) contained most phages confronted with a single
683 bacterial strain, with positive control wells containing 180 µL of bacterial suspension and 20
684 µL of SM buffer. Experimental wells comprised 180 µL of bacterial suspension and 20 µL of
685 phage suspension at 10^7 PFU/mL, establishing a multiplicity of infection (MOI) of 10. Positive
686 control measurements included between two and seven technical replicates per biological
687 replicate (except for bacterial strains 41A and 93D, which had only one technical replicate per
688 biological replicate). Measurements were taken using the SPECTROstar® Omega
689 spectrophotometer (BMG Labtech, France) every 5 minutes for 48 hours at 25°C, with
690 intermittent orbital shaking at 200 RPM (200 seconds "ON" followed by 400 seconds "OFF").

691 To assess the host specificity of the 23 phages, their infectivity was tested on four
692 *Pseudomonas* spp. strains outside the *P. syringae* species complex using the double agar layer
693 method on solid medium. The selected strains were *P. aeruginosa* CFBP 2466, *Pseudomonas*
694 *fluorescens* CFBP 2102, *Pseudomonas graminis* 38B9, and *Pseudomonas rhizosphaerae* 6B4.
695 Strains were grown on King's B (KB) agar for 48 h at 25°C (*P. fluorescens*, *P. graminis*) or 28°C
696 (*P. aeruginosa*, *P. rhizosphaerae*). Biological triplicates were inoculated into 5 mL of KB broth
697 and incubated overnight at 25°C with shaking at 200 rpm. Then, 500 µL of each culture was
698 transferred to 5 mL fresh KB and incubated under the same conditions for 4 h. The optical
699 density was adjusted to 0.2 to 0.8 at 600 nm, and 300 µL of the adjusted culture was mixed
700 with 6 mL of 0.7% KB semi-solid agar for the overlay. Each phage at an initial concentration of
701 10^7 PFU/mL was serially diluted, and 10 µL of each dilution were spotted onto the bacterial
702 lawn. Plates were allowed to dry and incubated overnight at 25°C before evaluation of plaque
703 lysis.

704 *b. Calculation of bacterial growth inhibition percentages and probability of*
705 *significant inhibitory effects of phages on bacteria*

706 Bacterial growth curve data from each plate were analysed using R software (version 4.4.0)
707 with the *growthcurver* package (78). The area under the empirical growth curve (auc*e*) was
708 the primary parameter analysed. For each biological replicate of the bacterial control
709 condition (without phage), the mean and standard deviation of the auc*e* were calculated. The
710 percentage of bacterial growth inhibition caused by each phage was calculated for each
711 biological replicate of bacteria exposed to a given phage using the following formula:

712 $$Bacterial\ growth\ inhibition(\%) = \frac{auce\ bacteria - auce\ phage}{auce\ bacteria} \times 100$$

713 Where auc*e* bacteria is the area under the curve of the empirical growth curve of each
714 bacterial replicate, and auc*e* phage is the area under the empirical growth curve of the tested
715 bacteria with each phage. This calculation standardises differences in bacterial growth rates,
716 allowing for a direct comparison of phage effects across different bacterial strains.

717 Phage impact on bacterial growth was assessed using a probability value derived from a
718 normalised distribution, calibrated based on the mean and standard deviation of the auc*e* for
719 each biological replicate of each bacterium exposed to phages. The probability ranges from 0
720 to 1 and is calculated using:

721 $$auce\ phage \sim N(\mu(aucebacteria), \sigma(aucebacteria))$$

722 A probability value below 0.05 indicates a significant growth reduction in presence of phage,
723 while a value exceeding 0.95 indicates a significant growth stimulation of bacteria in presence
724 of phage. The distribution of inhibition percentages within the 0.05 to 0.95 probability range
725 is approximately symmetric and unimodal (S8 Fig), resembling a Gaussian distribution.

726 However, the Shapiro-Wilk test indicated a significant deviation from strict normality ($p <$
727 0.01), likely due to the large sample size. Despite this, the distribution shape supports the use
728 of this probability window to robustly summarise phage effects on bacterial growth. Bacterial
729 strains 41A and 93D had only one biological replicate, precluding probability calculation.

730 Generally, bacterial populations decreased in the presence of phages. In 7% of observed
731 phage-bacteria combinations, a "stimulatory" effect was identified, with inhibition values
732 reaching up to -40%. Given the lack of reproducibility of the stimulatory effect (S9 Fig), only
733 positive inhibition values were retained for all analyses.

734 *Statistical analysis*

735 *a. Analysis of the structure of phage-bacteria interaction networks*

736 The analysis of networks representing bacterial growth inhibition percentages of 44 bacterial
737 strains combined with 23 phages was computed using the mean of the three biological
738 replicates for each combination. The analyses were performed using the R software (version
739 4.4.0) with the *bipartite* (for nestedness and modularity analysis) and *igraph* (for modularity
740 analysis) packages, following Moury *et al.* (32) and Torres-Barceló *et al.* (48).

741 The nestedness of the inhibition percentage network was evaluated using two algorithms:
742 *wNODF* (a weighted nestedness metric based on overlap and decreasing fill) (79) and *WINE*
743 (weighted-interaction nestedness estimator) (80). Modularity was assessed using six
744 algorithms: *spinglass* (81–83), *edge betweenness* (81), *fast greedy* (84), *leading eigenvector*
745 (85), *Beckett* (86), and *louvain* (87).

746 Prior to nestedness and modularity analyses, the continuous inhibitory percentage data of the
747 matrix were transformed into discrete integer values spanning from 0 to 100.

748 In addition, we analysed a network of binary data (0 and 1 values) representing the statistical
749 significance of bacterial growth inhibition by phages. A phage was considered capable of
750 inhibiting a bacterial strain if at least one of the three bacterial replicates was significantly
751 inhibited (probability < 0.05); in this case, the phage-bacterium combination was assigned a
752 value of 1. If no replicates showed significant inhibition, the combination was assigned a value
753 of 0. For this network, three algorithms were used for nestedness: *WINE* as applied to
754 quantitative data, *Binmatnest* (88), and *NODF2* (79) and three algorithms were used for
755 modularity (*spinglass, leading eigenvector,* and *Beckett)*.

756 For each algorithm and each network, the calculated nestedness and modularity scores were
757 compared with the scores obtained with seven (network of continuous inhibition data) or six
758 (network of binary inhibition significance data) different null models (i.e. simulated networks
759 for which no particular nestedness or modularity patterns are expected) in order to estimate
760 the statistical significance of nestedness and modularity of the actual networks. Null models
761 are described in Moury *et al.* (89,32).

762 *b. Statistical analyses to explain bacterial sensitivity and susceptibility to phages*

763 All statistical analyses were conducted in R version 4.4.2. The phylogenetic signal of the
764 number of phages infecting each bacterial strain was calculated using the *phylosig* function
765 from the *phytools* package, based on a phylogenetic tree reconstructed from bacterial core-
766 genome analysis (76).

767 The bacterial susceptibility profiles were tested with a PERmutational Multivariate ANalysis Of
768 VAriance (PERMANOVA), a non-parametric method that assesses the influence of categorical
769 variables on multivariate data by partitioning variation within a distance matrix. Unlike
770 traditional ANOVA, it makes no assumptions about data distribution and relies on

771 permutations to evaluate the statistical significance of group differences (90). Here, Bray–
772 Curtis dissimilarity was used to construct a distance matrix based on the mean percentage of
773 bacterial growth inhibition (from three replicates per combination). This matrix was then
774 analysed using the *adonis2* function from the *vegan* R package, with 9,999 permutations to
775 ensure robust p-value estimation (91). "*Phylogroup", "ecological origin",* and "*whether the*
776 *strain was an isolation bacterial strain for the phages*" were included as explanatory variables.
777 Multiple comparisons were conducted using the function *pairwise.adonis2* from the
778 *pairwiseAdonis* package (92), which is specifically adapted for PERMANOVA post-hoc testing.

