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.

66 Phage interactions with bacteria depends on a multitude of factors. The phage must
67 recognise and infect a given bacterium, integrate into the host genome and/or replicate
68 while avoiding or blocking bacterial defence systems (9,10). A range of approaches have
69 been employed to unravel the determinants involved in phage-bacterial interactions,
70 including ecological and evolutionary analyses of natural populations, and experimental
71 evolution (11,12). Mechanistic studies have also investigated the bacterial receptors (13,14)
72 with more recent work focusing on defence systems associated with susceptibility to phage
73 (14–17). For example, a study showed that *Vibrio lentus* strains that share identical phage
74 receptors display varying levels of susceptibility depending on the phage defence elements
75 present in their genomes (18). As with other aspects of phage-bacterial interactions,
76 research on prophages and defence systems in phytopathogenic bacteria remains limited
77 and have yet to connect these elements to host range data (19–25).

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

90 bacterial species (31). A recent study has introduced new statistical methods that
91 incorporate quantitative host–parasite interactions to analyse the structures of ecological
92 networks, providing deeper insights into the processes shaping phage–bacterial interactions
93 (32). Rather than relying solely on binary data (presence or absence of infection), these
94 approaches consider the strength of interactions, reflecting how effective each phage is at
95 infecting each host, and vice versa. This gradient-based perspective enables a more nuanced
96 understanding of infection patterns, host susceptibility, and phage infectivity, revealing
97 subtler ecological and evolutionary dynamics that would be overlooked by binary analyses
98 alone.

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

111 Historically, phage host range profiling was used to differentiate *P. syringae* strains (37–41).
112 More recent studies have emphasised the isolation and characterisation of novel *P. syringae*

113 phages and their potential as biocontrol agents (42,43). When reported, host-range assays
114 commonly aim to identify broad-host-range phages for practical applications (44). However,
115 ecological questions remain understudied, even though they offer valuable insights that
116 advance both fundamental and applied research. Investigating the roles of local adaptation,
117 host phylogeny and bacterial genomic defence arsenal (defensome) could help to clarify the
118 mechanisms shaping these networks, thereby guiding more effective biocontrol strategies.

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

136 We focused on the causal agents of bacterial canker of apricot trees to investigate the
137 structure and ecological drivers of phage–bacterium infection networks. Based on previous
138 studies, we hypothesised that phages would be locally adapted to their bacterial hosts,
139 leading to a modular infection network structure. We further proposed that this structure
140 could be explained by the presence of prophages and diverse bacterial defence systems
141 encoded in host genomes. To test these hypotheses, we addressed the following questions:
142 i) Are phages from diseased apricot trees specifically adapted to resident *P. syringae*
143 compared to closely-related strains from other environmental reservoirs? ii) Do closely-
144 related bacteria from different environments differ in their adaptation to phage? iii) What
145 bacterial genomic traits underlie the observed infection network structure? We conducted a
146 quantitative host-range assay using 23 genetically-diverse *P. syringae* phages and 44 *P.*
147 *syringae* strains. Whole-genome sequencing of all bacterial strains enabled detailed
148 annotation of prophages and defence systems. This integrative approach provides
149 mechanistic insights into how ecological origin and bacterial genomic content shape phage–
150 bacterial interactions, contributing to our understanding of the roles that phages play in
151 agricultural ecosystems.

152 Results

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

154 We characterised the phage–bacterial infection network by testing 23 phages against 44
155 strains from the *P. syringae* species complex (hereafter referred to as *P. syringae* strains) in
156 liquid culture. This allowed us to go beyond the information provided by traditional plaque
157 assays – that reveal phage replication – by providing data on bacterial growth. Each
158 interaction was quantified using two complementary metrics: 1) a continuous variable,

159 corresponding to the percentage of bacterial growth inhibition by phages (0–100%) (S1A
160 Table), and 2) a binary variable, indicating whether the inhibition was statistically significant
161 (1) or not (0), based on triplicate measurements (S1B Table). Because these values can be
162 interpreted from either the phage or bacterial perspective, we provide clear terminology to
163 distinguish them. From the bacterial perspective, these values represent susceptibility
164 (percentage inhibition) and sensitivity (binary significance of inhibition). Conversely, from
165 the phage perspective, these values reflect virulence (percentage inhibition) and infectivity
166 (binary success of infection).

167 Overall, the tested bacterial strains showed continuous distributions of susceptibility and
168 sensitivity values, reflecting the diversity of interactions of the panel with phages. On
169 average, bacterial susceptibility to phages was moderate, with a 24% reduction in growth,
170 and strains were sensitive to 12.3 out of the 23 tested phages (53.5%) (S1A Fig), indicating a
171 broader sensitivity than suggested by inhibition levels alone. The most susceptible bacterial
172 strains were DG1a.7, GAW0089, and DG12.6 (all from PG02), with average inhibition levels
173 of 46%, 50%, and 42%, respectively (S1A Table). The most broadly sensitive strains, DG1a.7,
174 7C, and CCE0067 (all from PG02), were significantly inhibited by 20, 19, and 19 phages,
175 respectively (87%, 83%, and 83% of tested phages respectively) (S1B Table). In contrast,
176 bacterial strains CC1583, DG8.15, and DGA9CS5-8 (from PG10, PG07 and PG10 respectively)
177 showed the highest resistance, with average inhibition percentages of only 3%, 6%, and 6%,
178 respectively (S1A Table). DG8.15 (from PG07) resisted to 17 phages (74% of tested phages)
179 and seven other strains (19B, 41D, CST0086, DGA9CS5-15, FMU-107, and USA0088 from
180 PG01, PG02, PG02, PG10, PG07 and PG02) were resistant to 16 phages (70%) (S1B Table).

181 From the phage perspective, average bacterial growth inhibition across the strain panel was
182 23%, reflecting moderate virulence. However, each phage was able to infect an average of
183 22.4 out of 42 tested *P. syringae* strains (53.3%) (S1B Fig), indicating a relatively broad
184 infectivity despite moderate inhibition levels. The most virulent phages were Ghual01 (65%
185 inhibition), Arace01 (57%), and Draal03 (44%) (S1A Table). The most infecting phages were
186 Nican01, Arace01, and Ghual01 infecting 39 (93% of tested bacteria), 37 (88%), and 34 (81%)
187 strains, respectively (S1B Table). Moreover, Arace01 and Ghual01 were the only phages
188 capable of infecting strains from other species in the *Pseudomonas* genus. Arace01 infected
189 a *P. rhizosphaerae* strain at 4.5×10^4 PFU/mL, while Ghual01 infected a *P. graminis* strain at
190 1.7×10^5 PFU/mL. In contrast, Orimi01, Cygsa01, and Pavpe01 displayed the lowest
191 virulence, averaging 0%, 3%, and 4% respectively (S1A Table), while Orimi01, Cygsa01, and
192 Touem01 had the lowest infectivity, infecting only 2 (5% of tested bacteria), 5 (12%), and 11
193 (26%) strains, respectively (S1B Table).

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

199 *Nested structure of the phage-bacterial interaction network and association with*
200 *bacterial traits*

201 The structure of the phage-bacterial infection network was analysed using multiple
202 algorithms to assess its nestedness and modularity. Given the diversity of both phages and
203 bacterial hosts, spanning 13 phage genera and several phylogroups of *P. syringae*, and the

204 ecological origins of phages (diseased apricot trees) and bacteria (diseased apricot trees,
205 healthy plants and other environmental reservoirs), we hypothesised a modular structure
206 type. However, the network representing bacterial growth inhibition percentage by phages
207 displayed a strong nested structure, as indicated by scores of 0.48 and 0.44, on a scale
208 varying from 0 to 1, with the *WINE* and *wNODF* algorithms, respectively (Table 1).
209 Nestedness was highly significant ($p < 0.01$) against nearly all null models, except null model
210 B, which was shown to be excessively conservative (32). Modularity analyses differed
211 depending on the algorithm (Table 1). With the *spinglass* algorithm, selected for its good
212 performance in terms of false positive rate (32), the modularity score was low (0.01, on a
213 scale varying from -0.5 to 1) and was not significant, whatever the null model. The four
214 alternative modularity algorithms produced slightly higher scores (0.09 < modularity < 0.25)
215 and showed significant modularity ($p < 0.01$) with both null models C1 and R1. These four
216 algorithms show lower performances than *spinglass* in terms of false positive rates, but
217 combined with null models C1 and R1, they show generally acceptable performances in
218 terms of false positive and false negative rates (32). These results suggest globally a slight
219 degree of modular organisation within the network. The strong and significant nestedness
220 and the weak modularity were further supported by the analysis of the binary data
221 representing the significance of bacterial growth inhibition (S2 Table).

