



**HAL**  
open science

# Large-scale identification of viral quorum sensing systems reveals density-dependent sporulation-hijacking mechanisms in bacteriophages

Charles Bernard, Yanyan Li, Philippe Lopez, Eric Bapteste

► **To cite this version:**

Charles Bernard, Yanyan Li, Philippe Lopez, Eric Bapteste. Large-scale identification of viral quorum sensing systems reveals density-dependent sporulation-hijacking mechanisms in bacteriophages. 2021. hal-03384404

**HAL Id: hal-03384404**

**<https://hal.science/hal-03384404>**

Preprint submitted on 20 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Large-scale identification of viral quorum sensing systems reveals density-dependent**  
2 **sporulation-hijacking mechanisms in bacteriophages**

3

4 **AUTHORS**

5 Charles Bernard <sup>1,2,\*</sup>, Yanyan Li <sup>2</sup>, Philippe Lopez <sup>1</sup> and Eric Bapteste <sup>1</sup>

6

7 **AFFILIATIONS**

8 <sup>1</sup> Institut de Systématique, Evolution, Biodiversité (ISYEB), Sorbonne Université, CNRS, Museum

9 National d'Histoire Naturelle, EPHE, Université des Antilles, Campus Jussieu, Bâtiment A, 4eme et.

10 Pièce 429, 75005 Paris, France

11 <sup>2</sup> Unité Molécules de Communication et Adaptation des Micro-organismes (MCAM), CNRS, Museum

12 National d'Histoire Naturelle, CP 54, 57 rue Cuvier, 75005 Paris, France

13

14 **CORRESPONDING AUTHOR**

15 \* Correspondence to Charles Bernard (ORCID Number: 0000-0002-8354-5350);

16 Phone: +33 (01) 44 27 34 70; E-mail address: [charles.bernard@cri-paris.org](mailto:charles.bernard@cri-paris.org)

17

18 **SIGNIFICANCE**

19 Communication between viruses supported by quorum sensing systems (QSSs) is a brand new  
20 research area that has transformed our views of viral adaptation and virus-host co-evolution. The viral  
21 QSSs discovered so far were found to guide the lysis-lysogeny decision in temperate bacteriophages  
22 as a function of phage density. Here, we identified that quorum sensing-mediated communication  
23 between phages can not only guide the regulation of viral processes but also the manipulation of the  
24 bacterial sporulation pathway. Our finding introduces the new view that not only bacteria decide when  
25 it is time to sporulate, some bacteriophages are also key stakeholders in this dynamical decision-  
26 making process. Considering that spores are the transmissive form of many pathogens, these new  
27 insights have important applied implications.

28

29

30

## 31 ABSTRACT

32 Communication between viruses supported by quorum sensing systems (QSSs) were found to  
33 optimize the fitness of temperate bacteriophages of Bacilli by guiding the transition from the host-  
34 destructive lytic cycle to the host-protective lysogenic cycle in a density-dependent manner. All known  
35 phage-encoded QSSs consist of a communication propeptide and a cognate intracellular receptor that  
36 regulates the expression of adjacent target genes upon recognition of the matured peptide, a  
37 signature known as RRNPP and found in chromosomes, plasmids and phages of Firmicutes bacteria.  
38 Recently, we have introduced the RRNPP\_detector software to detect novel genetic systems matching  
39 the RRNPP signature, which unearthed many novel phage-encoded candidate QSSs. Here, by  
40 looking at the adjacent genes likely regulated by these viral candidate QSSs, we identified an  
41 unsuspected clustering of viral QSSs with viral genes whose bacterial homologs are key regulators of  
42 the last-resort bacterial sporulation initiation pathway (*rap*, *spo0E* or *abrB*). Consistently, we found  
43 evidence in published data that certain of these QSSs encoded by prophages (phage genomes  
44 inserted within a bacterial genome) dynamically manipulate the timing of sporulation in the host.  
45 Because these viral QSSs are genetically diverse and are found associated with different sporulation  
46 regulators, this suggests a convergent evolution in bacteriophages of density-dependent sporulation-  
47 hijacking mechanisms.

48

## 49 INTRODUCTION

50 If cell-cell communication via quorum sensing was discovered in 1970 in bacteria, the first  
51 characterization of a functional viral quorum sensing systems (QSSs) only dates back to 2017 (1).  
52 Erez *et al.* discovered that certain *Bacillus* phages encode a communication propeptide, which upon  
53 expression, secretion, and maturation by the host cellular machinery, accumulates in the extracellular  
54 environment. Accordingly, during the lytic cycle, the extracellular concentration of this phage-encoded  
55 peptide reflects the number of host cells that have expressed the QSS-encoding genome of the  
56 phage, which correlates with the number of hosts that have been lysed by the phage (**Fig1**). At high  
57 concentrations of this peptide, when lots of hosts have been lysed and the survival of the phage-host  
58 collective can be endangered, the imported-peptide is transduced by its cognate phage-encoded  
59 intracellular receptor, which coordinates a population-wide transition from the host-destructive lytic  
60 cycle to the host-protective lysogenic cycle. Indeed, during the lysogenic cycle, the phage genome is  
61 inserted within the genome of the host, replicates as part of it, and often confers upon the host an  
62 immunity towards free virions This viral QSS, coined arbitrium, thereby optimizes the fitness of a  
63 phage with respect to its social context.

64 Arbitrium QSSs are classified as RRNPP QSSs, which consist of a communication propeptide  
65 (secreted and matured into a quorum sensing peptide) coupled with an intracellular receptor, turned-  
66 on or -off upon binding with the imported mature peptide (**Fig1**) (2). RRNPP receptors often are

67 transcription factors that dynamically regulate the expression of adjacent genes. As genetically diverse  
68 as RRNPP QSSs are in chromosomes, plasmids and phages of Firmicutes, they however function  
69 according to the same canonical mechanism (**Fig1**), except with very few exceptions in the mode of  
70 secretion (Shp and PrgQ peptides) (2). We recently identified that this shared mechanism between  
71 different RRNPP QSSs underlies a common signature that can be detected in-silico (3). On this basis,  
72 we have developed RRNPP\_detector, a tool designed to identify novel candidate QSSs matching the  
73 RRNPP signature (3).

74 In a large-scale application of RRNPP\_detector, we identified many novel candidate RRNPP QSSs  
75 in sequenced genomes of temperate bacteriophages and in prophages inserted within bacterial  
76 genomes (3). During the development of RRNPP\_detector, we noticed an unsuspected clustering of  
77 viral candidate QSSs with adjacent viral genes whose bacterial homologs are key regulators of the  
78 bacterial sporulation initiation pathway. Because RRNPP QSSs tend to dynamically regulate adjacent  
79 genes (1, 2, 4, 5), this hinted at density-dependent manipulations of bacterial sporulation by phages.

80 In Firmicutes, sporulation leads to the formation of endospores, able to resist extreme  
81 environmental stresses for prolonged periods and to resume vegetative growth in response to  
82 favorable changes in environmental conditions (6). The sporulation pathway is initiated when  
83 transmembrane kinases sense stress stimuli, and thereupon transfer their phosphate, either directly  
84 (*Clostridium*) or via phosphorelay (*Bacillus*, *Brevibacillus*) to Spo0A, the master regulator of  
85 sporulation (7, 8). However, only high Spo0A-P concentrations, and therefore intense stresses, can  
86 commit a cell to sporulate (9). Because sporulation is costly, Spo0A-P accumulation is subjected to  
87 multiple regulative check points by the Rap, Spo0E and AbrB proteins (10, 11). In adverse  
88 circumstances, Rap, Spo0E and AbrB thereby form a decision-making regulatory circuit that controls  
89 the timing of sporulation in Bacilli (11).