779 Additionally, the summed sensitivity, determined from the significant inhibition network (see
780 above), was used as a response variable in a Generalised Linear Model (GLM) with a Poisson
781 distribution with the function *glmmTMB* from the package *glmmTMB* (93). The model
782 included "*phylogroup"*, "*ecological origin"*, and their interaction as fixed effects. As no
783 significant interaction was detected, the simplified additive model (*phylogroup + ecological*
784 *origin*) was retained. No overdispersion was observed in the data (deviance/degrees of
785 freedom = 1.01). The effects of each explanatory variable were assessed using Type III ANOVA
786 with the *Anova* function from the *car* package. Post-hoc pairwise comparisons were
787 performed using Tukey's test implemented via the *multcomp* and *multcompView* packages,
788 with significance groups assigned accordingly.

789 *c. Statistical analysis to explain infectivity and virulence of phages*

790 The Phylogenetic Host Range Index (PHRI), derived from the phylogenetic diversity index *Hp*
791 (94), was calculated using the *hilldiv* package (48). For this analysis, the bacterial phylogenetic
792 tree, constructed from core-genome analysis (see above), was transformed into an ultrametric
793 tree, and the PHRI index was determined using the matrix of percentage inhibition.

794 The ranking of phage genera based on inhibition percentage (offset by +0.1 to avoid zeros)
795 was assessed using a GLMM with a Gamma distribution with a log-link, suitable for strictly
796 positive, continuous, and skewed data. The model included a random effect accounting for
797 the "*phage*" tested, the "*bacterial*" strain tested, and "*replicates*" nested within each
798 "*phage*:*bacterium*" combination. No overdispersion was detected (χ^2/ddl = 0.67). Post-hoc
799 pairwise comparisons between genera were performed using Tukey's test.

800 The phage-host infection profiles were tested with a PERMANOVA test based on the distance
801 matrix based on the mean percentage of bacterial growth inhibition compared with "*phage*
802 *genus*" and "*whether the phage had been isolated using one of the bacterial strains included*
803 *in the experimental panel*" explanatory variables (Bray-Curtis distance matrix; 9,999
804 permutations). Pairwise comparisons of phage infection patterns were performed using the
805 *pairwise.adonis2* function (92).

806 *d. Statistical analysis of defence systems and prophages*

807 The phylogenetic signal of the total number of defence systems and the abundance of
808 prophages was calculated using the *phylosig* function from the *phytools* package, based on a
809 phylogenetic tree reconstructed from bacterial core-genome analysis (76).

810 The total number of bacterial defence systems per bacterial genome and the abundance of
811 prophages were tested against the explanatory variables "*phylogroup*", "*ecological origin*",
812 and "*growth rate*" in a GLM with a Poisson distribution with an additive model of the three
813 explanatory variables. No overdispersion was observed in the data (deviance/degrees of
814 freedom for defence systems = 0.61; for abundance of prophages = 0.47), and the effects of
815 each explanatory variable were assessed using Type III ANOVA. Finally, the post-hoc pairwise

816 comparisons were performed using Tukey's test and significance groups were assigned
817 accordingly.

818 The defence system profiles, the proportions of the defence system categories, and prophages
819 distribution profile across bacterial strains were analysed using a PERMANOVA test based on
820 a Bray-Curtis dissimilarity matrix with 9,999 permutations. "*Phylogroup*", "*ecological origin*",
821 and "*growth rate*" were included as explanatory variables. Pairwise comparisons of the
822 defence system and prophage composition between groups were conducted using the
823 *pairwise.adonis2* function (92).

824 Linear regressions were conducted to explain the number of phages infecting each bacterial
825 strain, using the total number of defence system subtypes and the number of prophages per
826 bacterium. Before conducting these analyses, all statistical test assumptions associated to the
827 linear regressions were verified.

828 *e. Statistical analysis of host range data with multimodel analysis*

829 To account for the widest range of explanatory variables influencing the percentage of
830 inhibition, we performed a multimodel analysis coupled with a generalised linear mixed model
831 (GLMM). To enhance model efficiency, the number of explanatory variables was reduced
832 through a multi-step process. Variables with more than 20% missing data were excluded.
833 Redundant variables were then removed based on their correlations, assessed using
834 Spearman's correlation for quantitative variables, chi-squared tests for categorical variables,
835 and Kruskal-Wallis tests for associations between categorical and quantitative variables. To
836 further refine the selection, Principal Component Analyses (PCA) were conducted separately
837 on the matrices representing the distributions of defence systems and prophages across

838 bacterial genomes. The most informative variables contributing to the first principal
839 components were retained for subsequent modelling.

840 The PCA on defence systems was performed without centring or scaling the data. Two
841 dimensions accounted for 32.5% of the variance and led to the selection of the variables
842 "*Septu*", "*Gabija*", and "*RloC*", which were among the most frequent defence systems in the
843 genomes. Similarly, the PCA on the distribution of prophage species in bacterial genomes, also
844 conducted without centring or scaling (already binary), captured 14% of the variance in the
845 first two dimensions and identified two perpendicular variable groups composed respectively
846 of prophage species "*17*", "*54*", "*26*", and "*29*" on one side, and species "*7*", "*47*", "*57*", "*22*",
847 and "*31*" on the other. One representative species from each group was selected for the
848 multimodel analysis (species "*17*" and "*22*").

849 To identify factors explaining variations in bacterial growth inhibition percentage, we fitted a
850 GLMM. The response variable (offset by +0.1 to avoid zeros) was modelled using a Gamma
851 distribution with a log-link, suitable for strictly positive, continuous, and skewed data. Random
852 effects included "*phage*", "*bacterium*", and their "*imbricated replicates*".

853 To compare two approaches for modelling the effects of bacterial defence systems and
854 prophages on the percentage of inhibition, two separate GLMMs were constructed: one using
855 total counts of defence systems and prophages, and another incorporating principal
856 components PCA-derived of the distribution matrices (of prophages and defence systems). No
857 overdispersion was detected in either model (total counts: $\chi^2/df = 0.65$; PCA-derived: $\chi^2/df =$
858 0.66), supporting the appropriateness of the Gamma distribution for modelling the response
859 variable.

To identify the most informative predictors explaining variation in phage replication efficiency, we used a multimodel analysis. Model selection was performed using the *dredge* function from the *MuMIn* package, ranking models based on corrected Akaike information criterion (AIC). All possible sub-models were generated and evaluated. Models whose cumulative AIC weights accounted for 95% of the total were retained for model averaging using the *model.avg* function. The relative importance of each explanatory variable was then estimated using the *sw* function, which computes the sum of the AIC weights across all models in which the variable appears.

Acknowledgments

We thank Elisabeth Villard, Cécile Monteil, and Charlotte Chandeysson for their assistance with experimental procedures. We are also grateful to Dominique Holtappels for valuable discussions, and to Hannah Toth for her assistance in depositing bacterial genomes into NCBI.

875 **References**

876 1. Brown TL, Charity OJ, Adriaenssens EM. Ecological and functional roles of bacteriophages
877 in contrasting environments: marine, terrestrial and human gut. Curr Opin Microbiol. 2022
878 Dec 1;70:102229.

879 2. Wilhelm SW, Suttle CA. Viruses and Nutrient Cycles in the Sea: Viruses play critical roles in
880 the structure and function of aquatic food webs. BioScience. 1999 Oct 1;49(10):781–8.

881 3. Wommack KE, Colwell RR. Virioplankton: Viruses in Aquatic Ecosystems. Microbiol Mol
882 Biol Rev. 2000 Mar;64(1):69–114.

883 4. Suttle CA. Marine viruses — major players in the global ecosystem. Nat Rev Microbiol.
884 2007 Oct;5(10):801–12.

885 5. Brussaard CPD, Wilhelm SW, Thingstad F, Weinbauer MG, Bratbak G, Heldal M, et al.
886 Global-scale processes with a nanoscale drive: the role of marine viruses. ISME J. 2008
887 June;2(6):575–8.

888 6. Mirzaei MK, Maurice CF. Ménage à trois in the human gut: interactions between host,
889 bacteria and phages. Nat Rev Microbiol. 2017 July;15(7):397–408.

890 7. De Sordi L, Lourenço M, Debarbieux L. The Battle Within: Interactions of Bacteriophages
891 and Bacteria in the Gastrointestinal Tract. Cell Host Microbe. 2019 Feb;25(2):210–8.