222

223 **Table 1. Analysis of nestedness and modularity of the bacterial growth inhibition network**
224 **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

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

To assess whether the nested structure of phage–bacterial networks is shaped by bacterial evolutionary history, we first tested for a phylogenetic signal in bacterial sensitivity summed by bacteria (Fig 1). The analysis revealed a strong phylogenetic signal ($\lambda = 1.0$; LR ($\lambda = 0$) = 48.26; $p < 0.01$), indicating that closely-related bacterial strains tend to be infected by a similar number of phages. To explore this pattern further, we evaluated the effect of phylogroup identity using both a Generalised Linear Model (GLM) with the bacterial sensitivity summed by bacteria (Fig 1C) and a PErmutational Multivariate ANalysis Of VAriance (PERMANOVA) on bacterial susceptibility profiles (S2A Fig). Both analyses confirmed a significant influence of phylogroup on bacterial susceptibility, explaining almost 10% of the variance (GLM: $p < 0.01$; PERMANOVA: $R^2 = 0.09$; $p < 0.01$). In contrast, ecological origin had no significant effect in either test (GLM: $p = 0.87$; PERMANOVA: $R^2 = 0.05$; $p = 0.29$). Specifically, PG02 strains emerged as the most sensitive to phages, being infected on

242 average by 14.6 phages and showing the highest susceptibility, with 20.7% growth inhibition.
243 PG01 strains showed intermediate sensitivity being infected by 10.4 phages and an average
244 of 14.6% inhibition, followed closely by PG10 strains with 15.3% inhibition. In contrast, PG07
245 strains were the most resistant overall, infected on average by only 8.4 phages and
246 displaying the lowest susceptibility with 8.73% inhibition.

247

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

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

259

260 The broad taxonomic diversity of the phages studied here, covering 13 genera, includes both
261 specialists and generalists. To characterise phage host range, we calculated the Phylogenetic
262 Host Range Index (PHRI), which incorporates the number, the evenness, and the
263 phylogenetic diversity of host bacteria (48). The PHRI values ranged from 0 (Orimi01, a
264 specialist phage) to 0.38 (Arace01, a generalist phage), forming a gradient that reflects the

265 nested structure of the phage-bacteria interaction network (Fig 1D). Based on PHRI
266 classification, 9% of phages were categorised as specialists, 65% as generalists, and 26% as
267 having an intermediate host range (Fig 1D). At the genetic level, significant differences in
268 virulence were observed between phage genera (GLMM; $p < 0.01$). The generalist genus
269 *Aravirus* exhibited the highest virulence (52.8%), while the specialist genera *Cygnusvirus* and
270 *Orionvirus* displayed markedly lower virulence levels, averaging 1.1% and 4.2%, respectively
271 (S2B Fig). Additionally, phages with similar host infection profiles clustered by taxonomy,
272 with those sharing >70% nucleotide identity (same genus) explaining 63.5% of the variation
273 in virulence (PERMANOVA: $R^2 = 0.64$, $p < 0.01$), indicating genus-level host specificity (Fig
274 1A). For example, the two phages from the *Lyravirus sulafat* species and the *Uliginvirus*
275 genus exhibited highly similar infection host profiles.

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

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

283 The 44 strains of *P. syringae* analysed in this study were examined for their accessory
284 genome content, with a particular focus on their repertoire of defence systems to
285 counteract mobile genetic elements such as phages. Across these strains, a total of 48
286 defence system types and 86 subtypes were identified (Fig 2A). The most abundant defence
287 subtypes included Septu ($n = 81$), type IV Restriction-Modification (RM) ($n = 37$), and Gabija

288 (*n* = 31) (Fig 2A). In contrast, 31 subtypes were sporadically detected across the bacterial
289 genomes, including rare defence systems such as three Abi subtypes (AbiH, AbiJ, and AbiO),
290 two Asv subtypes (AsvI and AsvII), and three DTR subtypes (DTR6, DTR8, and DTR3) (Fig 2A).
291 Defence systems are present at an average rate of 9.93 per strain, with two strains (AF0015
292 and P3.01.09C09, both from PG01) harbouring the highest number (15 systems), and four
293 strains (DG1A_7 from PG02, CSZ0293, SZ0119, and SZB0008 from PG10) the lowest number
294 (5 systems) (Fig 2A). Each defence system can be classified into different functional
295 categories to better characterise its role. Specifically, 38.4% of defence system subtypes are
296 associated with abortive infection, 24.4% with nucleic acid degradation, and 37.2% remain
297 unclassified (Fig 2A).

298

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

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

311

312 To assess whether the distribution of defence systems across bacterial genomes was
313 influenced by phylogeny, we tested for a phylogenetic signal in the total number of encoded
314 systems per genome. A strong signal indicated that closely related strains tend to harbour
315 similar numbers of defence systems ($\lambda = 1.0$; LR test ($\lambda = 0$): 44.18; $p < 0.01$) (S3 Fig), as
316 described in a previous study (24). Phylogroup significantly explained variation in defence
317 system profiles (PERMANOVA, $R^2 = 0.14$, $p < 0.01$), with a weaker effect from ecological
318 origin ($R^2 = 0.06$, $p < 0.01$), while growth rate of the bacteria had no significant impact ($R^2 =$
319 0.02, $p = 0.51$) (Fig 2A). Post-hoc comparisons based on this analysis revealed that strains
320 from PG10 displayed significantly distinct defence system profiles compared to those from
321 PG02 and PG07, while strains isolated from healthy plants exhibited distinct defence system
322 profiles compared to other ecological origins. However, none of these factors significantly
323 affected the total number of defence systems, although the effect of phylogroup
324 approached significance (GLM: $p = 0.06$) (Fig 2B). This phylogroup trend is supported by a
325 larger dataset of 590 *P. syringae* genome assemblies, where PG01 strains encode the highest
326 number of defence systems, while PG07 and PG10 harbour some of the lowest counts (24).

327 While closely-related strains tend to possess similar numbers of defence systems, we further
328 investigated whether they also share the same subtypes of defence systems. To this end, we
329 calculated the percentage of shared defence system subtypes between each pair of bacterial
330 strains, where 100% indicates complete identity in the sets of subtypes. However, no
331 significant correlation was observed between the percentage of shared defence system
332 subtypes and the genome-wide Average Nucleotide Identity (ANI) between strains (linear
333 regression; adjusted $R^2 = 0.03$; $p < 0.01$) (S4 Fig). This suggests that, although phylogeny and

334 ecological origins influence the profile of defence systems, their inheritance does not follow
335 a strict vertical transmission. Furthermore, only 1.14% of the identified defence systems
336 were linked to prophages within bacterial genomes, indicating that the vast majority are
337 located outside prophage regions. This low prevalence indicates that while defence systems
338 may spread horizontally among bacterial populations, their dissemination is likely mediated
339 by mobile genetic elements other than prophages.

340 Defence systems may impose varying energetic costs on bacterial cells, potentially
341 influencing their allocation of resources. In strains of *P. syringae*, defence systems related to
342 nucleic acid degradation comprised, on average, 54% of the total defence systems, while
343 abortive infection systems, which are considered more metabolically costly (49), made up an
344 average of 25% (S3 Fig). We then revealed that the proportion of defence system categories
345 were significantly shaped by phylogroup, explaining 16% of the variance (PERMANOVA; R^2 =
346 0.16; p = 0.03), and growth rate (R^2 = 0.07; p = 0.04), but not ecological origin (R^2 = 0.08; p =
347 0.11). Pairwise comparisons based on this analysis showed that strains from PG07 had
348 significantly different defence category profiles compared to those from PG01 and PG02.

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

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

355 We identified 88 prophages among the 44 genomes of *P. syringae*. On average, each strain
356 carried two prophages, with one strain (FMU-107) harbouring five prophages and 18 strains

357 (41%) containing a single cryptic prophage (Fig 3A). Among these prophages, 36% were
358 classified as high quality (complete or nearly complete according to CheckV software), 39%
359 as medium quality, and 25% as low quality. Prophage sizes ranged from 12.0 kb to 89.8 kb,
360 with an average size of 32.2 kb. The size distribution exhibited three peaks at approximately
361 18 kb, 44 kb, and 57 kb. To classify these prophages taxonomically, we used VIRIDIC, which
362 identified 65 novel phage species and 42 novel genera (S3 Table), while proteomic tree
363 analysis with Viptree revealed six distinct prophage clusters, corresponding approximately to
364 family-level taxonomic divisions (Fig 3A) (50). One of these, the CL4 family-like prophage,
365 was found in all 44 *P. syringae* strains, which span four distinct phylogroups (Fig 3A). CL4
366 family-like prophage is a tailocin, specifically an R-type syringacin, a bacteriocin derived from
367 a bacteriophage (51,52). These tailocins are widely distributed across the *P. syringae* species
368 complex and are thought to play a key role in inter-bacterial competition by enabling strains
369 to kill closely-related competitors (53,54). Moreover, we found that the taxonomy of the CL4
370 family-like prophage mirrors the phylogeny of their bacterial hosts across 68 strains from 10
371 phylogroups (including the 44 strains analysed in this study), with a significant
372 cophylogenetic congruence signal (Procrustes Approach to Cophylogeny (PACo): residual
373 sum of squares = 0.15; $p < 0.01$) (S5 Fig).