90 Here, we present in more details our results regarding the association of sporulation regulators  
91 (either *rap*, *spo0E* or *abrB*) with candidate QSSs found in temperate bacteriophages of Firmicutes,  
92 their mechanistic consequence on the host biology, and their fundamental and applied implications.

93

## 94 **RESULTS AND DISCUSSION**

### 95 **Identification of 384 viral candidate RRNPP QSSs, distributed into 26 families, of which only 6** 96 **were previously known**

97 To detect viral QSSs, we followed the study design displayed in **FigS1**. We applied  
98 RRNPP\_detector (15-65aa and 250-460aa length thresholds for propeptides and receptors) against  
99 the Gut Phage Database (12), 32,327 NCBI complete genomes of Viruses, and 3,577 NCBI complete  
100 genomes of Firmicutes (because phage genomes can be inserted within bacterial genomes). This led  
101 to the identification of 16 candidate RRNPP QSSs on intestinal phage genomes, 10 on sequenced  
102 genomes of temperate phages and 2671 on bacterial genomes, respectively. Prophage regions within

103 bacterial genomes were subsequently detected by Phaster (13) and Prophage Hunter (14) to  
104 distinguish between genuine bacterial QSSs and prophage-encoded QSSs inserted within bacterial  
105 chromosomes or plasmids. This enabled identifying 358 additional candidate viral QSSs: 174 on  
106 intact/active prophages, 68 on questionable/ambiguous prophages, 116 on incomplete prophages.  
107 The genomic and taxonomic details of these 384 viral candidate RRNPP QSSs can be found in  
108 **TableS1**.

109 In a blast all vs all, the receptors of all the 2697 candidate QSSs were further classified into 64  
110 groups of homologs / QSS families (sequence identity  $\geq 30\%$ ; mutual length coverage of the  
111 alignment  $\geq 80\%$ ), as in (3). Of these 64 QSS families, 26 families comprised at least one  
112 phage/prophage-encoded QSS and served as the focus of this study (metadata of the 26 families in  
113 **TableS1**; distribution of the 384 viral candidate RRNPP QSSs into these 26 families in **Fig2**). 6 of  
114 these 26 QSS families had already been described prior to our study: Rap-Npr of *Bacilli* (2, 15), AimR  
115 of *B. subtilis* phages (1), AloR of *Paenibacilli* (16), AimR of *B. cereus* phages (5), the family of QSSs  
116 characterized by Feng *et al.* in *Clostridium saccharoperbutylacetonicum* (17), the family of QSSs  
117 characterized by Kotte *et al.* in *Clostridium acetobutylicum* (18) (**Fig2**). Accordingly, the 20 other  
118 families show great promise to expand the known diversity of viral RRNPP QSSs. As expected, the  
119 arbitrium systems of *B. subtilis* (N=15) and *B. cereus* (N=9) phages form two 100% viral families  
120 (**Fig2B**). The two biggest families, Rap-Npr of Bacilli (N=2258) and the novel candidate QSS1 family  
121 of Brevibacilli (N=18) are found to be shared between bacterial chromosomes, plasmids and  
122 phages/prophages (**Fig2**). The Rap-Phr family notably is the family in which (pro)phage-encoded  
123 QSSs are the most prevalent (326 viral Rap-Phr QSSs (**TableS1 and Fig2A**)).

124

### 125 **Rap-Phr QSSs that delay the timing of sporulation are found in many, diverse Bacillus** 126 **bacteriophages**

127 Importantly, bacterial Rap-Phr QSSs are known to regulate the competence, the sporulation and/or  
128 the production of public goods in *Bacilli*. Notably, these communication systems ensure that Spo0A-P  
129 only accumulates when the Rap-Phr encoding subpopulation reaches high densities (19). On this  
130 basis, Rap-Phr QSSs have been proposed as a means for a cell to delay a costly commitment to  
131 sporulation as long as the ratio of available food per kin-cell is compatible with individual survival in  
132 periods of nutrient limitation (19). However, the fact that Rap-Phr QSSs are found on phages implies  
133 that this delay in the timing of sporulation can sometimes be dependent on the intracellular density of  
134 phage genomes or prophages rather than on actual cell densities, presumably for the evolutionary  
135 benefit of the phage or the prophage-host collective (20).

136 Here, we inferred the maximum-likelihood phylogeny of the 2258 detected Rap receptors (**Fig2A**)  
137 and identified that the 326 viral Rap-Phr QSSs are genetically diverse and polyphyletic (scattered

138 between bacterial leaves). Remarkably, this suggests that bacteria and phages frequently exchange  
139 these communication systems.

140 Indeed, from a phage's perspective, a dynamical modulation of sporulation can be advantageous  
141 because on the one hand, sporulation can elicit the lytic cycle (21) and trigger cannibalistic behaviors  
142 that may reduce the amount of potential hosts (22) and on the other hand, spores can protect the  
143 phage genome under unfavorable environmental conditions. From a lysogenized host perspective, the  
144 *rap-phr* QSS acquired from the prophage can be considered as adaptive because they might support  
145 a cell-cell communication between lysogenized bacteria (since prophage density correlates with host  
146 density) that enables to cheat (delay production of public goods) and delay the costly sporulation  
147 program in a density-dependent manner, to the benefit of the prophage-host collective (23). A delay of  
148 sporulation might notably offer some additional time to resume growth if a peculiar stress happens to  
149 be relieved from the environment (24) or to benefit from the last bite of food present in the  
150 environment to replicate before sporulating, hence maximizing the number of representatives in the  
151 future heterogeneous population of spores.

152

### 153 **Multiple occurrences of the quorum sensing-mediated sporulation-hijacking genomic** 154 **signature in temperate bacteriophages**

155 Remarkably, the phage-encoded Rap-Phr QSSs were not the only potential host-hijacking viral  
156 QSSs. Indeed, we identified 5 additional candidate phage/prophage-encoded QSSs, distributed into 5  
157 different QSSs families, predicted to manipulate the host sporulation initiation pathway in a density-  
158 dependent manner (**Fig3**). This prediction lies on the observation that their receptor harbors a DNA  
159 binding domain (**TableS1**) and thus likely regulates the expression of adjacent genes (like the  
160 arbitrium system), and that either the *spo0E* or *arabB* sporulation regulator is found in the vicinity of the  
161 QSSs on the phage/prophage DNA (**Fig3**). The same genomic context, albeit not encoded by a  
162 phage, was shown to be linked to sporulation regulation in *Paenibacillus* bacteria (16). By either  
163 activating or repressing the viral *spo0E* or *abrB* genes once a quorum of phages/prophages is  
164 reached, our candidate viral QSSs could influence the total concentration of Spo0E or AbrB within  
165 hosts in a density-dependent-manner, thus influencing the dynamics of Spo0A-P accumulation in  
166 these hosts and thereby modulating the sporulation initiation pathway to the benefit of the phage or  
167 the prophage (**FigS2**). Importantly, as the sporulation initiation pathway can trigger a wide range of  
168 biological processes (sporulation, biofilm formation, cannibalism, toxin production and solventogenesis  
169 (22, 25)), these viral candidate QSSs would also manipulate, in a density-dependent manner, a  
170 substantially broader spectrum of the host biology than spore formation alone (**FigS2**). Because these  
171 viral QSSs belong to distinct QSS families and are carried by phages infecting different hosts, this  
172 suggests a convergent evolution of density-dependent sporulation-hijacking in bacteriophages.