892 8. Svircev A, Roach D, Castle A. Framing the future with bacteriophages in agriculture.
893 Viruses. 2018 May;10(5):218.

894 9. Hyman P, Abedon ST. Bacteriophage host range and bacterial resistance. In: Advances in
895 Applied Microbiology [Internet]. Elsevier; 2010 [cited 2024 Aug 12]. p. 217–48. Available
896 from: https://linkinghub.elsevier.com/retrieve/pii/S0065216410700071

897 10. Georjon H, Bernheim A. The highly diverse antiphage defence systems of bacteria. Nat Rev
898 Microbiol. 2023 Oct;21(10):686–700.

899 11. Piel D, Bruto M, Labreuche Y, Blanquart F, Goudenège D, Barcia-Cruz R, et al. Phage–host
900 coevolution in natural populations. Nat Microbiol. 2022 July;7(7):1075–86.

901 12. Borin JM, Lee JJ, Lucia-Sanz A, Gerbino KR, Weitz JS, Meyer JR. Rapid bacteria-phage
902 coevolution drives the emergence of multiscale networks. Science. 2023 Nov
903 10;382(6671):674–8.

904 13. Beamud B, García-González N, Gómez-Ortega M, González-Candelas F, Domingo-Calap P,
905 Sanjuan R. Genetic determinants of host tropism in *Klebsiella* phages. Cell Rep. 2023 Feb
906 6;42(2):112048.

907 14. Gaborieau B, Vaysset H, Tesson F, Charachon I, Dib N, Bernier J, et al. Prediction of strain
908 level phage–host interactions across the *Escherichia* genus using only genomic
909 information. Nat Microbiol. 2024 Nov;9(11):2847–61.

15. Johnson MC, Laderman E, Huiting E, Zhang C, Davidson A, Bondy-Denomy J. Core defense hotspots within *Pseudomonas aeruginosa* are a consistent and rich source of anti-phage defense systems. Nucleic Acids Res. 2023 June 9;51(10):4995–5005.

16. Burke KA, Urick CD, Mzhavia N, Nikolich MP, Filippov AA. Correlation of *Pseudomonas aeruginosa* phage resistance with the numbers and types of antiphage systems. Int J Mol Sci. 2024 Jan;25(3):1424.

17. Costa AR, van den Berg DF, Esser JQ, Muralidharan A, van den Bossche H, Bonilla BE, et al. Accumulation of defense systems in phage-resistant strains of *Pseudomonas aeruginosa*. Sci Adv. 2024 Feb 23;10(8):eadj0341.

18. Hussain FA, Dubert J, Elsherbini J, Murphy M, VanInsberghe D, Arevalo P, et al. Rapid evolutionary turnover of mobile genetic elements drives bacterial resistance to phages. Science. 2021 Oct 22;374(6566):488–92.

19. Czajkowski R. May the phage be with you? Prophage-like elements in the genomes of soft rot Pectobacteriaceae: *Pectobacterium* spp. and *Dickeya* spp. Front Microbiol [Internet]. 2019 Feb 14 [cited 2025 Apr 28];10. Available from: https://www.frontiersin.orghttps://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2019.00138/full

20. Dominguez-Mirazo M, Jin R, Weitz JS. Functional and comparative genomic analysis of integrated prophage-like sequences in "*Candidatus Liberibacter asiaticus*". mSphere. 2019 Nov 13;4(6):10.1128/msphere.00409-19.

21. Castillo JA, Secaira-Morocho H, Maldonado S, Sarmiento KN. Diversity and evolutionary dynamics of antiphage defense systems in *Ralstonia solanacearum* species complex. Front Microbiol [Internet]. 2020 May 20 [cited 2024 Sept 24];11. Available from: https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2020.00961/full

22. Greenrod STE, Stoycheva M, Elphinstone J, Friman VP. Global diversity and distribution of prophages are lineage-specific within the *Ralstonia solanacearum* species complex. BMC Genomics. 2022 Oct 6;23(1):689.

23. Dougherty PE, Nielsen TK, Riber L, Lading HH, Forero-Junco LM, Kot W, et al. Widespread and largely unknown prophage activity, diversity, and function in two genera of wheat phyllosphere bacteria. ISME J. 2023 Dec;17(12):2415–25.

24. Holtappels D, Rickus GEJ, Morgan T, Rezende RR de, Koskella B, Alfenas-Zerbini P. Evidence for the key roles of the *Pseudomonas syringae* mobilome in shaping biotic interactions [Internet]. bioRxiv; 2024 [cited 2024 Aug 20]. p. 2024.03.19.585818. Available from: https://www.biorxiv.org/content/10.1101/2024.03.19.585818v2

25. Jimenez D, Beltran D, Castillo JA. Mobile genetic elements of *Xylella fastidiosa* and their contribution to pathogenicity. Plant Pathol. 2024 Dec;73(9):2490–9.

947 26. Flores CO, Meyer JR, Valverde S, Farr L, Weitz JS. Statistical structure of host-phage
948 interactions. Proc Natl Acad Sci U S A. 2011 July 12;108(28):E288–97.

949 27. Gupta A, Peng S, Leung CY, Borin JM, Medina SJ, Weitz JS, et al. Leapfrog dynamics in
950 phage-bacteria coevolution revealed by joint analysis of cross-infection phenotypes and
951 whole genome sequencing. Ecol Lett. 2022;25(4):876–88.

952 28. Molina F, Menor-Flores M, Fernández L, Vega-Rodríguez MA, García P. Systematic analysis
953 of putative phage-phage interactions on minimum-sized phage cocktails. Sci Rep. 2022
954 Feb 14;12(1):2458.

955 29. Kauffman KM, Chang WK, Brown JM, Hussain FA, Yang J, Polz MF, et al. Resolving the
956 structure of phage–bacteria interactions in the context of natural diversity. Nat Commun.
957 2022 Jan 18;13(1):372.

958 30. Flores CO, Valverde S, Weitz JS. Multi-scale structure and geographic drivers of cross-
959 infection within marine bacteria and phages. ISME J. 2013 Mar;7(3):520–32.

960 31. Cahier K, Piel D, Barcia-Cruz R, Goudenège D, Wegner KM, Monot M, et al. Environmental
961 *vibrio* phage-bacteria interaction networks reflect the genetic structure of host
962 populations. Environ Microbiol [Internet]. 2023 Mar 6 [cited 2023 Mar 9];n/a(n/a).
963 Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/1462-2920.16366

964 32. Moury B, Audergon JM, Baudracco-Arnas S, Ben Krima S, Bertrand F, Boissot N, et al. The
965 quasi-universality of nestedness in the structure of quantitative plant-parasite
966 interactions. Peer Community J. 2021 Nov 30;1:e44.

967 33. Morris CE, Monteil CL, Berge O. The life history of *Pseudomonas syringae*: linking
968 agriculture to earth system processes. Annu Rev Phytopathol. 2013;51:85–104.

969 34. Morris CE, Sands DC, Vinatzer BA, Glaux C, Guilbaud C, Buffière A, et al. The life history of
970 the plant pathogen *Pseudomonas syringae* is linked to the water cycle. ISME J. 2008
971 Mar;2(3):321–34.

972 35. Morris CE, Lamichhane JR, Nikolić I, Stanković S, Moury B. The overlapping continuum of
973 host range among strains in the *Pseudomonas syringae* complex. Phytopathol Res. 2019
974 Jan 16;1(1):4.

975 36. Morris CE, Sands DC, Vanneste JL, Montarry J, Oakley B, Guilbaud C, et al. Inferring the
976 evolutionary history of the plant pathogen *Pseudomonas syringae* from its biogeography
977 in headwaters of rivers in North America, Europe, and New Zealand. mBio. 2010 June
978 29;1(3):e00107-10.

979 37. Thomas MD, Leary JV. Bacteriophages from sewage specific for fluorescent phytogenic
980 Pseudomonads. Phytopathology. 1983 Sept;(73):403–6.

981 38. Cuppels DA. Isolation and characterization of phages useful for identifying *Pseudomonas*
982 *syringae pv. tomato*. Phytopathology. 1983 Apr;(73):1376–81.

39. Gross DC, Cody YS, Proebsting EL, Radamaker GK, Spotts RA. Ecotypes and pathogenicity of ice-nucleation-active *Pseudomonas syringae* isolated from deciduous fruit tree orchards. 1983 Sept;(74):241–8.