374

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

377 (A) Distribution of prophages in the genomes of the 44 *P. syringae* strains, categorised by
378 genus and family-like cluster (VipTree). Bacterial strains are organised based on a maximum
379 likelihood phylogenetic tree based on their core-genomes, with their phylogroups and

380 ecological origins indicated on the right. A node with a maximum likelihood support value
381 below 0.90 is marked with a red dot. (B) Abundance of prophages per bacterial strain
382 according to their phylogroups and ecological origins. ns: not significant ($p > 0.05$), according
383 to post-hoc Tukey tests from a Generalised Linear Model with a Poisson distribution. (C)
384 Bacterial sensitivity to phages as a function of the number of prophages, coloured by
385 phylogroups.

386

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

402 lowest, and PG01 and PG10 showed intermediate levels (24). In contrast, differences across
403 ecological origins were minimal (Fig 3B).

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

408 *Contribution of bacterial defence systems and prophages to bacterial sensitivity to*
409 *phages*

410 Since bacterial defence systems target phages (49), a negative correlation is expected
411 between their number and bacterial sensitivity to phages. To test this, we examined the
412 correlation between bacterial sensitivity and the number of encoded defence systems, while
413 accounting for phylogenetic background. For both PG01 and PG02, no significant correlation
414 was detected (linear regression for PG01: adjusted $R^2 = 0.12$; $p = 0.18$; for PG02: adjusted R^2
415 $= -0.07$; $p = 0.94$) (Fig 2C). In contrast, a negative correlation was detected in PG10,
416 supporting our expectation (linear regression for PG10: adjusted $R^2 = 0.34$; $p = 0.05$) (Fig 2C).
417 In contrast, PG07 exhibited an unexpected positive correlation, suggesting that strains with a
418 higher number of defence systems were more sensitive to phages (linear regression for
419 PG07: adjusted $R^2 = 0.92$; $p < 0.01$) (Fig 2C).

420 While prophages can contribute to bacterial immunity against phage infection (55), our
421 results revealed no significant correlation between prophage abundance and bacterial
422 sensitivity to phages. This suggests that the detected prophages do not play a defining role
423 in resistance, either overall (linear regression; adjusted $R^2 = -0.02$; $p = 0.81$) or within
424 individual phylogroups (PG01: adjusted $R^2 = 0.01$, $p = 0.33$; PG02: adjusted $R^2 = -0.02$, $p =$

425 0.43; PG07: adjusted R^2 = 0.01, p = 0.38; PG10: adjusted R^2 = –0.12, p = 0.94) (Fig 3C).
426 Moreover, no correlation was observed between prophage abundance and the number of
427 defence systems encoded in bacterial genomes, suggesting that these systems do not
428 provide protection against prophage integration or maintenance (linear regression; adjusted
429 R^2 = –0.02; p = 0.64).

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

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

441 Consistent with the above findings, bacterial phylogeny and phage taxonomy emerged as
442 the strongest explanatory variables for inhibition percentage, with an importance weight of
443 1.0 across the 256 models (Table 2). In addition, a variable indicating whether the phage had
444 been isolated using one of the tested bacterial strains was significantly associated with
445 inhibition level. Specifically, phages isolated from strains within the tested host range
446 exhibited a slightly lower inhibition capacity (mean = 15.1%) compared to those isolated
447 from external strains (mean = 17.7%).

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

Explanatory variable	*Sum of AIC weights*	*N Models*
Phage genera ⍰	**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 ⍰	**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 ⍰; bacterial variables are unmarked.

460 The model identified the total number of prophages in bacterial genomes as a significant
461 predictor of inhibition of bacterial growth (sum of weights = 0.72; Table 2). Inhibition levels
462 varied nonlinearly with prophage abundance: inhibition levels were 15.6%, 14.7%, 17.5%,
463 28.4%, and 10.9% for strains carrying 1 to 5 prophages, respectively. While prophage
464 abundance did not predict bacterial sensitivity, it appeared to influence bacterial
465 susceptibility, suggesting that prophage-rich bacteria may be more susceptible to highly-
466 effective phages. In contrast, the total number of defence systems did not significantly
467 explain bacterial growth inhibition (sum of weights = 0.38; Table 2), despite its correlation
468 with the bacterial sensitivity for some phylogroups.

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

478

479 Discussion

480 Our study uncovered a strong nested structure in the phage–bacterial infection network
481 involving genetically-diverse *P. syringae* strains and phages. This discovery is in strong,

482 contrast with the modular structure frequently reported (11,29). This pronounced
483 nestedness likely reflects the ubiquitous nature of the highly diverse *P. syringae* complex,
484 mirrored by a phage community ranging from specialists to generalists in host range. While
485 bacterial phylogeny had a marked influence on the nestedness of the infection network, we
486 found no evidence of local phages-bacteria co-adaptation. Defence systems and prophages
487 contributed to phage resistance, but their distribution was largely shaped by bacterial
488 phylogeny and ecological origin, rather than phage pressure alone. Together, these findings
489 highlight how evolutionary and ecological forces shape phage–bacterial infection networks
490 in natural environments.

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

492 The observed nested infection network is unexpected when we consider the high genetic
493 diversity of the phages (13 genera, 17 species) and bacteria (4 phylogroups and 7 clades,
494 corresponding to genus and species levels, respectively) used in this work as well as the
495 varied ecological origins of the bacterial strains. Such a diversity of organisms could be
496 expected to reveal local adaptation, where phages perform better on sympatric than
497 allopatric bacterial hosts, in line with what has been demonstrated previously (56,11).
498 However, our findings suggest that phage–host interactions in this pathosystem transcend
499 ecological boundaries. Phages isolated from apricot tree soil infected bacterial strains from
500 both diseased orchards and aquatic environments with comparable efficiency, and bacterial
501 origin did not broadly influence bacterial sensitivity/susceptibility to phages. This absence of
502 modularity, combined with a lack of ecological signal in host range, points to broader
503 ecological connectivity and population flow within the *P. syringae* species complex. These
504 results align with previous work (36) showing that *P. syringae* population structures are

505 shaped more by fluctuating ecological dynamics than by geographic isolation, likely due to
506 documented exchanges between agricultural and aquatic environments (57). Overall, the
507 nested structure observed here implies that interaction dynamics in this system are
508 maintained by widespread generalism and ongoing ecological mixing, rather than niche-
509 specific adaptation.

510 Bacterial phylogeny plays a central role in shaping both the composition of defence systems,
511 the distribution of prophages, and the sensitivity of *P. syringae* strains to phages. In our
512 study, PG02 strains, known for their ubiquitous presence across habitats and for containing
513 strains capable of infecting a broad range of plant species (46,35), were also the most
514 sensitive to phages, carried a moderate number of defence systems, and encoded one of the
515 highest prophage loads (24). In contrast, PG07 strains, which are predominantly associated
516 with diverse environmental reservoirs (46), were the most resistant to phages despite
517 encoding relatively few defence systems and prophages. PG01 strains, mainly isolated from
518 diseased plants (46), showed intermediate bacterial sensitivity to phages, possessed the
519 greatest number of defence systems, and encoded a moderate prophage load (24). Finally,
520 PG10 strains, exclusively environmental, showed moderate bacterial sensitivity to phages,
521 low numbers of defence systems, and a moderate prophage count (24). These findings
522 suggest a potential trade-off between ecological ubiquity and bacterial sensitivity to phages,
523 a hypothesis that requires direct experimental investigation. In a similar vein, a recent study
524 reported that effector gene counts in *P. syringae*, which are key to pathogenicity, positively
525 correlate with phage defences and show partial conservation within phylogroups (24).

526 While phylogroups show consistent trends in defence system profiles, closely-related strains
527 often do not share identical defence systems, suggesting acquisition through horizontal gene

528 transfers (49,58). However, strains within the same phylogroup tend to retain consistent
529 features in their defensome, including the number and functional categories of defence
530 systems. This pattern suggests that each *P. syringae* phylogroup may follow a common anti-
531 phage strategy, while relying on distinct sets of defence systems to implement it. For strains
532 belonging to phylogroup PG10, we observed a negative correlation between the number of
533 defence systems and sensitivity to phages, suggesting a protective effect from accumulating
534 diverse systems in their genome (17). Conversely, strains in PG07, although generally the
535 most resistant to phages, showed a positive correlation between bacterial sensitivity to
536 phages and defence systems. This counter-intuitive pattern could reflect historical predation
537 by phages, where strains subjected to intense phage attack would have accumulated
538 defence systems (59). Alternatively, the apparent lack of protection may arise because these
539 systems target phages or mobile genetic elements (MGEs) other than those tested here.
540 However, it is important to interpret these trends cautiously, as they are based on relatively
541 small sample sizes (n=6 for PG07; n=10 for PG10), and our understanding remains partial
542 given the continual discovery of novel defence systems (60).