173 We next found published data that support our prediction of a dynamical sporulation-hijacking  
174 mediated by phage-encoded QSSs. Indeed, the RapBL5-PhrBL5 of *Bacillus licheniformis* (26), the  
175 Qss5R-Qss5P of *Clostridium saccharoperbutylacetonicum* (17), and the QsrG-QsrG of *Clostridium*  
176 *acetobutylicum* (18) were previously found to dynamically regulate sporulation and we identified that  
177 these QSSs are predicted by Prophage Hunter to be encoded by prophages inserted within bacterial  
178 chromosomes (**Fig4**). Moreover, 2 of these 3 prophages are predicted by Phaster to be intact and by  
179 Prophage Hunter to be active and thus able to re-initiate the lytic cycle upon excision while the other  
180 probably was domesticated by its host (27). These observations are the evidence that some genomes  
181 of temperate bacteriophages encode a communication system that guides the manipulation of  
182 bacterial sporulation.

183

#### 184 **Concluding remarks**

185 Here, we computationally characterized distinct candidate QSSs presenting a genomic signature  
186 for a density-dependent hijacking of the bacterial sporulation initiation pathway in sequenced genomes  
187 of temperate bacteriophages (*Bacillus* phage phi3T, *Brevibacillus* phage Sundance) and in latent  
188 prophages. Moreover, we found published data that supports the validity of this signature. Because  
189 these QSSs are genetically diverse and belong to phages/prophages infecting different species of  
190 bacteria, this suggests multiple independent acquisitions of quorum sensing-mediated sporulation-  
191 hijacking genetic systems in phages. Accordingly, phage-encoded QSSs would not only dynamically  
192 regulate viral processes (as in the arbitrium QSSs) but also bacterial processes, for the benefit of the  
193 phage or the prophage-host collective. Our study also highlights that if bacteria decide when it's time  
194 to sporulate, some phages also are stakeholders in the decision-making process. Some  
195 bacteriophages had already been reported to be either spore-restricting or spore-converting (28–31),  
196 but either way, these activations/impairments of sporulation were not the result of a decision making-  
197 process, unlike the density-dependent modulations described here.

198 From an applied viewpoint, these findings are also important because as sporulation enables  
199 bacteria to resist harsh environmental conditions, it represents a route for bacteria to travel between  
200 environments, and notably to end up within human bodies. Consequently, endospores are the  
201 transmissive form of many bacteria, be they commensal or pathogen (32–34). Furthermore, 50-60% of  
202 the bacterial genera from the human gut are estimated to produce resilient spores, specialized for  
203 host-to-host transmission (35). Accordingly, by dynamically interfering with sporulation, phage-  
204 encoded QSSs could influence the dynamics of transmission of bacteria in humans. With this respect,  
205 we also report a high prevalence of genes in the human gut phage database matching the HMM  
206 models of Spo0E and AbrB (**TableS2**), strengthening the case of a regulative effect of bacteriophages  
207 on the host-to-host transmission of gastrointestinal bacteria.

208

## 209 **METHODS**

### 210 **Construction of the target data sets**

211 The complete genomes of Viruses and Firmicutes were queried from the NCBI 'Assembly'  
212 database (36), as of 28/04/2020 and 10/04/2020, respectively. The features tables (annotations) and  
213 the encoded protein sequences of these genomes were downloaded using 'GenBank' as source  
214 database. The Gut Phage Database (12) was downloaded as of 29/10/2020, from the following url:  
215 [http://ftp.ebi.ac.uk/pub/databases/metagenomics/genome\\_sets/gut\\_phage\\_database/](http://ftp.ebi.ac.uk/pub/databases/metagenomics/genome_sets/gut_phage_database/)

216

### 217 **Detection of candidate RRNPP QSSs**

218 We launched a systematical search of the RRNPP-type signature using RRNPP\_detector  
219 (3) against the complete genomes of Viruses and Firmicutes available on the NCBI and the MAGs of  
220 bacteriophages from the Gut Phage Database. RRNPP\_detector defines candidate RRNPP-type  
221 quorum sensing systems as tandems of adjacent ORFs encoding a candidate receptor (protein  
222 matching HMMs of peptide-binding tetratricopeptide repeats (TPRs)) and a candidate pro-peptide (10-  
223 small protein predicted to be excreted via the SEC-translocon), consistent with the genetic features  
224 that are specific to the RRNPP quorum sensing mechanism and that are common between different  
225 experimentally-validated RRNPP QSSs (3). We specified a length range of 15-65aa for candidate  
226 propeptides and 250-460aa for candidate receptors.

227

### 228 **Classification of the candidate RRNPP QSSs into families**

229 Because quorum sensing pro-peptides offer few amino acids to compare, are versatile and  
230 subjected to intragenic duplication (26), we classified the QSSs based on sequence homology of the  
231 receptors as in (3). We launched a BLASTp (37) All vs All of the receptors of the 2697 candidate QSSs  
232 identified in the complete genomes of Viruses and Firmicutes. The output of BLASTp was filtered to  
233 retain only the pairs of receptors giving rise to at least 30% sequence identity over more than 80% of  
234 the length of the two proteins. These pairs were used to build a sequence similarity network and the  
235 families were defined based on the connected components of the graph.

236

### 237 **Identification of already known QSS families**

238 A BLASTp search was launched using as queries the RapA (NP\_389125.1), NprR  
239 (WP\_001187960.1), PlcR (WP\_000542912.1), Rgg2 (WP\_002990747.1), AimR (APD21232.1), AimR-  
240 like (AID50226.1), PrgX (WP\_002366018.1), TraA (BAA11197.1), AloR13 (IMG: 2547357582), QsrB  
241 (AAK78305.1), Qss5R (AGF59421.1) reference receptors, and as a target database, the 2697  
242 candidate QSS receptors previously identified. If the best hit of a reference RRNPP-type receptor

243 gave rise to a sequence identity  $\geq 30\%$  over more than 80% mutual coverage, then the family to  
244 which this best hit belongs was considered as an already known family.

245

### 246 **Prophage detection**

247 All the NCBI ids of the genomic accessions of chromosomes or plasmids of Firmicutes encoding  
248 one or several candidate QSSs were retrieved and automatically submitted to the Phaster webtool  
249 (13). Eventually, each QSS was defined as viral if its genomic coordinates on a given  
250 chromosome/plasmid fell within a region predicted by Phaster to belong to a prophage (qualified as  
251 either 'intact', 'questionable' or 'incomplete' prophage). Phaster results were complemented by  
252 ProphageHunter (14), a webtool that computes the likelihood that a prophage is active (able to  
253 reinitiate the lytic cycle by excision). Because ProphageHunter cannot be automatically queried, we  
254 only called upon this webtool for chromosomes/plasmids which encode QSSs that are not part of the  
255 two biggest of the 64 detected QSS families, namely Rap-Phr (2258 candidate QSSs) and PlcR-PapR  
256 (223 candidate QSSs, data not shown). Likewise, coordinates of candidate QSSs were eventually  
257 intersected with predicted prophage regions to detect potential prophage-encoded candidate QSSs  
258 that could have been missed by Phaster (**TableS1**). Based on the results of Phaster and Prophages,  
259 the focus of the study were further restricted to the 26 families comprising at least one phage or  
260 prophage-encoded QSSs.