40. Cuppels DA. The use of pathovar-indicative bacteriophages for rapidly detecting *Pseudomonas syringae pv. tomato* in tomato leaf and fruit lesions. 1984 Feb;(74):891–4.

41. Taylor JD, Teverson DM, Allen DJ, Pastor-Corrales MA. Identification and origin of races of *Pseudomonas syringae pv. phaseolicola* from Africa and other bean growing areas. Plant Pathol. 1996;45(3):469–78.

42. Frampton RA, Taylor C, Holguín Moreno AV, Visnovsky SB, Petty NK, Pitman AR, et al. Identification of bacteriophages for biocontrol of the kiwifruit canker phytopathogen *Pseudomonas syringae pv. actinidiae*. Appl Environ Microbiol. 2014 Apr;80(7):2216–28.

43. James SL, Rabiey M, Neuman BW, Percival G, Jackson RW. Isolation, characterisation and experimental evolution of phage that infect the horse chestnut tree pathogen, *Pseudomonas syringae pv. aesculi*. Curr Microbiol. 2020 Aug 1;77(8):1438–47.

44. Rombouts S, Volckaert A, Venneman S, Declercq B, Vandenheuvel D, Allonsius CN, et al. Characterization of novel bacteriophages for biocontrol of bacterial blight in leek caused by *Pseudomonas syringae pv. porri*. Front Microbiol [Internet]. 2016 Mar 15 [cited 2024 Oct 8];7. Available from: https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2016.00279/full

45. Parisi L, Morgaint B, Blanco-Garcia J, Guilbaud C, Chandeysson C, Bourgeay JF, et al. Bacteria from four phylogroups of the *Pseudomonas syringae* complex can cause bacterial canker of apricot. Plant Pathol. 2019 Sept;68(7):1249–58.

46. Berge O, Monteil CL, Bartoli C, Chandeysson C, Guilbaud C, Sands DC, et al. A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex. PLOS ONE. 2014 Sept 3;9(9):e105547.

47. Omrani M, Roth M, Roch G, Blanc A, Morris CE, Audergon JM. Genome-wide association multi-locus and multi-variate linear mixed models reveal two linked loci with major effects on partial resistance of apricot to bacterial canker. BMC Plant Biol. 2019 Jan 21;19(1):31.

48. Torres-Barceló C, Boyer C, Garneau JR, Poussier S, Robène I, Moury B. A Phylogenetic Host-Range Index Reveals Ecological Constraints in Phage Specialisation and Virulence. Mol Ecol. 2025 July 28;n/a(n/a):e70052.

49. van Houte S, Buckling A, Westra ER. Evolutionary ecology of prokaryotic immune mechanisms. Microbiol Mol Biol Rev. 2016 July 13;80(3):745–63.

50. Turner D, Kropinski AM, Adriaenssens EM. A roadmap for genome-based phage taxonomy. Viruses. 2021 Mar;13(3):506.

1019 51. Mavrodi DV, Loper JE, Paulsen IT, Thomashow LS. Mobile genetic elements in the genome
1020 of the beneficial rhizobacterium *Pseudomonas fluorescens* Pf-5. BMC Microbiol. 2009 Jan
1021 13;9(1):8.

1022 52. Hockett KL, Renner T, Baltrus DA. Independent Co-Option of a Tailed Bacteriophage into a
1023 Killing Complex in *Pseudomonas*. mBio. 2015 Aug 11;6(4):10.1128/mbio.00452-15.

1024 53. Backman T, Latorre SM, Symeonidi E, Muszyński A, Bleak E, Eads L, et al. A phage tail–like
1025 bacteriocin suppresses competitors in metapopulations of pathogenic bacteria. Science.
1026 2024 June 14;384(6701):eado0713.

1027 54. Fautt C, Hockett K, Delattre S, Baltrus D, Couradeau E. Tailocin tail fiber diversity correlates
1028 with rfbD variation in the *Pseudomonas syringae* species complex. ISME Commun. 2025
1029 Jan 1;5(1):ycaf099.

1030 55. Hunter M, Fusco D. Superinfection exclusion: A viral strategy with short-term benefits and
1031 long-term drawbacks. PLOS Comput Biol. 2022 May 10;18(5):e1010125.

1032 56. Koskella B, Lin DM, Buckling A, Thompson JN. The costs of evolving resistance in
1033 heterogeneous parasite environments. Proc R Soc B Biol Sci. 2011 Dec
1034 14;279(1735):1896–903.

1035 57. Monteil CL, Yahara K, Studholme DJ, Mageiros L, Méric G, Swingle B, et al. Population-
1036 genomic insights into emergence, crop adaptation and dissemination of *Pseudomonas*
1037 *syringae* pathogens. Microb Genomics [Internet]. 2016 Oct 1 [cited 2021 Sept 15];2(10).
1038 Available from:
1039 https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000089

1040 58. Koonin EV, Makarova KS, Wolf YI. Evolutionary genomics of defense systems in archaea
1041 and bacteria. Annu Rev Microbiol. 2017 Sept 8;71(Volume 71, 2017):233–61.

1042 59. Bernheim A, Sorek R. The pan-immune system of bacteria: antiviral defence as a
1043 community resource. Nat Rev Microbiol. 2020 Feb;18(2):113–9.

1044 60. Makarova KS, Wolf YI, Koonin EV. Comparative genomics of defense systems in archaea
1045 and bacteria. Nucleic Acids Res. 2013 Apr 1;41(8):4360–77.

1046 61. Nobrega FL, Vlot M, de Jonge PA, Dreesens LL, Beaumont HJE, Lavigne R, et al. Targeting
1047 mechanisms of tailed bacteriophages. Nat Rev Microbiol. 2018 Dec;16(12):760–73.

1048 62. Bertozzi Silva J, Storms Z, Sauvageau D. Host receptors for bacteriophage adsorption.
1049 FEMS Microbiol Lett. 2016 Feb 1;363(4):fnw002.

1050 63. Müller DM, Pourtois JD, Kim MK, Targ B, Burgener EB, Milla C, et al. Bacterial receptors
1051 but not anti-phage defence mechanisms determine host range for a pair of *Pseudomonas*
1052 *aeruginosa* lytic phages [Internet]. bioRxiv; 2024 [cited 2024 Sept 23]. p.
1053 2024.04.30.591980. Available from:
1054 https://www.biorxiv.org/content/10.1101/2024.04.30.591980v2

1055 64. Kering KK, Kibii BJ, Wei H. Biocontrol of phytobacteria with bacteriophage cocktails. Pest
1056 Manag Sci. 2019;75(7):1775–81.

1057 65. Feltin C, Garneau JR, Morris CE, Bérard A, Torres-Barceló C. Novel phages of *Pseudomonas*
1058 *syringae* unveil numerous potential auxiliary metabolic genes. J Gen Virol.
1059 2024;105(6):001990.

1060 66. King EO, Ward MK, Raney DE. Two simple media for the demonstration of pyocyanin and
1061 fluorescin. J Lab Clin Med. 1954 Aug 1;44(2):301–7.

1062 67. Dieckmann MA, Beyvers S, Nkouamedjo-Fankep RC, Hanel PHG, Jelonek L, Blom J, et al.
1063 EDGAR3.0: comparative genomics and phylogenomics on a scalable infrastructure. Nucleic
1064 Acids Res. 2021 July 2;49(W1):W185–92.

1065 68. Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum-likelihood trees for
1066 large alignments. PLOS ONE. 2010 Mar 10;5(3):e9490.

1067 69. Tesson F, Hervé A, Mordret E, Touchon M, d'Humières C, Cury J, et al. Systematic and
1068 quantitative view of the antiviral arsenal of prokaryotes. Nat Commun. 2022 May
1069 10;13(1):2561.

1070 70. Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, Hu B, et al. Identification of mobile genetic
1071 elements with geNomad. Nat Biotechnol. 2023 Sept 21;1–10.

1072 71. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpides NC. CheckV assesses
1073 the quality and completeness of metagenome-assembled viral genomes. Nat Biotechnol.
1074 2021 May;39(5):578–85.