543 *Limitations and alternative approaches for understanding phage-bacterium*
544 *interactions*

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

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

558 A complementary strategy to uncover bacterial genetic determinants of sensitivity to phages
559 and resistance phenotypes involves genotype–phenotype association analyses, including
560 approaches based on gene presence–absence matrices, phylogenetic trait correlation
561 methods and genome-wide association studies (GWAS). This latter method has recently
562 been applied to 80 *P. aeruginosa* strains, identifying host genes that influence susceptibility
563 to two phages, particularly those related to flagella and LPS/O-antigens associated with
564 lipopolysaccharide biosynthesis (63). While our dataset of 44 *P. syringae* genomes may
565 provide a sufficient basis for GWAS, challenges remain. The high genomic diversity of the *P.*
566 *syringae* species complex and the large number of phages tested (n = 23), each potentially
567 interacting with different receptors or resistance pathways, add considerable complexity.
568 Future studies may address this by focusing on specific phage-host combinations or by
569 applying comparative genomics approaches that can help disentangle the genetic basis of
570 resistance across diverse strains.

571 *Applied perspectives and future directions*

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

583 To further test our hypothesis that ecological connectivity shapes the nested structure of
584 phage–*P. syringae* interactions, future studies should isolate phages from diverse non-
585 agricultural environmental, such as water and healthy plants. These newly isolated phages
586 should then be tested against the same bacterial strains used in our study, or against new
587 strains from corresponding habitats. If similar nested network structures emerge, it would
588 strengthen the idea that phage and *P. syringae* populations continuously migrate between
589 agricultural and non-agricultural environments, forming an ecologically-connected meta-
590 community. Moving forward, several key research priorities should guide future
591 investigations: verification of the role of specific defence systems through genetic
592 manipulation like gene deletion and heterologous expression; systematic investigation of
593 receptor usage by our studied phages to reveal molecular determinants of host range; and

594 strategic expansion of the bacterial strain collection to improve phylogenetic and ecological
595 coverage, with particular focus on understudied phylogroups. These complementary
596 approaches would collectively strengthen our understanding of the molecular and ecological
597 factors driving the phage-bacteria interactions observed in this system. Extending such
598 analyses to other ecologically and phylogenetically-distinct plant-associated bacteria, such as
599 *Xanthomonas*, *Ralstonia*, or *Agrobacterium*, could reveal whether the patterns observed in
600 *P. syringae*-phages infection network are generalisable or reflect unique evolutionary
601 dynamics of this highly diverse species complex.

602

603 **Materials and Methods**

604 *Phages panel*

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

616 *Bacterial strains panel*

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

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

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

632 Of the 44 *P. syringae* strains, 39 were sequenced using 20 µL of genomic DNA, with
633 sequencing performed by Plasmidsaurus Inc (San Francisco, USA) using Oxford Nanopore
634 Technology (ONT) and a custom analysis and annotation pipeline
635 (https://plasmidsaurus.com/). The long-read sequencing library was prepared without
636 amplification using v14 library preparation chemistry, and sequencing was conducted with
637 R10.4.1 flow cells. The genomes of five additional strains were obtained from the RefSeq
638 library at NCBI: CC1559 (GCF_000452685.1), USA0007 (GCF_000452545.1), 41A

639 (GCF_000935775.1), MAFF302273^{PT} (also known as JD08) (GCF_000145925.1), and CC1583
640 (GCF_000452665.1).

641 The core genome analysis was performed using a private database within the EDGAR 3.5
642 platform (67). The core genome was defined as the set of orthologous genes present in all
643 332 genomes of the *P. syringae* species complex, including the 44 strains analysed in this
644 study. A total of 1,346 core genes were identified per genome, resulting in a total of 446,872
645 aligned genes across 332 genomes. Each set of orthologous genes was aligned individually at
646 the protein level using MUSCLE (67). The resulting alignment were concatenated into a
647 single alignment, resulting in a total of 512,223 amino acid residues per genome and
648 170,058,036 residues across all genomes. Phylogenetic relationships were inferred using the
649 FastTree algorithm within EDGAR, which constructs maximum-likelihood phylogenetic trees
650 based on the core genome alignment (68). All phylogenetic trees of bacterial strains shown
651 in the figures were rooted at the midpoint of the longest path and the nodes with a
652 maximum likelihood support value below 0.90 were marked with red dots.

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

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

658 The prophages within the 44 *P. syringae* strains were identified using *geNomad* version 1.5.0
659 (70) based on assembled nucleotide sequences of bacterial genomes. Among the 97
660 identified prophage sequences, only those classified as viral by *geNomad* and *CheckV* (71)
661 were retained. Specifically, sequences labelled as "No Terminal Repeats" in the topology

662 column and those with a "viral_genes" value of 0 were excluded. This filtering resulted in a
663 final dataset of 88 verified viral prophages, accounting for 91% of the total identified
664 prophages across the 44 strains. Additionally, a genetic analysis of the prophages was
665 performed using *VIRIDIC* (72) to classify them into genera and species (S3 Table). *VipTree*
666 was used to generate prophage clusters based on a proteomic tree analysis using distance-
667 based methods (BIONJ from tBLASTx scores) (73). *geNomad* and *VipTree* were then
668 employed to identify prophages belonging to the CL4 cluster and to construct a proteomic
669 tree encompassing 68 CL4 prophages (24 reference strains + 44 analysed strains) (74).

670 To assess the congruence between the taxonomy of CL4 prophages and the phylogeny of
671 their bacterial hosts, we performed a PACo (Procrustes Approach to Cophylogeny) analysis.
672 Phylogenetic distance matrices were generated from the host and prophage trees using the
673 *cophenetic* function (75). These matrices were then centred and corrected for negative
674 eigenvalues using the "Cailliez" correction method. Congruence was tested using the *PACo*
675 function from the *paco* package with 1,000 permutations to assess statistical significance.
676 The co-phylogeny visualisation was performed with *phytools* package and *cophylo* function
677 (76).

678 *Phage host range analysis*

679 *a. Experimental setup*

680 To analyse the host range of phages infecting *P. syringae*, bacterial growth kinetics in the
681 presence of phages were measured using 96-well plates (77). The dataset comprised 44
682 bacterial strains tested in biological triplicate without phage as a positive control, and the
683 same 44 strains in biological triplicate exposed individually to each of the 23 phages (S6 Fig).
684 Each phage was tested against each biological triplicate of bacterial strain, resulting in 3,036

685 growth kinetic measurements (23 phages × 44 bacterial strains × 3 replicates) (S6 Fig).
686 Negative controls (culture medium) were used to check for contamination, with plates
687 containing more than two contaminated wells out of four discarded (S6 Fig and S7A Fig). The
688 combination of the two internal controls were added to each plate: the *P. syringae* strain
689 P1.01.01B03 (PG01a, isolated from apricot buds in France in 2014) and phage Ghuch01,
690 chosen because it completely inhibited the growth of strain P1.01.01B03 under the
691 experimental conditions of this study (45,65) (S6 Fig). A plate was considered valid if the two
692 internal controls had all three replicates above the first quartile threshold of bacterial
693 growth and growth inhibition (S7B-C Fig).

694 Bacteria were incubated in 5 mL of liquid KB medium at 25°C and 200 rpm for 4 hours. Each
695 replicate was standardised to an OD_{600nm} = 0.01 (~10^7 CFU/mL) in phosphate buffer and
696 diluted in liquid KB medium to a final concentration of 10^5 CFU/mL. Each 96-well
697 spectrophotometric plate (NUNC, Thermo Scientific™ 267334) contained most phages
698 confronted with a single bacterial strain, with positive control wells containing 180 µL of
699 bacterial suspension and 20 µL of SM buffer. Experimental wells comprised 180 µL of
700 bacterial suspension and 20 µL of phage suspension at 10^7 PFU/mL, establishing a
701 multiplicity of infection (MOI) of 10. Positive control measurements included between two
702 and seven technical replicates per biological replicate (except for bacterial strains 41A and
703 93D, which had only one technical replicate per biological replicate). Measurements were
704 taken using the SPECTROstar® Omega spectrophotometer (BMG Labtech, France) every 5
705 minutes for 48 hours at 25°C, with intermittent orbital shaking at 200 rpm (200 seconds
706 "ON" followed by 400 seconds "OFF").

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

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

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

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

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

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

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

736 A probability value below 0.05 indicates a significant growth reduction in presence of phage,
737 while a value exceeding 0.95 indicates a significant growth stimulation of bacteria in
738 presence of phage. The distribution of inhibition percentages within the 0.05 to 0.95
739 probability range is approximately symmetric and unimodal (S8 Fig), resembling a Gaussian
740 distribution. However, the Shapiro-Wilk test indicated a significant deviation from strict
741 normality ($p < 0.01$), likely due to the large sample size. Despite this, the distribution shape
742 supports the use of this probability window to robustly summarise phage effects on bacterial
743 growth. Bacterial strains 41A and 93D had only one biological replicate, precluding
744 probability calculation.