261

### 262 **Prediction of the mature quorum sensing peptides**

263 For each uncharacterized families of candidate receptors of size  $>1$  with at least one (pro)phage-  
264 encoded member referenced in the NCBI, the cognate pro-peptides were aligned in a multiple  
265 sequence alignment (MSA) using MUSCLE *version 3.8.31* (38). Each MSA was visualized with  
266 Jalview *version 1.8.0\_201* under the ClustalX color scheme which colors amino acids based on  
267 residue type conservation (39). The region of RRNPP-type pro-peptides encoding the mature quorum  
268 sensing peptide usually corresponds to a small sequence (5-6aa), located in the C-terminal of the pro-  
269 peptide, with conserved amino-acids types in at least 3 positions (1, 18, 20, 40). Based on the amino-  
270 acid profile of C-terminal residues in each MSA, putative mature quorum sensing peptides were  
271 manually determined (**FigS3**).

272

### 273 **Phylogenetic tree of Rap**

274 A multiple sequence alignment (MSA) of the protein sequences of the Rap receptors forming a  
275 candidate Rap-Phr QSS was performed using MUSCLE *version 3.8.31* (38). The MSA was then  
276 trimmed using trimmal *version 1.4.rev22* with the option '-automated1', optimized for maximum  
277 likelihood phylogenetic tree reconstruction (41). The trimmed MSA was then given as input to IQ-  
278 TREE *version multicore 1.6.10* to infer a maximum likelihood phylogenetic tree under the LG+G model

279 with 1000 ultrafast bootstraps (42). The tree was further edited via the Interactive Tree Of Life (ITOL)  
280 online tool (43).

281

## 282 **Analysis of the genomic context of candidate QSSs**

283 The genomic context of the (pro)phage-encoded candidate QSSs that are not part of the arbitrium  
284 families (functions already known in phages) and the Rap-Phr family (Rap is a protein inhibitor and not  
285 a transcription factor) and whose receptors matched the DNA binding domain profiles of RRNPP  
286 transcription factors (Hidden Markov Models PFAM PF01381, Superfamily SSF47413, SMART  
287 SM00530, CATH 1.10.260.40) were investigated by analyzing the functional annotation of their  
288 adjacent protein-coding genes, or when missing, by launching a “Conserved Domains” search within  
289 their sequence and a BLASTp search of their sequence against the NR (non-redundant) protein  
290 database maintained by the NCBI.

291

## 292 **Identification of *rap*, *spo0E* and *abrB* genes in the Gut Phage Database**

293 With HMMER, we launched an HMM search of reference HMMs of Rap (PFAM PF18801), Spo0E  
294 (PFAM PF09388) and AbrB (SMART SM00966) against all the protein sequences predicted from the  
295 ORFs of the MAGs from the Gut Phage Database. The hits were retained only if they gave rise to an  
296 E-value < 1E-5.

297

## 298 **DECLARATIONS**

### 299 **Availability of data and materials**

300 All the NCBI or Gut Phage Database IDs of the proteins discussed in this manuscript are available in  
301 the supplementary tables.

### 302 **Competing Interests**

303 The authors of this manuscript have no competing interests to disclose.

### 304 **Fundings**

305 This research did not receive any specific grant from funding agencies in the public, commercial, or  
306 not-for-profit sectors. C. Bernard was supported by a PhD grant from the Ministère de l'Enseignement  
307 supérieur, de la Recherche et de l'Innovation.

### 308 **Authors' Contribution**

309 C.B, Y.L, E.B and P.L conceived the study. C.B performed the analyses. C.B, Y.L and E.B wrote the  
310 manuscript with input from all authors. All documents were edited and approved by all authors.