1075 72. Moraru C, Varsani A, Kropinski AM. VIRIDIC—a novel tool to calculate the intergenomic
1076 similarities of prokaryote-infecting viruses. Viruses. 2020 Nov 6;12(11):1268.

1077 73. Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H, Goto S. ViPTree: the viral
1078 proteomic tree server. Bioinforma Oxf Engl. 2017 Aug 1;33(15):2379–80.

1079 74. Orfei B, Pothier JF, Fenske L, Blom J, Moretti C, Buonaurio R, et al. Race-specific genotypes
1080 of *Pseudomonas syringae pv. tomato* are defined by the presence of mobile DNA elements
1081 within the genome. Front Plant Sci [Internet]. 2023 July 5 [cited 2025 June 7];14. Available
1082 from: https://www.frontiersin.org/journals/plant-
1083 science/articles/10.3389/fpls.2023.1197706/full

1084 75. Balbuena JA, Míguez-Lozano R, Blasco-Costa I. PACo: A Novel Procrustes Application to
1085 Cophylogenetic Analysis. PLOS ONE. 2013 Apr 8;8(4):e61048.

1086 76. Revell LJ. phytools 2.0: an updated R ecosystem for phylogenetic comparative methods
1087 (and other things). PeerJ. 2024 Jan 5;12:e16505.

1088 77. Xie Y, Wahab L, Gill JJ. Development and validation of a microtiter plate-based assay for
1089 determination of bacteriophage host range and virulence. Viruses. 2018 Apr;10(4):189.

78. Sprouffske K, Wagner A. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. BMC Bioinformatics. 2016 Apr 19;17:172.

79. Almeida-Neto M, Guimarães P, Guimarães Jr PR, Loyola RD, Ulrich W. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. Oikos. 2008;117(8):1227–39.

80. Galeano J, Pastor JM, Iriondo JM. Weighted-Interaction Nestedness Estimator (WINE): A new estimator to calculate over frequency matrices. Environ Model Softw. 2009 Nov 1;24(11):1342–6.

81. Newman MEJ, Girvan M. Finding and evaluating community structure in networks. Phys Rev E. 2004 Feb 26;69(2):026113.

82. Reichardt J, Bornholdt S. Statistical mechanics of community detection. Phys Rev E. 2006 July 18;74(1):016110.

83. Traag VA, Bruggeman J. Community detection in networks with positive and negative links. Phys Rev E. 2009 Sept 21;80(3):036115.

84. Clauset A, Newman MEJ, Moore C. Finding community structure in very large networks. Phys Rev E. 2004 Dec 6;70(6):066111.

85. Newman MEJ. Finding community structure in networks using the eigenvectors of matrices. Phys Rev E. 2006 Sept 11;74(3):036104.

86. Beckett SJ. Improved community detection in weighted bipartite networks. R Soc Open Sci. 2016 Jan;3(1):140536.

87. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech Theory Exp. 2008 Oct;2008(10):P10008.

88. Rodríguez-Gironés MA, Santamaría L. A new algorithm to calculate the nestedness temperature of presence–absence matrices. J Biogeogr. 2006;33(5):924–35.

89. Moury B, Fabre F, Hébrard E, Froissart R. Determinants of host species range in plant viruses. J Gen Virol. 2017 Apr;98(4):862–73.

90. Anderson MJ. Permutational multivariate analysis of variance (PERMANOVA). In: Wiley StatsRef: Statistics Reference Online [Internet]. John Wiley & Sons, Ltd; 2017 [cited 2025 June 4]. p. 1–15. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat07841

91. Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci. 2003;14(6):927–30.

92. Arbizu PM. pmartinezarbizu/pairwiseAdonis [Internet]. 2025 [cited 2025 Apr 14]. Available from: https://github.com/pmartinezarbizu/pairwiseAdonis

1124 93. Brooks M E, Kristensen K, Benthem K J ,van, Magnusson A, Berg C W, Nielsen A, et al.
1125 glmmTMB balances speed and flexibility among packages for zero-inflated generalized
1126 linear mixed modeling. R J. 2017;9(2):378.

1127 94. Allen B, Kon M, Bar-Yam Y. A new phylogenetic diversity measure generalizing the
1128 Shannon index and its application to phyllostomid bats. Am Nat. 2009 Aug;174(2):236–43.

1129

1130 ## Supporting information

1131 **S1 Fig. Bacterial resistance to phages and host range breadth of the tested phages.**

1132 (A) Number of phages to which bacteria are sensitive or resistant, with the phylogeny of the
1133 strains indicated at the bottom of the histogram. (B) Number of bacteria infected or not by
1134 the phages, with the taxonomy of the phages listed at the bottom of the histogram. Values
1135 are derived from the binary inhibition matrix.

1136 **S2 Fig. Taxonomic and phylogenetic characteristics, and ecological origins of phages and**
1137 **bacteria explaining host range.**

1138 (A) Bacterial percentage of inhibition by phages across bacterial phylogroups and ecological
1139 origins. Significance letters indicate results from the PERMANOVA test and associated pairwise
1140 comparisons. (B) Bacterial inhibition percentage of different phage genera. Significance letters
1141 represent results from a GLMM with a Gamma regression distribution and random effect
1142 (phage, bacteria, and the replicate imbricated in each of the phage-bacteria interactions), and
1143 post hoc pairwise comparisons performed using Tukey's test.

1144 **S3 Fig. Relative proportions of defence systems by functional category (%) and total number**
1145 **of defence systems across bacterial genomes.**

1146 On the left, a phylogenetic tree based on the core genomes of the bacterial strains is shown,
1147 with corresponding phylogroups indicated in the adjacent column.

1148 **S4 Fig. Percentage of similarity of defence system subtypes among bacterial strains in**
1149 **relation to their genome-wide average nucleotide identity (ANI).**

1150 **S5 Fig. Co-phylogeny of the CL4 prophage family-like (R-type syringacin).**

1151 Co-phylogenetic analysis between 68 strains of the *P. syringae* species complex (left),
1152 represented by a maximum likelihood phylogenetic tree based on their core-genomes, and
1153 their associated 68 CL4 family-like prophages (right), represented by a proteomic tree. The 44
1154 bacterial strains included in this study are marked with dots at the tips of the bacterial tree.
1155 Phylogroups and prophage-carrying strains are colour-coded accordingly. Two nodes with a
1156 maximum likelihood support value below 0.90 are marked with red dots.

1157 **S6 Fig. Experimental setup illustrating bacterial growth inhibition by phages.**

1158 Each biological replicate per bacterial strain is represented by a different colour. For clarity,
1159 only 3 of the 23 tested phages are shown. Three types of internal controls are included:
1160 positive controls with bacteria only, negative controls with medium only, and internal
1161 reference controls using the bacterial strain P1.01.01B03 (growth shown in S7B Fig) and the
1162 internal phage Ghuch01, which consistently inhibits this strain (see S7C Fig).

1163 **S7 Fig. Analysis of the different controls used in the phage host range *in vitro* tests.**

1164 (A) Analysis of culture medium controls by plate (labelled by bacteria). (B) The internal
1165 bacterial control, *P. syringae* strain P1.01.01B03, with an area under the empirical curve (AUC)
1166 threshold of 36 (first quartile), shows that for seven dates, all three replicates are below this
1167 threshold (the dates are marked with a red dot). (C) The internal phage control, Ghuch01, with

1168 an inhibition percentage threshold of 98% (first quartile), indicates that for two dates, all three
1169 replicates are below this threshold (the dates are marked with a red dot).

1170 **S8 Fig. Histogram showing the distribution of inhibition percentages within the 0.05 to 0.95**
1171 **probability range, displaying an approximately symmetric and unimodal shape.**

1172 **S9 Fig. Results of additional experiments on five potentially stimulating phages and five**
1173 **potentially stimulated bacteria**.