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

749 *Statistical analysis*

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

751 The analysis of networks representing bacterial growth inhibition percentages of 44 bacterial
752 strains combined with 23 phages was computed using the mean of the three biological
753 replicates for each combination. The analyses were performed using the R software (version
754 4.4.0) with the *bipartite* (for nestedness and modularity analysis) and *igraph* (for modularity
755 analysis) packages, following Moury *et al.* (32) and Torres-Barceló *et al.* (48).

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

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

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

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

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

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

782 The bacterial susceptibility profiles were tested with a PERmutational Multivariate ANalysis
783 Of VAriance (PERMANOVA), a non-parametric method that assesses the influence of
784 categorical variables on multivariate data by partitioning variation within a distance matrix.
785 Unlike traditional ANOVA, it makes no assumptions about data distribution and relies on
786 permutations to evaluate the statistical significance of group differences (90). Here, Bray–
787 Curtis dissimilarity was used to construct a distance matrix based on the mean percentage of
788 bacterial growth inhibition (from three replicates per combination). This matrix was then
789 analysed using the *adonis2* function from the *vegan* R package, with 9,999 permutations to
790 ensure robust p-value estimation (91). "*Phylogroup", "ecological origin",* and "*whether the*
791 *strain was an isolation bacterial strain for the phages*" were included as explanatory
792 variables. Multiple comparisons were conducted using the function *pairwise.adonis2* from

793 the *pairwiseAdonis* package (92), which is specifically adapted for PERMANOVA post-hoc
794 testing.

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

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

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

811 The ranking of phage genera based on inhibition percentage (offset by +0.1 to avoid zeros)
812 was assessed using a GLMM with a Gamma distribution with a log-link, suitable for strictly
813 positive, continuous, and skewed data. The model included a random effect accounting for
814 the "*phage*" tested, the "*bacterial*" strain tested, and "*replicates*" nested within each

815 *"phage:bacterium"* combination. No overdispersion was detected (χ^2/ddl = 0.67). Post-hoc
816 pairwise comparisons between genera were performed using Tukey's test.

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

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

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

827 The total number of bacterial defence systems per bacterial genome and the abundance of
828 prophages were tested against the explanatory variables *"phylogroup"*, *"ecological origin",*
829 and *"growth rate"* in a GLM with a Poisson distribution with an additive model of the three
830 explanatory variables. No overdispersion was observed in the data (deviance/degrees of
831 freedom for defence systems = 0.61; for abundance of prophages = 0.47), and the effects of
832 each explanatory variable were assessed using Type III ANOVA. Finally, the post-hoc pairwise
833 comparisons were performed using Tukey's test and significance groups were assigned
834 accordingly.

835 The defence system profiles, the proportions of the defence system categories, and
836 prophages distribution profile across bacterial strains were analysed using a PERMANOVA
837 test based on a Bray-Curtis dissimilarity matrix with 9,999 permutations. *"Phylogroup"*,

838 *"ecological origin"*, and *"growth rate"* were included as explanatory variables. Pairwise
839 comparisons of the defence system and prophage composition between groups were
840 conducted using the *pairwise.adonis2* function (92).

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

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

846 To account for the widest range of explanatory variables influencing the percentage of
847 inhibition, we performed a multimodel analysis coupled with a generalised linear mixed
848 model (GLMM). To enhance model efficiency, the number of explanatory variables was
849 reduced through a multi-step process. Variables with more than 20% missing data were
850 excluded. Redundant variables were then removed based on their correlations, assessed
851 using Spearman's correlation for quantitative variables, chi-squared tests for categorical
852 variables, and Kruskal-Wallis tests for associations between categorical and quantitative
853 variables. To further refine the selection, Principal Component Analyses (PCA) were
854 conducted separately on the matrices representing the distributions of defence systems and
855 prophages across bacterial genomes. The most informative variables contributing to the first
856 principal components were retained for subsequent modelling.

857 The PCA on defence systems was performed without centring or scaling the data. Two
858 dimensions accounted for 32.5% of the variance and led to the selection of the variables
859 *"Septu"*, *"Gabija"*, and *"RloC"*, which were among the most frequent defence systems in the
860 genomes. Similarly, the PCA on the distribution of prophage species in bacterial genomes,

861 also conducted without centring or scaling (already binary), captured 14% of the variance in
862 the first two dimensions and identified two perpendicular variable groups composed
863 respectively of prophage species *"17"*, *"54"*, *"26",* and *"29"* on one side, and species *"7"*,
864 *"47"*, *"57"*, *"22"*, and *"31"* on the other. One representative species from each group was
865 selected for the multimodel analysis (species *"17"* and *"22"*).

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

870 To compare two approaches for modelling the effects of bacterial defence systems and
871 prophages on the percentage of inhibition, two separate GLMMs were constructed: one
872 using total counts of defence systems and prophages, and another incorporating principal
873 components PCA-derived of the distribution matrices (of prophages and defence systems).
874 No overdispersion was detected in either model (total counts: $\chi^2/df = 0.65$; PCA-derived:
875 $\chi^2/df = 0.66$), supporting the appropriateness of the Gamma distribution for modelling the
876 response variable.

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

883 estimated using the *sw* function, which computes the sum of the AIC weights across all
884 models in which the variable appears.

885

886 **Acknowledgments**

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

891

892

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

940 19. Czajkowski R. May the phage be with you? Prophage-like elements in the genomes of
941 soft rot Pectobacteriaceae: *Pectobacterium* spp. and *Dickeya* spp. Front Microbiol
942 [Internet]. 2019 Feb 14 [cited 2025 Apr 28];10. Available from:
943 https://www.frontiersin.orghttps://www.frontiersin.org/journals/microbiology/articles/
944 10.3389/fmicb.2019.00138/full

945 20. Dominguez-Mirazo M, Jin R, Weitz JS. Functional and comparative genomic analysis of
946 integrated prophage-like sequences in "*Candidatus Liberibacter asiaticus*". mSphere.
947 2019 Nov 13;4(6):10.1128/msphere.00409-19.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

993 36. Morris CE, Sands DC, Vanneste JL, Montarry J, Oakley B, Guilbaud C, et al. Inferring the
994 evolutionary history of the plant pathogen *Pseudomonas syringae* from its biogeography
995 in headwaters of rivers in North America, Europe, and New Zealand. mBio. 2010 June
996 29;1(3):e00107-10.

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

999 38. Cuppels DA. Isolation and characterization of phages useful for identifying *Pseudomonas*
1000 *syringae pv. tomato*. Phytopathology. 1983 Apr;(73):1376–81.

1001 39. Gross DC, Cody YS, Proebsting EL, Radamaker GK, Spotts RA. Ecotypes and pathogenicity
1002 of ice-nucleation-active *Pseudomonas syringae* isolated from deciduous fruit tree
1003 orchards. 1983 Sept;(74):241–8.

1004 40. Cuppels DA. The use of pathovar-indicative bacteriophages for rapidly detecting
1005 *Pseudomonas syringae pv. tomato* in tomato leaf and fruit lesions. 1984 Feb;(74):891–4.

1006 41. Taylor JD, Teverson DM, Allen DJ, Pastor-Corrales MA. Identification and origin of races
1007 of *Pseudomonas syringae pv. phaseolicola* from Africa and other bean growing areas.
1008 Plant Pathol. 1996;45(3):469–78.

1009 42. Frampton RA, Taylor C, Holguín Moreno AV, Visnovsky SB, Petty NK, Pitman AR, et al.
1010 Identification of bacteriophages for biocontrol of the kiwifruit canker phytopathogen
1011 *Pseudomonas syringae pv. actinidiae*. Appl Environ Microbiol. 2014 Apr;80(7):2216–28.

1012 43. James SL, Rabiey M, Neuman BW, Percival G, Jackson RW. Isolation, characterisation and
1013 experimental evolution of phage that infect the horse chestnut tree pathogen,
1014 *Pseudomonas syringae pv. aesculi*. Curr Microbiol. 2020 Aug 1;77(8):1438–47.

1015 44. Rombouts S, Volckaert A, Venneman S, Declercq B, Vandenheuvel D, Allonsius CN, et al.
1016 Characterization of novel bacteriophages for biocontrol of bacterial blight in leek caused
1017 by *Pseudomonas syringae pv. porri*. Front Microbiol [Internet]. 2016 Mar 15 [cited 2024
1018 Oct 8];7. Available from:
1019 https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2016.00279/f
1020 ull

1021 45. Parisi L, Morgaint B, Blanco-Garcia J, Guilbaud C, Chandeysson C, Bourgeay JF, et al.
1022 Bacteria from four phylogroups of the *Pseudomonas syringae* complex can cause
1023 bacterial canker of apricot. Plant Pathol. 2019 Sept;68(7):1249–58.