### 311 **Acknowledgments**

312 We would like to thank Dr. A. K. Watson for critical reading and discussion.

313

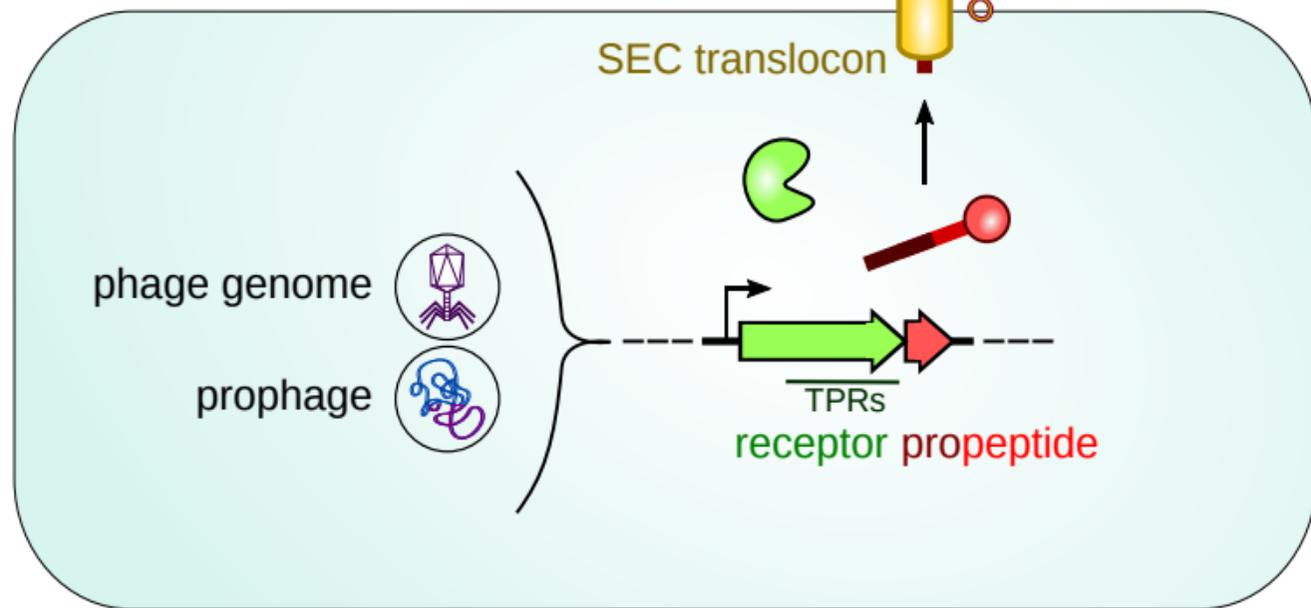
## 314 **REFERENCES**

- 315 1. Z. Erez, *et al.*, Communication between viruses guides lysis-lysogeny decisions. *Nature* **541**,  
316 488–493 (2017).
- 317 2. M. B. Neiditch, G. C. Capodagli, G. Prehna, M. J. Federle, Genetic and Structural Analyses of  
318 RRNPP Intercellular Peptide Signaling of Gram-Positive Bacteria. *Annu. Rev. Genet.* **51**, 311–  
319 333 (2017).
- 320 3. C. Bernard, Y. Li, E. Bapteste, P. Lopez, RRNPP\_detector: a tool to detect RRNPP quorum  
321 sensing systems in chromosomes, plasmids and phages of gram-positive bacteria AUTHORS.  
322 *bioRxiv*, 2021.08.18.456871 (2021).
- 323 4. V. Kohler, W. Keller, E. Grohmann, Regulation of gram-positive conjugation. *Front. Microbiol.*  
324 **10**, 1134 (2019).
- 325 5. A. Stokar-Avihail, N. Tal, Z. Erez, A. Lopatina, R. Sorek, Widespread Utilization of Peptide  
326 Communication in Phages Infecting Soil and Pathogenic Bacteria. *Cell Host Microbe* **25**, 746-  
327 755.e5 (2019).
- 328 6. M. Y. Galperin, Genome Diversity of Spore-Forming Firmicutes. *Microbiol. Spectr.* **1**, TBS-0015-  
329 2012 (2013).
- 330 7. I. S. Tan, K. S. Ramamurthi, Spore formation in *Bacillus subtilis*. *Environ. Microbiol. Rep.* **6**,  
331 212–225 (2014).
- 332 8. M. A. Al-Hinai, S. W. Jones, E. T. Papoutsakis, The *Clostridium* Sporulation Programs: Diversity  
333 and Preservation of Endospore Differentiation. *Microbiol. Mol. Biol. Rev.* **79**, 19–37 (2015).
- 334 9. M. Fujita, R. Losick, Evidence that entry into sporulation in *Bacillus subtilis* is governed by a  
335 gradual increase in the level and activity of the master regulator Spo0A. *Genes Dev.* **19**, 2236–  
336 2244 (2005).
- 337 10. S. H. Shafikhani, T. Leighton, AbrB and Spo0E Control the Proper Timing of Sporulation in  
338 *Bacillus subtilis*. *Curr. Microbiol.* **48**, 262–269 (2004).
- 339 11. D. Schultz, M. Lu, T. Stavropoulos, J. Onuchic, E. Ben-Jacob, Turning oscillations into  
340 opportunities: Lessons from a bacterial decision gate. *Sci. Rep.* **3** (2013).
- 341 12. L. F. Camarillo-Guerrero, A. Almeida, G. Rangel-Pineros, R. D. Finn, T. D. Lawley, Massive  
342 expansion of human gut bacteriophage diversity. *Cell* **184**, 1098-1109.e9 (2021).
- 343 13. D. Arndt, *et al.*, PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic*  
344 *Acids Res.* **44**, W16 (2016).
- 345 14. W. Song, *et al.*, Prophage Hunter: an integrative hunting tool for active prophages. *Nucleic*  
346 *Acids Res.* **47**, W74–W80 (2019).
- 347 15. S. Perchat, *et al.*, NprR, a moonlighting quorum sensor shifting from a phosphatase activity to a  
348 transcriptional activator. *Microb. Cell* **3**, 573–575 (2016).
- 349 16. M. Voichek, S. Maaß, T. Kroniger, D. Becher, R. Sorek, Peptide-based quorum sensing systems  
350 in *Paenibacillus polymyxa*. *Life Sci. Alliance* **3** (2020).

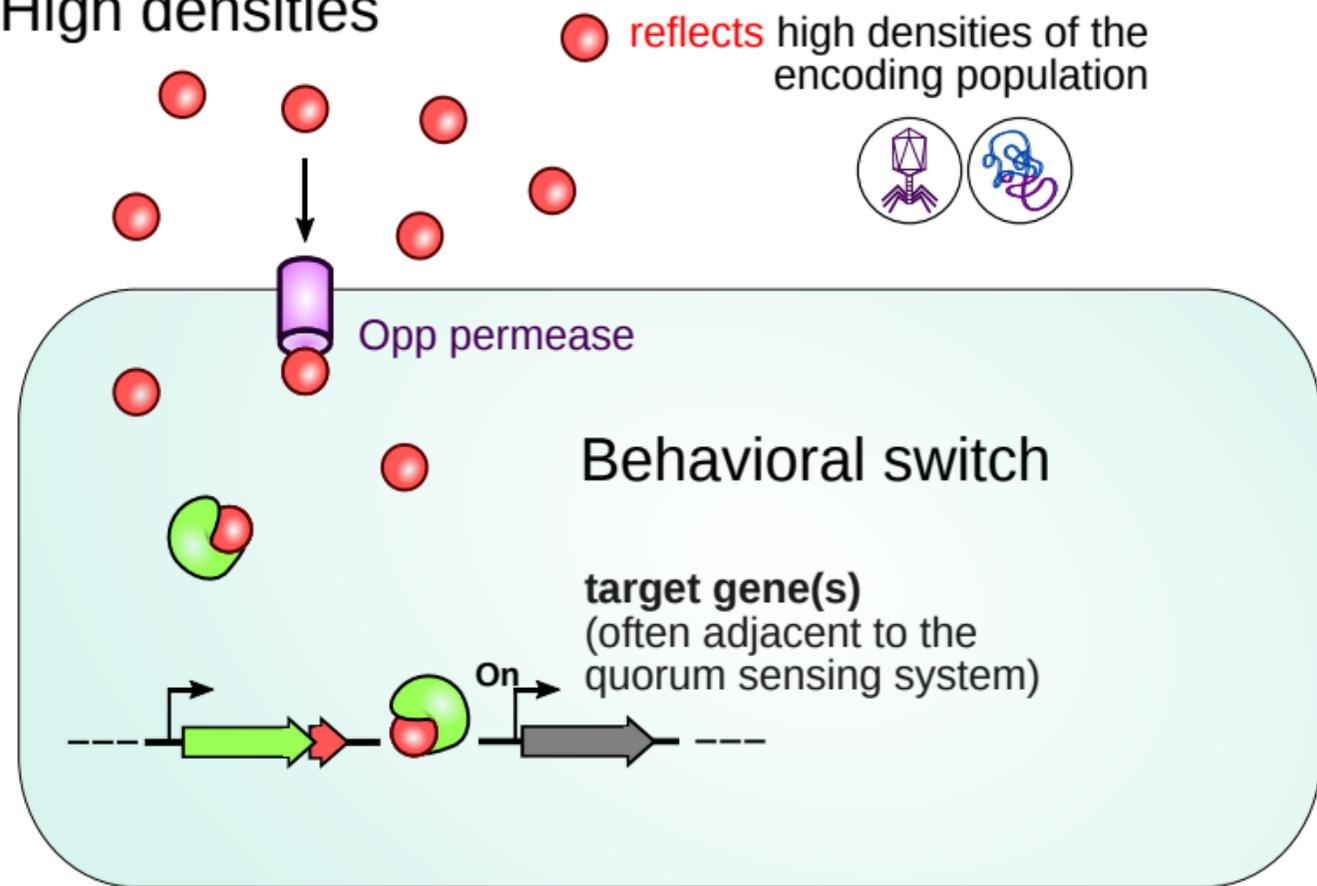
- 351 17. J. Feng, *et al.*, RRNPP-Type quorum-sensing systems regulate solvent formation, sporulation  
352 and cell motility in *Clostridium saccharoperbutylacetonicum*. *Biotechnol. Biofuels* **13**, 1–16  
353 (2020).
- 354 18. A. K. Kotte, *et al.*, RRNPP-type quorum sensing affects solvent formation and sporulation in  
355 *clostridium acetobutylicum*. *Microbiol. (United Kingdom)* **166**, 579–592 (2020).
- 356 19. I. B. Bischofs, J. A. Hug, A. W. Liu, D. M. Wolf, A. P. Arkin, Complexity in bacterial cell-cell  
357 communication: quorum signal integration and subpopulation signaling in the *Bacillus subtilis*  
358 phosphorelay. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 6459–64 (2009).
- 359 20. C. Bernard, Y. Li, P. Lopez, E. Bapteste, Beyond arbitrium: identification of a second  
360 communication system in *Bacillus* phage phi3T that may regulate host defense mechanisms.  
361 *ISME J.* (2020) <https://doi.org/10.1038/s41396-020-00795-9>.
- 362 21. W. J. Meijer, *et al.*, Molecular basis for the exploitation of spore formation as survival  
363 mechanism by virulent phage  $\phi$ 29. *EMBO J.* **24**, 3647–3657 (2005).
- 364 22. J. E. González-Pastor, Cannibalism: A social behavior in sporulating *Bacillus subtilis*. *FEMS*  
365 *Microbiol. Rev.* **35**, 415–424 (2011).
- 366 23. M. Kalamara, M. Spacapan, I. Mandic-Mulec, N. R. Stanley-Wall, Social behaviours by *Bacillus*  
367 *subtilis*: quorum sensing, kin discrimination and beyond. *Mol. Microbiol.* **110**, 863 (2018).
- 368 24. J. E. González-Pastor, E. C. Hobbs, R. Losick, Cannibalism by Sporulating Bacteria. *Science*  
369 (80-. ). **301**, 510–513 (2003).
- 370 25. P. Dürre, *et al.*, Transcriptional regulation of solventogenesis in *Clostridium acetobutylicum* in  
371 *Journal of Molecular Microbiology and Biotechnology*, (2002), pp. 295–300.
- 372 26. E. Even-Tov, S. Omer Bendori, S. Pollak, A. Eldar, Transient Duplication-Dependent Divergence  
373 and Horizontal Transfer Underlie the Evolutionary Dynamics of Bacterial Cell–Cell Signaling.  
374 *PLOS Biol.* **14**, e2000330 (2016).
- 375 27. L. M. Bobay, M. Touchon, E. P. C. Rocha, Pervasive domestication of defective prophages by  
376 bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 12127–12132 (2014).
- 377 28. D. P. Boudreaux, V. R. Srinivasan, “Bacteriophage-induced Sporulation in *Bacillus cereus* T.”
- 378 29. M. G. Bramucci, K. M. Keggin, P. S. Lovett, Bacteriophage conversion of spore-negative  
379 mutants to spore-positive in *Bacillus pumilus*. *J. Virol.* **22**, 194–202 (1977).
- 380 30. T. H. Silver-Mysliwiec, M. G. Bramucci, Bacteriophage-enhanced sporulation: Comparison of  
381 spore-converting bacteriophages PMB12 and SP10. *J. Bacteriol.* **172**, 1948–1953 (1990).
- 382 31. R. Schuch, V. A. Fischetti, The secret life of the anthrax agent *Bacillus anthracis*:  
383 bacteriophage-mediated ecological adaptations. *PLoS One* **4**, e6532 (2009).
- 384 32. M. Mallozzi, V. K. Viswanathan, G. Vedantam, Spore-forming Bacilli and Clostridia in human  
385 disease. *Future Microbiol.* **5**, 1109–1123 (2010).
- 386 33. F. Postollec, *et al.*, Tracking spore-forming bacteria in food: From natural biodiversity to  
387 selection by processes. *Int. J. Food Microbiol.* **158**, 1–8 (2012).