1174 Generally, bacterial populations decreased in the presence of phages. In 7% of observed
1175 phage-bacteria combinations, a "stimulatory" effect was identified, with inhibition values
1176 reaching up to -40%. Five phages showing strong stimulatory effect and five bacteria showing
1177 significant stimulation were chosen to test the reproducibility of the stimulation effect in
1178 subsequent experiments. (A) Matrix of mean inhibition percentages for the five *P. syringae*
1179 strains (rows) and five phages (columns), with their phylogenetic and taxonomic affiliations
1180 indicated by colour. The number of stars indicates the number of replicates showing a
1181 significantly inhibiting or stimulating percentage. Of ten initially identified significant
1182 stimulating interactions (delimited by red lines), only three were confirmed in subsequent
1183 experiments. (B) Size of the five bacterial populations in response to the phages. Stars
1184 correspond to significant differences from the control (no phage) based on the results of the
1185 GLM model, family = Gamma. Arrows indicate interactions with at least one of the replicates
1186 showing a significant inhibition or stimulation percentage. Out of the three significant
1187 interactions of stimulatory phenotype, only 2 phage-bacteria interactions resulted in slightly
1188 higher bacterial densities in the presence of phages compared to phage-free conditions,
1189 though this difference was not statistically significant. (C) Spearman correlation between

1190 bacterial concentration and inhibition percentage, by bacterium and phage. No correlation
1191 was significant and bacterial population size only poorly explains the inhibition percentages.

1192

1193 **S1 Table. Phage-bacterium matrices representing the percentage of inhibition of bacterial**
1194 **growth (A) and the statistical significance of inhibition (B).**

1195 Rows correspond to bacterial strains, while columns represent phage clones, each annotated
1196 with their respective phylogroups and genera.

1197 **S2 Table: Analysis of nestedness and modularity of the network of binary data representing**
1198 **the significance or non-significance of bacterial growth inhibition by phages (S1B Table).**

1199 Null models are described in Moury *et al.* (32) for B, N, C2, R2 and S, and in Moury *et al.* (89)
1200 for PD. For each null model and each algorithm, the frequency of simulations, among 100,
1201 showing a strictly higher nestedness (or modularity) score than that of the actual matrix is
1202 indicated. Significantly nested or modular networks ($p \leq 0.05$) are indicated in red.

1203 **S3 Table. Genetic analysis of the 88 prophages identified in the 44 *P. syringae* strains using**
1204 **VIRIDIC.**

1205 The 27 reference phages are highlighted in yellow (25 from the collection described in Feltin
1206 *et al.* (65) and 3 from NCBI). Species and genus-level clusters of the prophages are indicated
1207 on the right in colour. High genetic similarities are highlighted in green.

1208 **S4 Table. Phylogenetic characteristics, ecological origins, countries, and dates of sampling**
1209 **of the 44 *P. syringae* strains used for the host range study.**

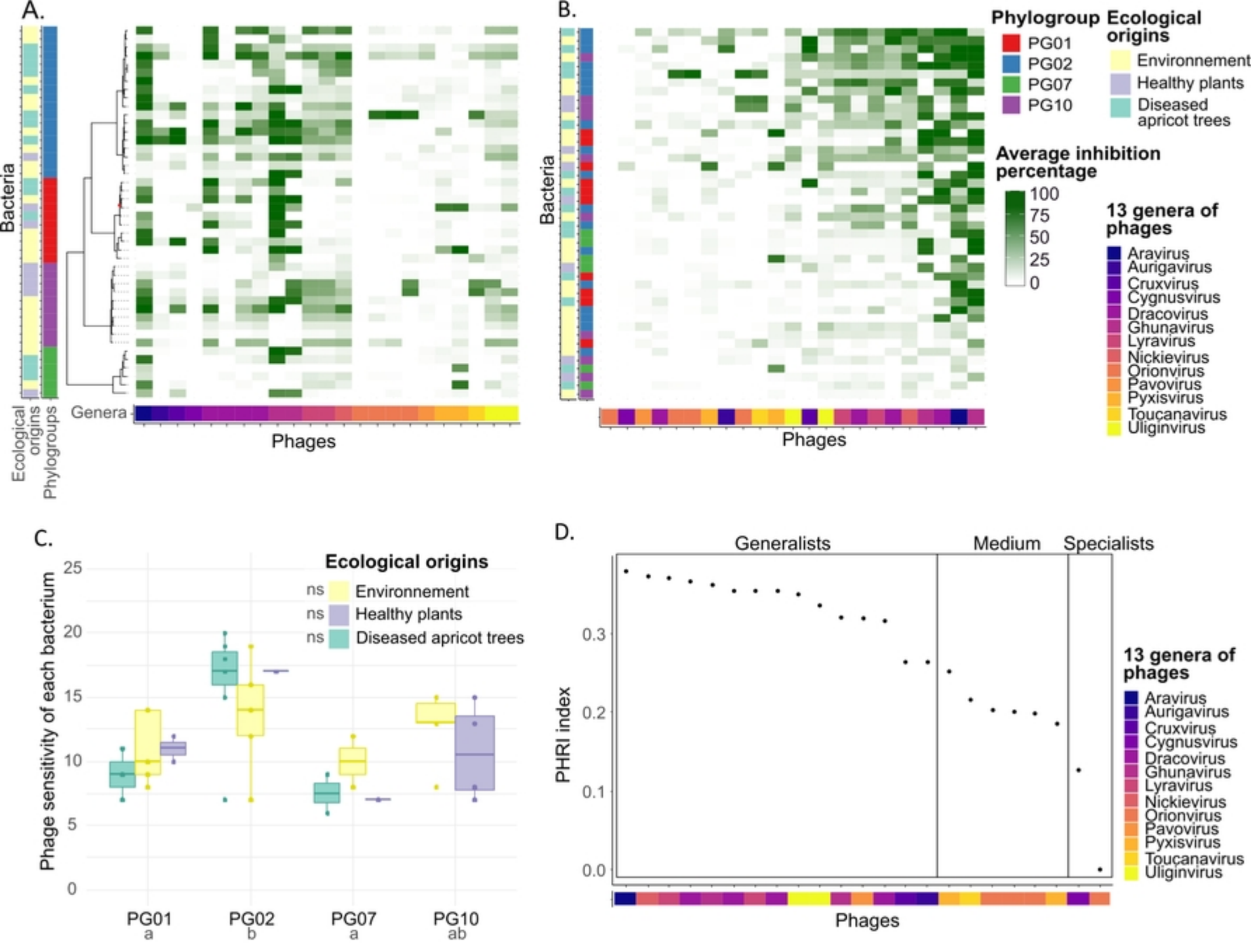
1210 The four *P. syringae* strains highlighted in bold are strains that were used to isolate some of

1211 the phages tested in the host range assay. Four strains are in the CIRM-CFBP, the French