1024 46. Berge O, Monteil CL, Bartoli C, Chandeysson C, Guilbaud C, Sands DC, et al. A user's guide
1025 to a data base of the diversity of *Pseudomonas syringae* and its application to classifying
1026 strains in this phylogenetic complex. PLOS ONE. 2014 Sept 3;9(9):e105547.

1027 47. Omrani M, Roth M, Roch G, Blanc A, Morris CE, Audergon JM. Genome-wide association
1028 multi-locus and multi-variate linear mixed models reveal two linked loci with major
1029 effects on partial resistance of apricot to bacterial canker. BMC Plant Biol. 2019 Jan
1030 21;19(1):31.

1031 48. Torres-Barceló C, Boyer C, Garneau JR, Poussier S, Robène I, Moury B. A Phylogenetic
1032 Host-Range Index Reveals Ecological Constraints in Phage Specialisation and Virulence.
1033 Mol Ecol. 2025 July 28;n/a(n/a):e70052.

1034 49. van Houte S, Buckling A, Westra ER. Evolutionary ecology of prokaryotic immune
1035 mechanisms. Microbiol Mol Biol Rev. 2016 July 13;80(3):745–63.

1036 50. Turner D, Kropinski AM, Adriaenssens EM. A roadmap for genome-based phage
1037 taxonomy. Viruses. 2021 Mar;13(3):506.

1038 51. Mavrodi DV, Loper JE, Paulsen IT, Thomashow LS. Mobile genetic elements in the
1039 genome of the beneficial rhizobacterium *Pseudomonas fluorescens* Pf-5. BMC Microbiol.
1040 2009 Jan 13;9(1):8.

1041 52. Hockett KL, Renner T, Baltrus DA. Independent Co-Option of a Tailed Bacteriophage into
1042 a Killing Complex in *Pseudomonas*. mBio. 2015 Aug 11;6(4):10.1128/mbio.00452-15.

1043 53. Backman T, Latorre SM, Symeonidi E, Muszyński A, Bleak E, Eads L, et al. A phage tail–like
1044 bacteriocin suppresses competitors in metapopulations of pathogenic bacteria. Science.
1045 2024 June 14;384(6701):eado0713.

1046 54. Fautt C, Hockett K, Delattre S, Baltrus D, Couradeau E. Tailocin tail fiber diversity
1047 correlates with rfbD variation in the *Pseudomonas syringae* species complex. ISME
1048 Commun. 2025 Jan 1;5(1):ycaf099.

1049 55. Hunter M, Fusco D. Superinfection exclusion: A viral strategy with short-term benefits
1050 and long-term drawbacks. PLOS Comput Biol. 2022 May 10;18(5):e1010125.

1051 56. Koskella B, Lin DM, Buckling A, Thompson JN. The costs of evolving resistance in
1052 heterogeneous parasite environments. Proc R Soc B Biol Sci. 2011 Dec
1053 14;279(1735):1896–903.

1054 57. Monteil CL, Yahara K, Studholme DJ, Mageiros L, Méric G, Swingle B, et al. Population-
1055 genomic insights into emergence, crop adaptation and dissemination of *Pseudomonas*
1056 *syringae* pathogens. Microb Genomics [Internet]. 2016 Oct 1 [cited 2021 Sept 15];2(10).
1057 Available from:
1058 https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000089

1059 58. Koonin EV, Makarova KS, Wolf YI. Evolutionary genomics of defense systems in archaea
1060 and bacteria. Annu Rev Microbiol. 2017 Sept 8;71(Volume 71, 2017):233–61.

1061 59. Bernheim A, Sorek R. The pan-immune system of bacteria: antiviral defence as a
1062 community resource. Nat Rev Microbiol. 2020 Feb;18(2):113–9.

1063 60. Makarova KS, Wolf YI, Koonin EV. Comparative genomics of defense systems in archaea
1064 and bacteria. Nucleic Acids Res. 2013 Apr 1;41(8):4360–77.

1065 61. Nobrega FL, Vlot M, de Jonge PA, Dreesens LL, Beaumont HJE, Lavigne R, et al. Targeting
1066 mechanisms of tailed bacteriophages. Nat Rev Microbiol. 2018 Dec;16(12):760–73.

1067 62. Bertozzi Silva J, Storms Z, Sauvageau D. Host receptors for bacteriophage adsorption.
1068 FEMS Microbiol Lett. 2016 Feb 1;363(4):fnw002.

1069 63. Müller DM, Pourtois JD, Kim MK, Targ B, Burgener EB, Milla C, et al. Bacterial receptors
1070 but not anti-phage defence mechanisms determine host range for a pair of
1071 *Pseudomonas aeruginosa* lytic phages [Internet]. bioRxiv; 2024 [cited 2024 Sept 23]. p.
1072 2024.04.30.591980. Available from:
1073 https://www.biorxiv.org/content/10.1101/2024.04.30.591980v2

1074 64. Kering KK, Kibii BJ, Wei H. Biocontrol of phytobacteria with bacteriophage cocktails. Pest
1075 Manag Sci. 2019;75(7):1775–81.

1076 65. Feltin C, Garneau JR, Morris CE, Bérard A, Torres-Barceló C. Novel phages of
1077 *Pseudomonas syringae* unveil numerous potential auxiliary metabolic genes. J Gen Virol.
1078 2024;105(6):001990.

1079 66. King EO, Ward MK, Raney DE. Two simple media for the demonstration of pyocyanin and
1080 fluorescin. J Lab Clin Med. 1954 Aug 1;44(2):301–7.

1081 67. Dieckmann MA, Beyvers S, Nkouamedjo-Fankep RC, Hanel PHG, Jelonek L, Blom J, et al.
1082 EDGAR3.0: comparative genomics and phylogenomics on a scalable infrastructure.
1083 Nucleic Acids Res. 2021 July 2;49(W1):W185–92.

1084 68. Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum-likelihood trees for
1085 large alignments. PLOS ONE. 2010 Mar 10;5(3):e9490.

1086 69. Tesson F, Hervé A, Mordret E, Touchon M, d'Humières C, Cury J, et al. Systematic and
1087 quantitative view of the antiviral arsenal of prokaryotes. Nat Commun. 2022 May
1088 10;13(1):2561.

1089 70. Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, Hu B, et al. Identification of mobile
1090 genetic elements with geNomad. Nat Biotechnol. 2023 Sept 21;1–10.

1091 71. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpides NC. CheckV assesses
1092 the quality and completeness of metagenome-assembled viral genomes. Nat Biotechnol.
1093 2021 May;39(5):578–85.

1094 72. Moraru C, Varsani A, Kropinski AM. VIRIDIC—a novel tool to calculate the intergenomic
1095 similarities of prokaryote-infecting viruses. Viruses. 2020 Nov 6;12(11):1268.

1096 73. Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H, Goto S. ViPTree: the viral
1097 proteomic tree server. Bioinforma Oxf Engl. 2017 Aug 1;33(15):2379–80.

1098 74. Orfei B, Pothier JF, Fenske L, Blom J, Moretti C, Buonaurio R, et al. Race-specific
1099 genotypes of *Pseudomonas syringae pv. tomato* are defined by the presence of mobile
1100 DNA elements within the genome. Front Plant Sci [Internet]. 2023 July 5 [cited 2025 June
1101 7];14. Available from: https://www.frontiersin.org/journals/plant-
1102 science/articles/10.3389/fpls.2023.1197706/full

1103 75. Balbuena JA, Míguez-Lozano R, Blasco-Costa I. PACo: A Novel Procrustes Application to
1104 Cophylogenetic Analysis. PLOS ONE. 2013 Apr 8;8(4):e61048.

1105 76. Revell LJ. phytools 2.0: an updated R ecosystem for phylogenetic comparative methods
1106 (and other things). PeerJ. 2024 Jan 5;12:e16505.

1107 77. Xie Y, Wahab L, Gill JJ. Development and validation of a microtiter plate-based assay for
1108 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

1143 93. Brooks M E, Kristensen K, Benthem K J ,van, Magnusson A, Berg C W, Nielsen A, et al.
1144 glmmTMB balances speed and flexibility among packages for zero-inflated generalized
1145 linear mixed modeling. R J. 2017;9(2):378.

1146 94. Allen B, Kon M, Bar-Yam Y. A new phylogenetic diversity measure generalizing the
1147 Shannon index and its application to phyllostomid bats. Am Nat. 2009 Aug;174(2):236–
1148 43.

1149

1150 ## Supporting information

1151 **S1 Fig. Bacterial resistance to phages and host range breadth of the tested phages.**

1152 (A) Number of phages to which bacteria are sensitive or resistant, with the phylogeny of the
1153 strains indicated at the bottom of the histogram. (B) Number of bacteria infected or not by
1154 the phages, with the taxonomy of the phages listed at the bottom of the histogram. Values
1155 are derived from the binary inhibition matrix.