- 388 34. M. C. Swick, T. M. Koehler, A. Driks, Surviving Between Hosts: Sporulation and Transmission.  
389 *Microbiol. Spectr.* **4** (2016).
- 390 35. H. P. Browne, *et al.*, Culturing of “unculturable” human microbiota reveals novel taxa and  
391 extensive sporulation. *Nature* **533**, 543–546 (2016).
- 392 36. NCBI Resource Coordinators, Database resources of the National Center for Biotechnology  
393 Information. *Nucleic Acids Res.* **44**, D7–D19 (2016).
- 394 37. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool.  
395 *J. Mol. Biol.* **215**, 403–410 (1990).
- 396 38. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput.  
397 *Nucleic Acids Res.* **32**, 1792–7 (2004).
- 398 39. A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, G. J. Barton, Jalview Version 2--a  
399 multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191  
400 (2009).
- 401 40. M. Pottathil, B. A. Lazazzera, The extracellular PHR peptide-Rap phosphatase signaling circuit  
402 of bacillus subtilis. *Front. Biosci.* **8**, 913 (2003).
- 403 41. S. Capella-Gutierrez, J. M. Silla-Martinez, T. Gabaldon, trimAl: a tool for automated alignment  
404 trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
- 405 42. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A Fast and Effective  
406 Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **32**, 268–  
407 274 (2015).
- 408 43. I. Letunic, P. Bork, Interactive Tree of Life (iTOL) v4: Recent updates and new developments.  
409 *Nucleic Acids Res.* **47**, W256–W259 (2019).
- 410
- 411

Low intracellular densities of free phage genomes or inserted prophages



High densities



412 **Figure1**

413 **Mechanism of RRNPP QSSs in phages/prophages.** The receptor and the communication  
414 propeptide of the phage-encoded QSS are in green and red, respectively. Upon bacterial expression,  
415 secretion and maturation, the concentration of the communication peptide reflects the intracellular  
416 density of phage genomes or prophages, which correlates with the number of lysed cells during the  
417 lytic cycle and the number of lysogenized cells during the lysogenic cycle. At high concentrations of  
418 the peptide, reflecting a quorum of phages/prophages, the receptor binds the imported mature peptide  
419 via TPRs motifs and gets either turned-on or -off as a transcription factor or a protein inhibitor, which is  
420 at the basis of density-dependent regulations of target genes (often adjacent to the QSS) or proteins.  
421 These regulations thereupon coordinates a behavioral transition at the scale of the entire  
422 phage/prophage population.

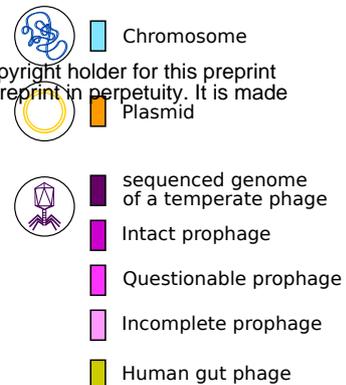
**A**

**Rap/NprR  
Bacillus**

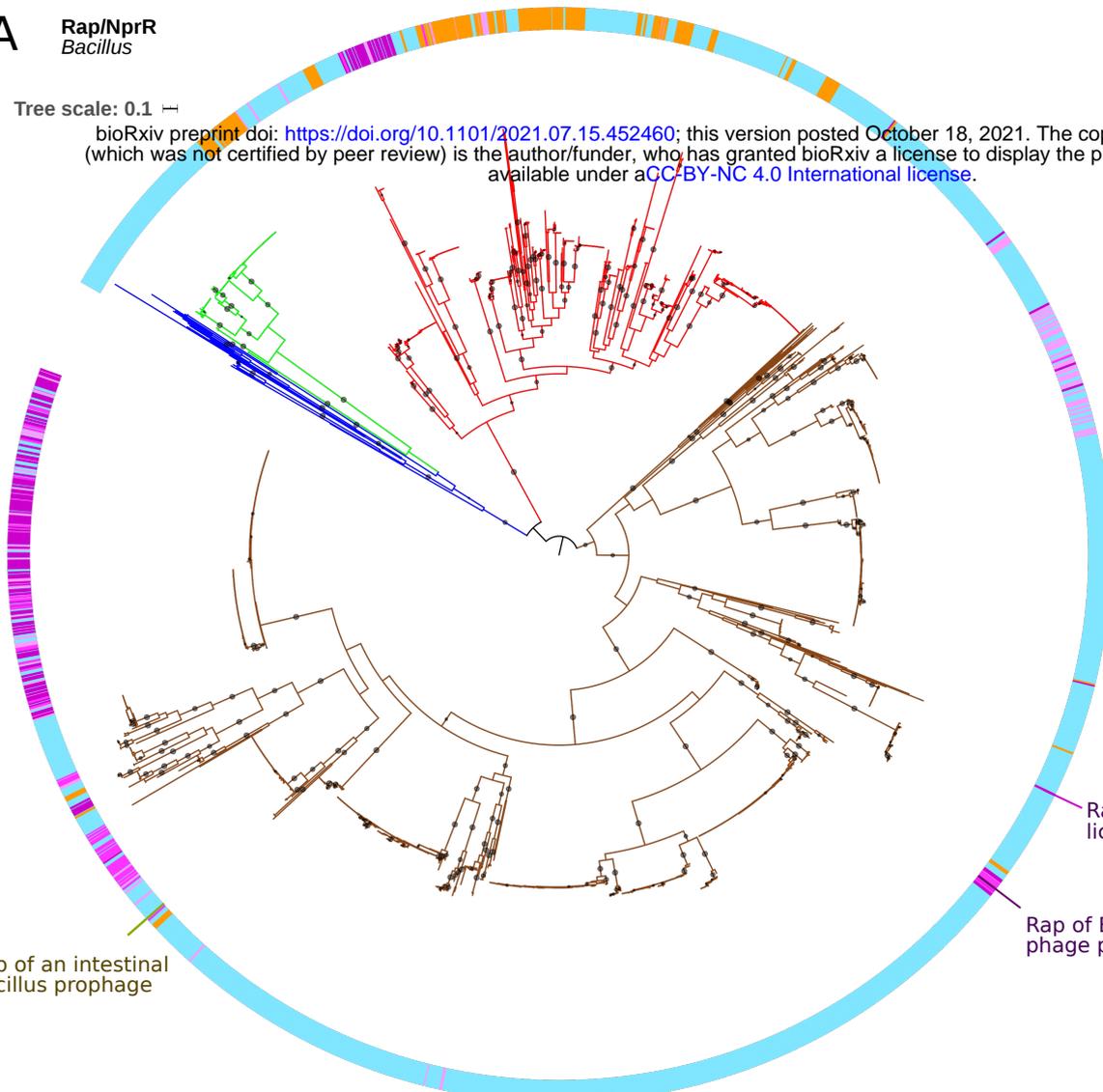
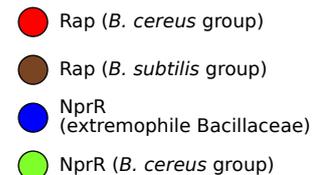
Tree scale: 0.1