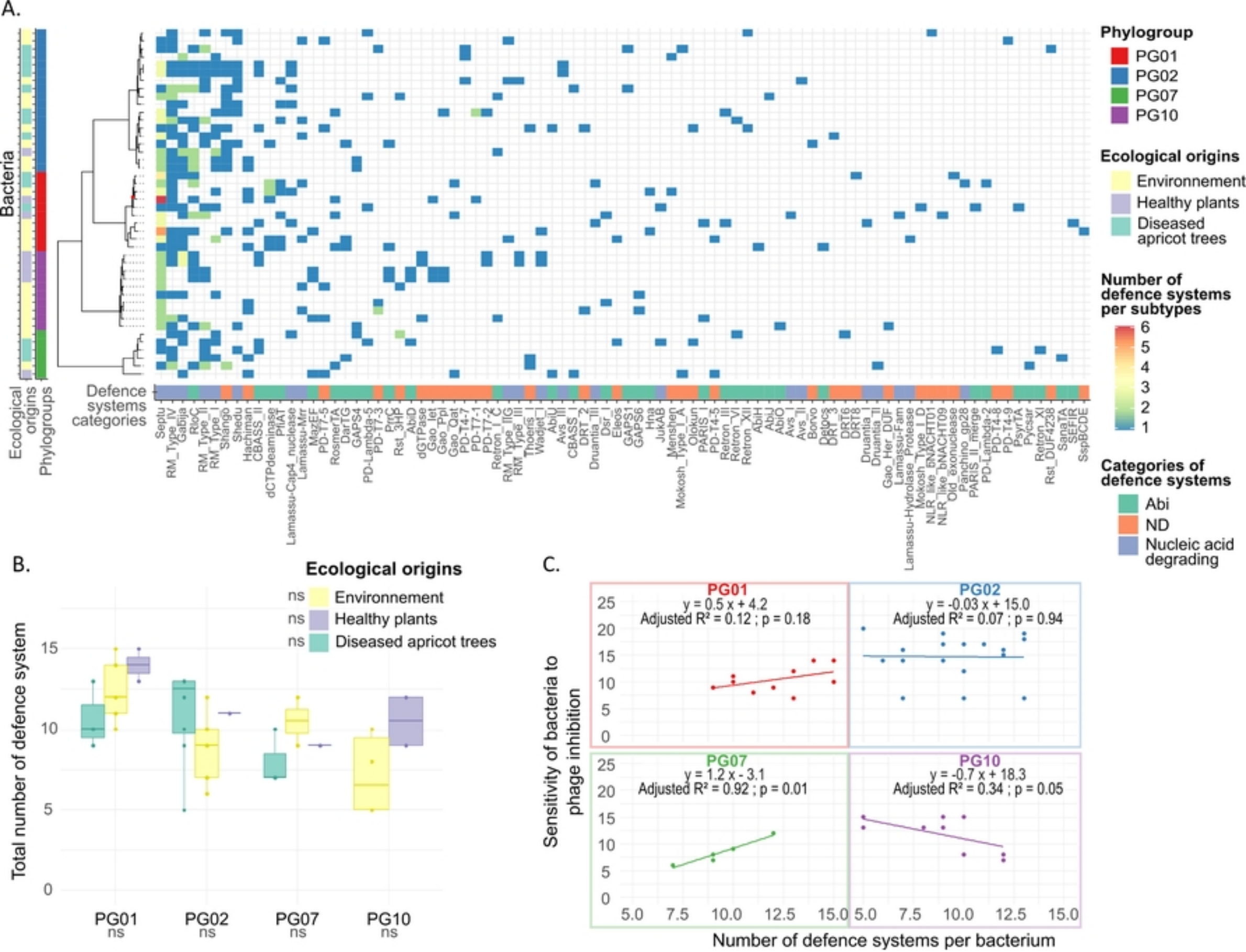
1212 collection for plant-associated bacteria (https://cirm-cfbp.fr/ ; June 2025): USA0007(CFBP

1213 8574), CC1559 (CFBP 8551), MAFF302273PT (CFBP 2339), and CC1583 (CFBP 8564).

1214

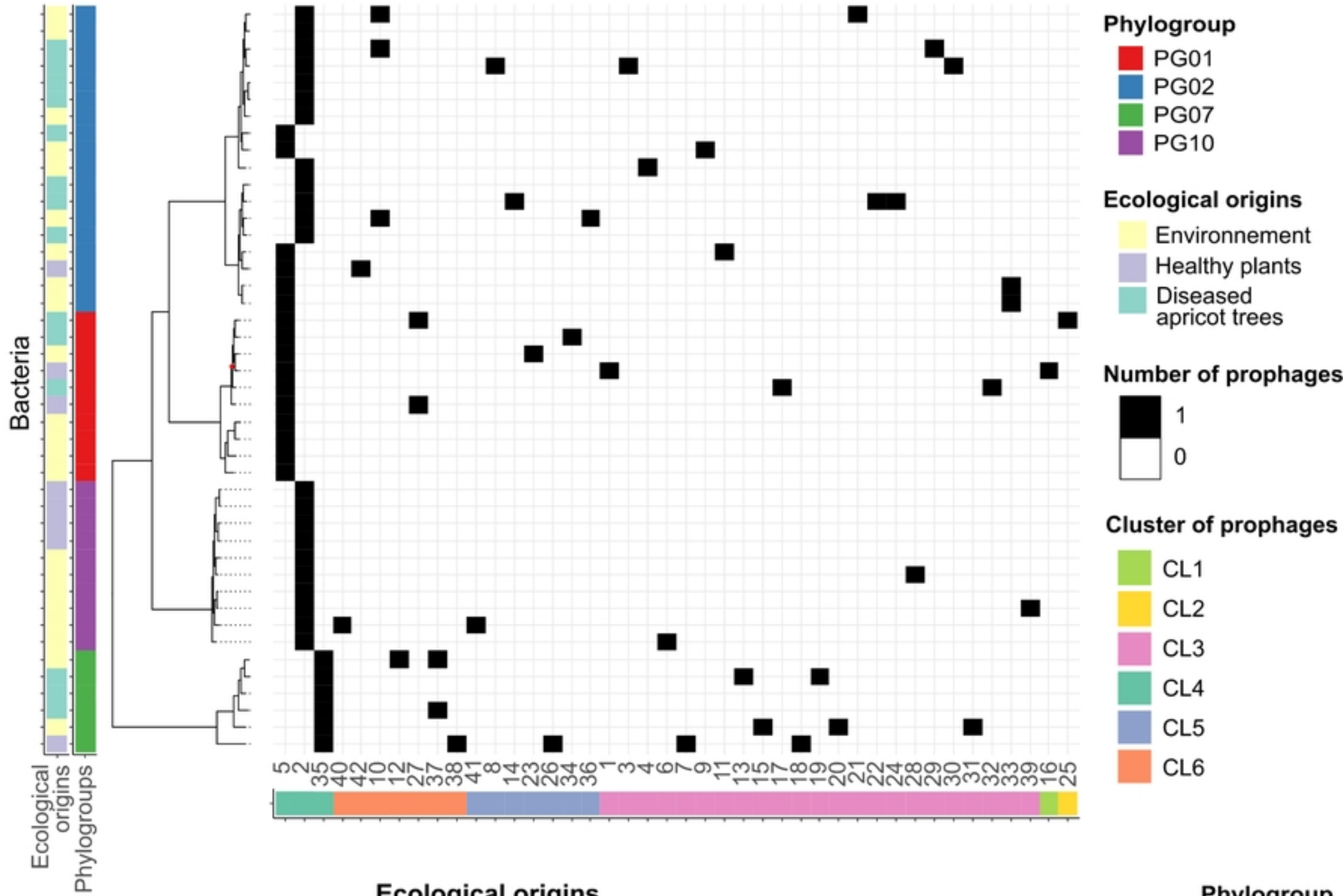

A.
Bacteria
Ecological origins
Phylogroups
Genera
Phages
B.
Bacteria
Phages
Phylogroup
PG01
PG02
PG07
PG10
Ecological origins
Environnement
Healthy plants
Diseased apricot trees
Average inhibition percentage
100
75
50
25
0
13 genera of phages
Aravirus
Aurigavirus
Cruxvirus
Cygnusvirus
Dracovirus
Ghunavirus
Lyravirus
Nickievirus
Orionvirus
Pavovirus
Pyxisvirus
Toucanavirus
Uliginvirus
C.
Phage sensitivity of each bacterium
Ecological origins
ns Environnement
ns Healthy plants
ns Diseased apricot trees
PG01 a
PG02 b
PG07 a
PG10 ab
D.
Generalists
Medium
Specialists
PHRI index
Phages
13 genera of phages
Aravirus
Aurigavirus
Cruxvirus
Cygnusvirus
Dracovirus
Ghunavirus
Lyravirus
Nickievirus
Orionvirus
Pavovirus
Pyxisvirus
Toucanavirus
Uliginvirus