1156 **S2 Fig. Taxonomic and phylogenetic characteristics, and ecological origins of phages and**
1157 **bacteria explaining host range.**

1158 (A) Bacterial percentage of inhibition by phages across bacterial phylogroups and ecological
1159 origins. Significance letters indicate results from the PERMANOVA test and associated
1160 pairwise comparisons. (B) Bacterial inhibition percentage of different phage genera.
1161 Significance letters represent results from a GLMM with a Gamma regression distribution
1162 and random effect (phage, bacteria, and the replicate imbricated in each of the phage-
1163 bacteria interactions), and post hoc pairwise comparisons performed using Tukey's test.

1164 **S3 Fig. Relative proportions of defence systems by functional category (%) and total**
1165 **number of defence systems across bacterial genomes.**

1166 On the left, a phylogenetic tree based on the core genomes of the bacterial strains is shown,
1167 with corresponding phylogroups indicated in the adjacent column.

1168 **S4 Fig. Percentage of similarity of defence system subtypes among bacterial strains in**
1169 **relation to their genome-wide average nucleotide identity (ANI).**

1170 **S5 Fig. Co-phylogeny of the CL4 prophage family-like (R-type syringacin).**

1171 Co-phylogenetic analysis between 68 strains of the *P. syringae* species complex (left),
1172 represented by a maximum likelihood phylogenetic tree based on their core-genomes, and
1173 their associated 68 CL4 family-like prophages (right), represented by a proteomic tree. The
1174 44 bacterial strains included in this study are marked with dots at the tips of the bacterial
1175 tree. Phylogroups and prophage-carrying strains are colour-coded accordingly. Two nodes
1176 with a maximum likelihood support value below 0.90 are marked with red dots.

1177 **S6 Fig. Experimental setup illustrating bacterial growth inhibition by phages.**

1178 Each biological replicate per bacterial strain is represented by a different colour. For clarity,
1179 only 3 of the 23 tested phages are shown. Three types of internal controls are included:
1180 positive controls with bacteria only, negative controls with medium only, and internal
1181 reference controls using the bacterial strain P1.01.01B03 (growth shown in S7B Fig) and the
1182 internal phage Ghuch01, which consistently inhibits this strain (see S7C Fig).

1183 **S7 Fig. Analysis of the different controls used in the phage host range *in vitro* tests.**

1184 (A) Analysis of culture medium controls by plate (labelled by bacteria). (B) The internal
1185 bacterial control, *P. syringae* strain P1.01.01B03, with an area under the empirical curve
1186 (AUC) threshold of 36 (first quartile), shows that for seven dates, all three replicates are
1187 below this threshold (the dates are marked with a red dot). (C) The internal phage control,

1188 Ghuch01, with an inhibition percentage threshold of 98% (first quartile), indicates that for
1189 two dates, all three replicates are below this threshold (the dates are marked with a red
1190 dot).

1191 **S8 Fig. Histogram showing the distribution of inhibition percentages within the 0.05 to**
1192 **0.95 probability range, displaying an approximately symmetric and unimodal shape.**

1193 **S9 Fig. Results of additional experiments on five potentially stimulating phages and five**
1194 **potentially stimulated bacteria**.

1195 Generally, bacterial populations decreased in the presence of phages. In 7% of observed
1196 phage-bacteria combinations, a "stimulatory" effect was identified, with inhibition values
1197 reaching up to -40%. Five phages showing strong stimulatory effect and five bacteria
1198 showing significant stimulation were chosen to test the reproducibility of the stimulation
1199 effect in subsequent experiments. (A) Matrix of mean inhibition percentages for the five *P.*
1200 *syringae* strains (rows) and five phages (columns), with their phylogenetic and taxonomic
1201 affiliations indicated by colour. The number of stars indicates the number of replicates
1202 showing a significantly inhibiting or stimulating percentage. Of ten initially identified
1203 significant stimulating interactions (delimited by red lines), only three were confirmed in
1204 subsequent experiments. (B) Size of the five bacterial populations in response to the phages.
1205 Stars correspond to significant differences from the control (no phage) based on the results
1206 of the GLM model, family = Gamma. Arrows indicate interactions with at least one of the
1207 replicates showing a significant inhibition or stimulation percentage. Out of the three
1208 significant interactions of stimulatory phenotype, only 2 phage-bacteria interactions resulted
1209 in slightly higher bacterial densities in the presence of phages compared to phage-free
1210 conditions, though this difference was not statistically significant. (C) Spearman correlation

1211 between bacterial concentration and inhibition percentage, by bacterium and phage. No
1212 correlation was significant and bacterial population size only poorly explains the inhibition
1213 percentages.

1214

1215 **S1 Table. Phage-bacterium matrices representing the percentage of inhibition of bacterial**
1216 **growth (A) and the statistical significance of inhibition (B).**

1217 Rows correspond to bacterial strains, while columns represent phage clones, each annotated
1218 with their respective phylogroups and genera.

1219 **S2 Table: Analysis of nestedness and modularity of the network of binary data**
1220 **representing the significance or non-significance of bacterial growth inhibition by phages**
1221 **(S1B Table).**

1222 Null models are described in Moury *et al.* (32) for B, N, C2, R2 and S, and in Moury *et al.* (89)
1223 for PD. For each null model and each algorithm, the frequency of simulations, among 100,
1224 showing a strictly higher nestedness (or modularity) score than that of the actual matrix is
1225 indicated. Significantly nested or modular networks ($p \leq 0.05$) are indicated in red.

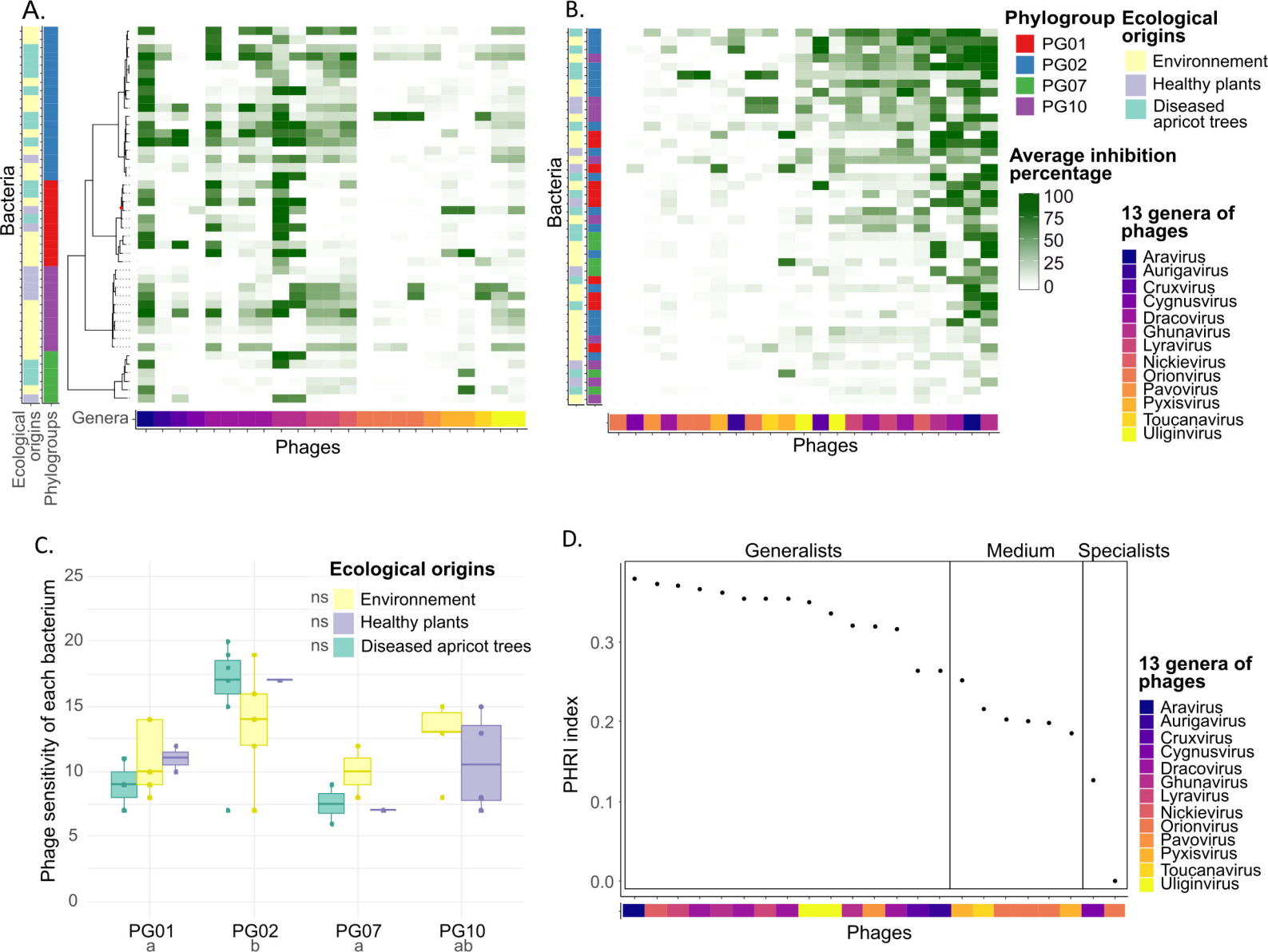
1226 **S3 Table. Genetic analysis of the 88 prophages identified in the 44 *P. syringae* strains using**
1227 **VIRIDIC.**