bioRxiv preprint doi: <https://doi.org/10.1101/2021.07.15.452460>; this version posted October 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

**Strip color**

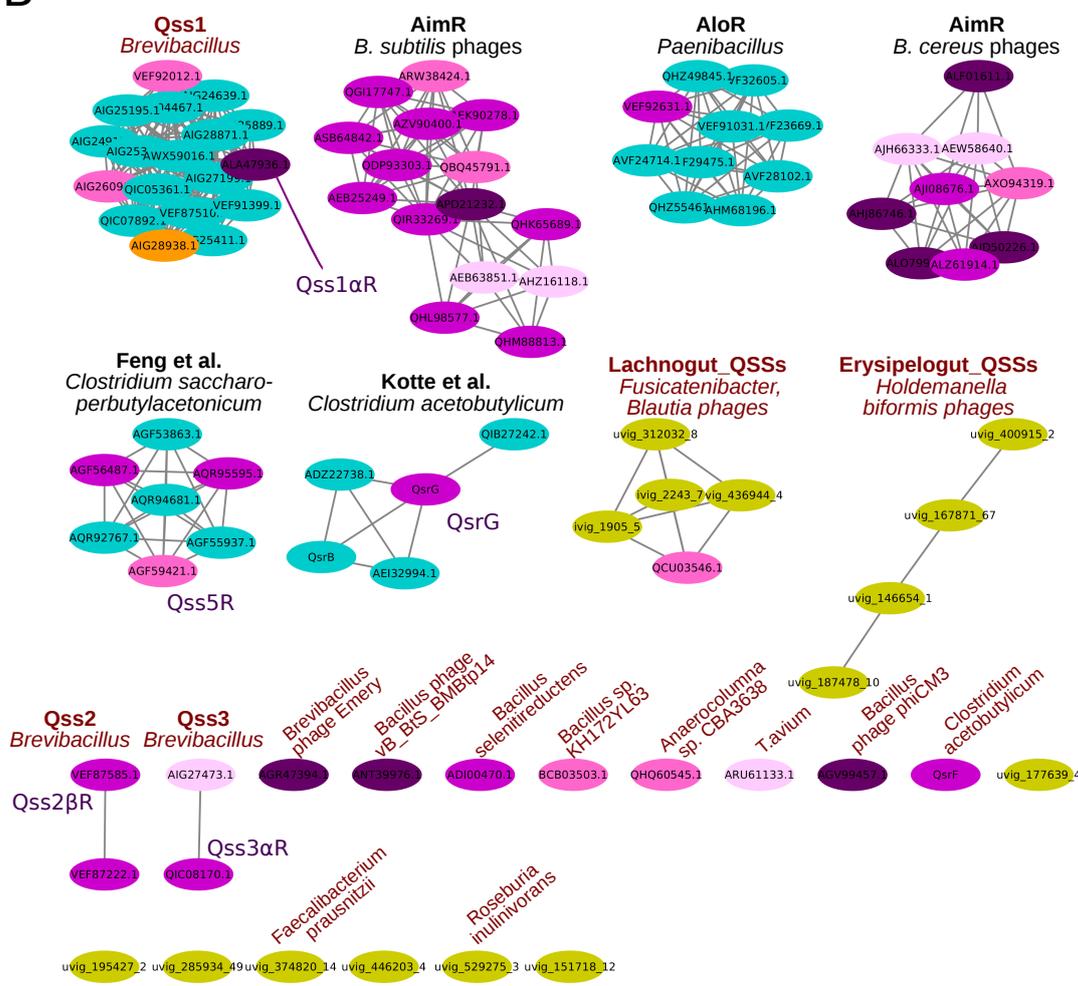
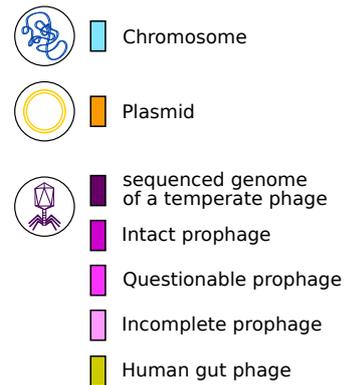


**Branch color**



**B**

**Node color**



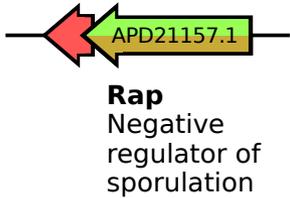
423 **Figure 2**

424 **QSS families comprising at least one phage/prophage-encoded candidate QSS**

425 **A. Maximum-likelihood phylogeny of receptors of the Rap/NprR family.** The figure displays the  
426 maximum-likelihood phylogenetic tree of the family comprising the Rap (no DNA binding domain) and  
427 the NprR (DNA binding domain) receptors that are part of a detected RRNPP-type QSS. The  
428 clustering of Rap and NprR into the same protein family is consistent with the common phylogenetic  
429 origin proposed for these receptors (S. Perchat, *et al.*, *Microb. Cell* 2016)). The tree was midpoint  
430 rooted and a small black circle at the middle of a branch indicates that the branch is supported by 90%  
431 of the 1000 ultrafast bootstraps performed. Branch colors are indicative of the type of receptor (Rap or  
432 NprR) and of the bacterial group that either directly encodes the QSS or can be lysogenized by a  
433 (pro)phage that encodes the QSS. The colorstrip surrounding the phylogenetic tree assigns a color to  
434 each leaf based on the type of genetic support that encodes the QSS: blue for chromosomes, orange  
435 for plasmids, dark purple for sequenced genomes of temperate bacteriophages, different levels of  
436 purple for Phaster-predicted intact, questionable and incomplete prophages. The Rap receptors of  
437 *Bacillus* phage phi3T (only Rap found in a sequenced genome of a temperate phage) and of *B.*  
438 *licheniformis* intact prophage (viral Rap inserted into the host chromosome and shown to modulate the  
439 bacterial sporulation and competence pathways) are outlined. **B. Sequence similarity network of the**  
440 **other candidate QSS families comprising at least one phage/prophage-encoded QSS.** Each  
441 node corresponds to a receptor sequence found adjacent to a candidate pro-peptide and is colored  
442 according to the type of genetic element encoding the QSS, as displayed in the legend. The label of  
443 each node indicates the NCBI id or a Gut Phage Database id of the candidate receptor. Each edge  
444 corresponds to a similarity link between two receptors defined according to the following thresholds:  
445 percentage identity  $\geq 30\%$ , alignment coverage  $\geq 80\%$  of the lengths of both receptors, E-value  $\leq$   
446  $1E-5$ . Each connected component of the graph thereby defines groups of homologous receptors and  
447 is considered as a QSS family. The families are ordered from the largest to the smallest. Families with  
448 a black label were already described before (but not necessarily in phages) while families with a red  
449 label are novel. The most prevalent encoding-taxon in the family is displayed on top of it. The nodes of  
450 the receptors that are part of a predicted sporulation-hijacking QSS are characterized by an additional  
451 label.