Figure

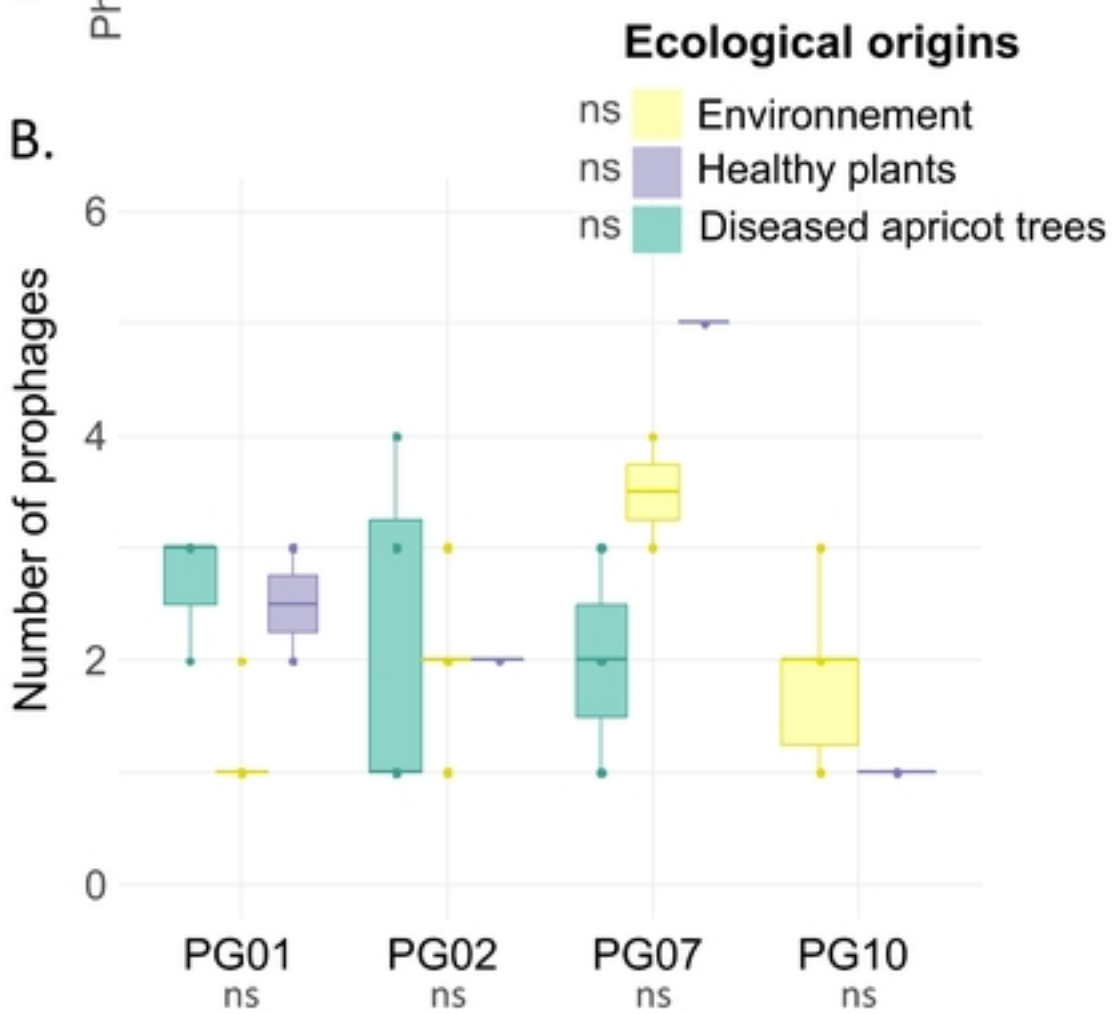

A.
Bacteria
Ecological origins
Phylogroups
Defence systems categories
Phylogroup
PG01
PG02
PG07
PG10
Ecological origins
Environnement
Healthy plants
Diseased apricot trees
Number of defence systems per subtypes
6
5
4
3
2
1
Categories of defence systems
Abi
ND
Nucleic acid degrading
B.
Ecological origins
ns Environnement
ns Healthy plants
ns Diseased apricot trees
Total number of defence system
PG01 ns
PG02 ns
PG07 ns
PG10 ns
C.
PG01
y = 0.5 x + 4.2
Adjusted R² = 0.12 ; p = 0.18
PG02
y = -0.03 x + 15.0
Adjusted R² = 0.07 ; p = 0.94
PG07
y = 1.2 x - 3.1
Adjusted R² = 0.92 ; p = 0.01
PG10
y = -0.7 x + 18.3
Adjusted R² = 0.34 ; p = 0.05
Sensitivity of bacteria to phage inhibition
Number of defence systems per bacterium

Figure

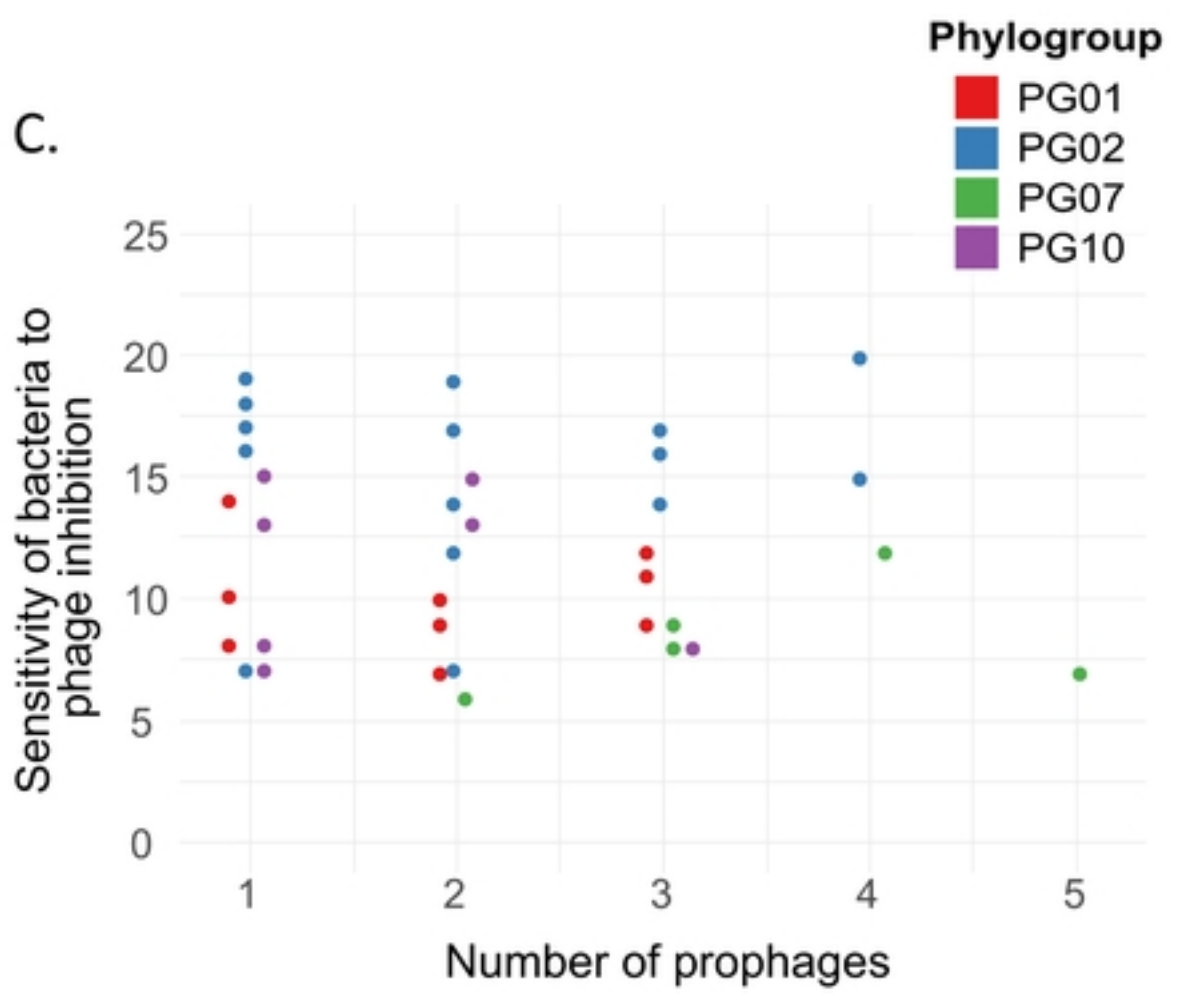
A.


Bacteria
Ecological origins
Phylogroups
Phylogroup
PG01
PG02
PG07
PG10
Ecological origins
Environnement
Healthy plants
Diseased apricot trees
Number of prophages
1
0
Cluster of prophages
CL1
CL2
CL3
CL4
CL5
CL6

B.


Ecological origins
ns Environnement
ns Healthy plants
ns Diseased apricot trees
Number of prophages
0
2
4
6
PG01
ns
PG02
ns
PG07
ns
PG10
ns

C.


Phylogroup
PG01
PG02
PG07
PG10
Sensitivity of bacteria to phage inhibition
0
5
10
15
20
25
1
2
3
4
5
Number of prophages

Figure