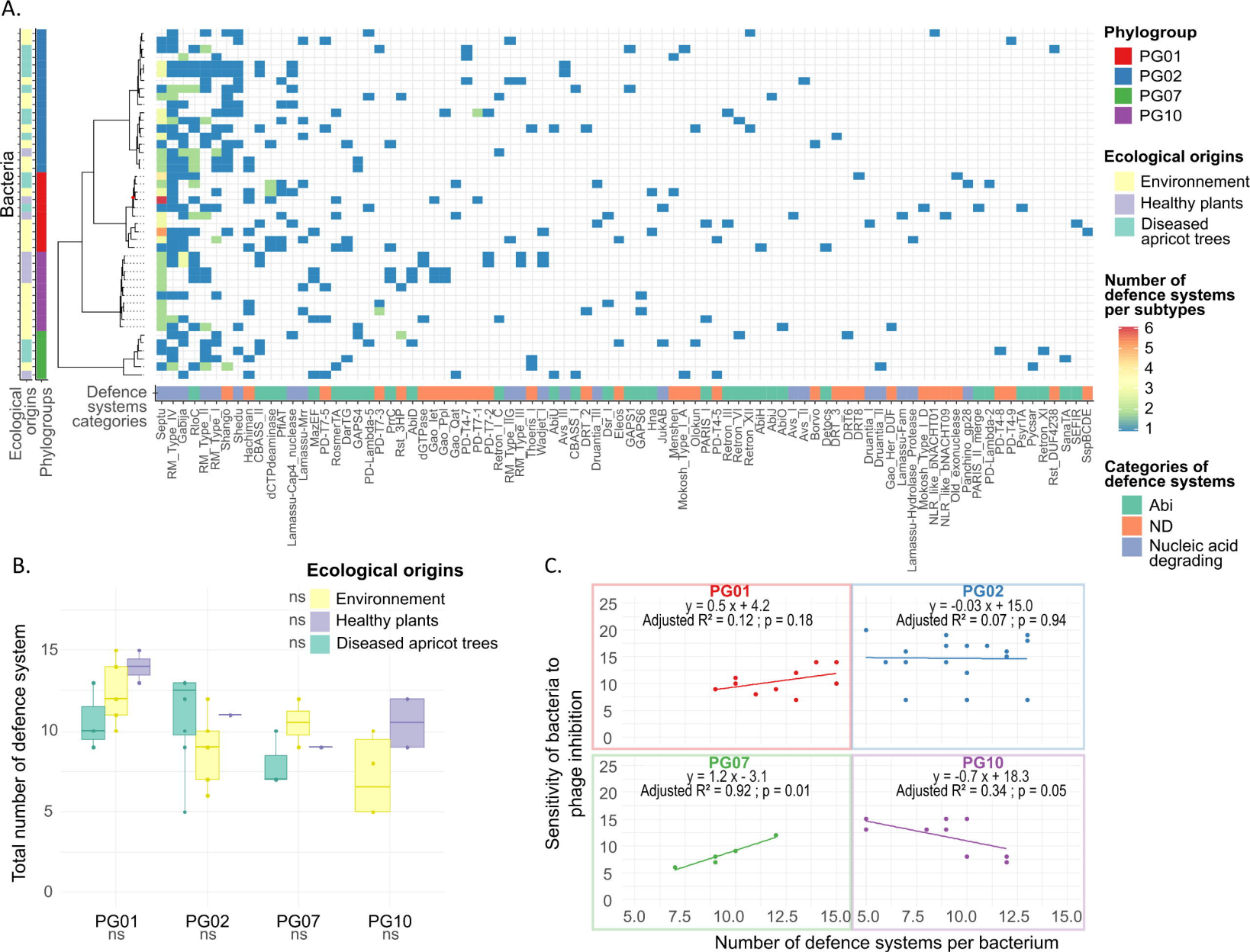
1228 The 27 reference phages are highlighted in yellow (25 from the collection described in Feltin
1229 *et al.* (65) and 3 from NCBI). Species and genus-level clusters of the prophages are indicated
1230 on the right in colour. High genetic similarities are highlighted in green.

1231 **S4 Table. Phylogenetic characteristics, ecological origins, countries, and dates of sampling**
1232 **of the 44 *P. syringae* strains used for the host range study.**

1233 The four *P. syringae* strains highlighted in bold are strains that were used to isolate some of
1234 the phages tested in the host range assay. Four strains are in the CIRM-CFBP, the French
1235 collection for plant-associated bacteria (https://cirm-cfbp.fr/ ; June 2025): USA0007(CFBP
1236 8574), CC1559 (CFBP 8551), MAFF302273PT (CFBP 2339), and CC1583 (CFBP 8564).

1237


A.
Bacteria
Genera
Phages
Ecological origins
Phylogroups
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
25
20
15
10
5
0
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
0.3
0.2
0.1
0.0
Phages
13 genera of phages
Aravirus
Aurigavirus
Cruxvirus
Cygnusvirus
Dracovirus
Ghunavirus
Lyravirus
Nickievirus
Orionvirus
Pavovirus
Pyxisvirus
Toucanavirus
Uliginvirus


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

A.

Bacteria

Ecological origins

Phylogroups

5 2 35 40 42 10 12 27 37 38 41 8 14 23 26 34 36 1 3 4 6 7 9 11 13 15 17 18 19 20 21 22 24 28 29 30 31 32 33 39 16 25

Phylogroup

- PG01
- PG02
- PG07
- PG10

Ecological origins

- Environnement
- Healthy plants
- Diseased apricot trees

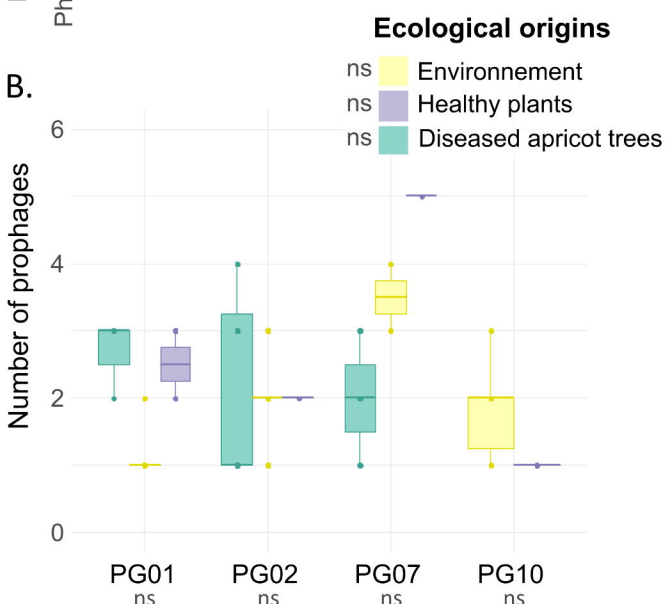
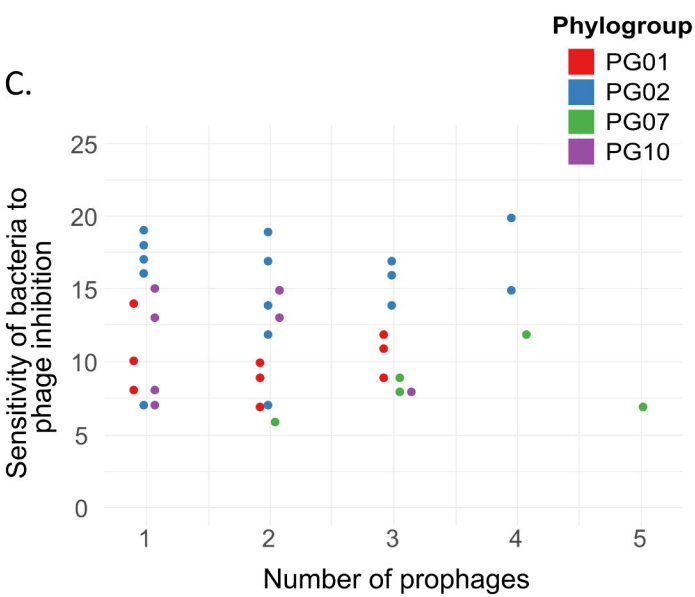
Number of prophages

- 1
- 0

Cluster of prophages

- CL1
- CL2
- CL3
- CL4
- CL5
- CL6

B.



C.