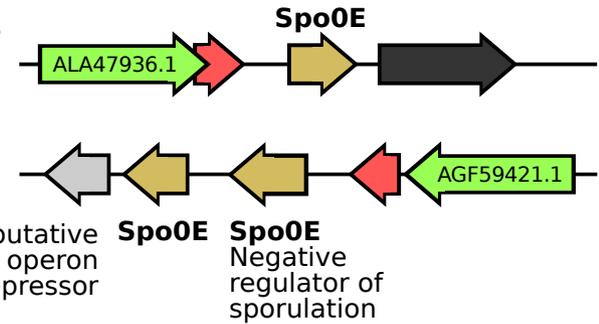
# NCBI Complete Genomes

 Bacillus phage phi3T  
**Rap-Phr**

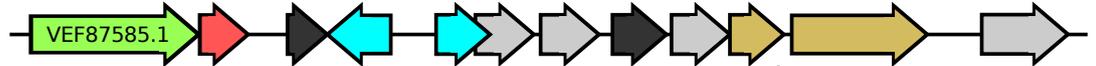


 Brevibacillus phage Sundance  
**Qss1α**

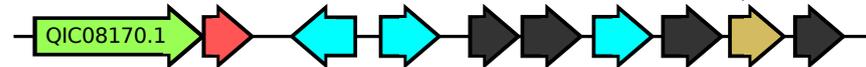
 Ambiguous prophage of *Clostridium saccharoperbutylacetonicum*  
**Feng et al.**  
**QssR5-QssP5**



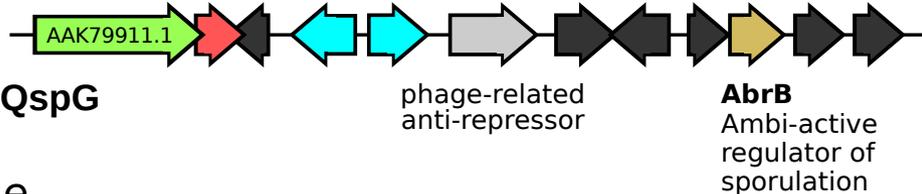
 Active prophage of *Brevibacillus brevis*  
**Qss2β**



 Active prophage of *Brevibacillus 7WMA2*  
**Qss3α**



 Active prophage of *Clostridium acetobutylicum*  
**Kotte et al.** **QsrG-QspG**



# Gut Phage Database

 Prophage of *Bacillus subtilis*  
**Rap-Phr**



452 **Figure 3.**

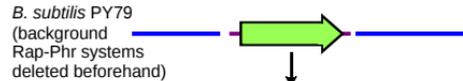
453 **Clustering of sporulation initiation regulators with QSSs found in sequenced genomes of**  
454 **phages** (virion icon) **or prophages** (lysogenized chromosome icon). The name and the genome of  
455 each QSS is written aside from each genomic context. Genes are colored according to their functional  
456 roles, as displayed in the legend. Within the QSS receptor gene, the NCBI id of the protein is shown.  
457 Because RRNPP QSSs tend to regulate adjacent genes, these genomic contexts hint at density-  
458 dependent sporulation-hijacking mechanisms.

## RapBL5-PhrBL5

Phaster: intact prophage (2,854,951-2,867,485)  
 ProphageHunter: active prophage (2,855,587-2,866,209)

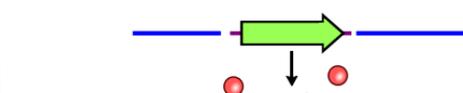


### rapBL5 cloned without phrBL5



Decreased sporulation

### Addition of PhrBL5<sub>mature</sub>

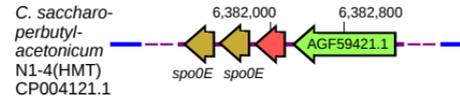


Sporulation Competence

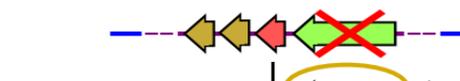
PhrBL5 --| RapBL5 --| Spo0F-P --> Spo0A-P

## Qss5R-Qss5P

Phaster: incomplete prophage (6,390,448-6,416,714)  
 ProphageHunter: ambiguous prophage (6,375,466-6,398,632)

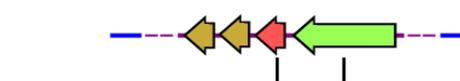


### Overexpression of qss5R



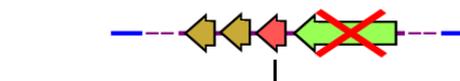
Increased sporulation plasmid

### Basal expression of the qss



Sporulation

### Mutation of qss5R

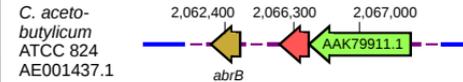


Decreased sporulation

Hypothesis: Qss5P --| Qss5R --| Spo0E --| Spo0A-P

## QsrG-QspG

Phaster: intact prophage (2,017,737-2,083,827)  
 ProphageHunter: active prophage (2,050,451-2,078,632)



### Basal expression of the qss



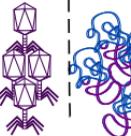
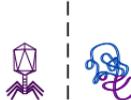
Sporulation  
 $1.15 \times 10^8$  heat resistant  
 CFU/mL (endospores)  
 after 7 days culture

### Mutation of qsrG



Decreased sporulation  
 only  $3.85 \times 10^7$  heat  
 resistant CFU/mL  
 after 7 days culture

phage or  
 prophage  
 density



Lytic  
 cycle

Lysogenic  
 cycle

459 **Figure 4.**

460 **Summary of the evidences that some viral QSSs dynamically influence bacterial sporulation.**

461 The functions of the three QSSs (one column each) were investigated in (17, 18, 26) as displayed in  
462 this figure. The prediction that each QSS is encoded by a prophage is shown in purple at the top of  
463 each column. A prophage qualified as active by Prophage Hunter means that it is predicted to be able  
464 to reinitiate the lytic cycle upon excision from the bacterial chromosome. The interaction between the  
465 quorum sensing peptide was only investigated for RapBL5-PhrBL5 and is still lacking for Qss5R-  
466 Qss5P and QsrG-QspG. However, there is evidence in all QSSs that the receptor regulates the  
467 sporulation pathway, via Spo0F-P inhibition for the Rap protein inhibitor, and likely via transcriptional  
468 regulation of the adjacent *spo0E* and *abrB* genes for the Qss5R and QsrG transcription factors.

469