Phage recombination drives evolution of spore-forming *Bacilli*

Anna Dragoš[1]*, Priyadarshini B.[1], Zahraa Hasan[1], Mikael Lenz-Strube[2], Paul J Kempen[3], Gergely Maróti[4], Charlotte Kaspar[5,6], Baundauna Bose[7], Briana M. Burton[8], Ilka B Bischofs[5,6], Ákos T. Kovács[1]*

[1] Bacterial Interactions and Evolution Group, Department for Biotechnology and Biomedicine, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

[2] Bacterial Ecophysiology and Biotechnology Group, Department for Biotechnology and Biomedicine, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

[3] Department of Health Technology, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

[4] Institute of Plant Biology, Biological Research Centre, Hungarian Academy of Sciences, H-6701 Szeged, Hungary

[5] BioQuant Center of the University of Heidelberg, 69120 Heidelberg, Germany

[6] Max-Planck-Institute for Terrestrial Microbiology, 35043, Marburg, Germany

[7] Gro Biosciences, Cambridge, MA 02139

[8] Department of Bacteriology, University of Wisconsin, WI 53706, Madison, USA

* Corresponding authors: Bacterial Interactions and Evolution Group, Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads Building 221, 2800 Kongens Lyngby, Denmark

Tel: +45 26272642; Fax: +45 45 932809; E-mail: adragos@dtu.dk;

Tel: +45 45 252527; Fax: +45 45 932809; E-mail: atkovacs@dtu.dk;

Abstract

Phages are the main source of within-species bacterial diversity and drivers of horizontal gene transfer. Prophage cargo can determine ecological interactions of a bacterium with other community members, and even its pathogenic potential, but we know little about the mechanisms that drive genetic diversity of these mobile genetic elements (MGEs). Recently, we showed that a sporulation selection regime promotes evolutionary changes within SPβ prophage of *Bacillus subtilis*, leading to direct antagonistic interactions within the population. Interestingly, SPβ belongs to the so-called phage regulatory switches that precisely excise from the chromosome of sporulating mother cells, a phenomenon observed for phages infecting diverse spore-forming species. Herein, we reveal that under a sporulation selection regime, SPβ recombines with low copy number phi3Ts phage DNA present within the *B. subtilis* population. Recombination results in a new prophage occupying a different integration site, as well as the spontaneous release of virulent phage hybrids. Analysis of *Bacillus* sp. strains suggests that SPβ and phi3T belong to a distinct cluster of unusually large phages inserted into sporulation-related genes that are equipped with a spore-related genetic arsenal. Comparison of *Bacillus* sp. genomes at global and local ecological scales indicates that these SPβ-like phages diversify rapidly, especially in the absence of other MGEs constraining their lytic cycle. Our study captures inter-phage recombination under an experimentally imposed selection regime, and reveals the ubiquity of similar phenomena in *Bacillus* sp. genomic data. Therefore, our work is a stepping stone toward empirical studies on phage evolution, and understanding the eco-evolutionary relationships between bacteria and their phages.

Introduction

Bacteriophages are major regulators of bacteria. Each day, approximately half of all bacterial biomass is killed by lytic phages, imposing a constant predator-prey evolutionary arms race[1–3]. Moreover, phages reside in 40–50% of bacterial genomes as prophage elements[4], serving as a main source of intra-species genetic diversity and gene transfer agents[5]. Although prophages can be considered genetic parasites, they can also benefit their host with new metabolic functions, survival strategies or weapons for inter-bacterial warfare[4–6]. Despite the abundance and relevance of prophages in bacterial genomes, we still understand very little about their ecological and evolutionary imprint. This knowledge gap is particularly paramount for non-pathogenic bacterial agents, such as widely applied biocontrol and probiotic bacterial strains of agricultural and medical importance. Notably, genomes of certain beneficial bacterial genera (e.g. *Bacillus* sp.) are extremely rich in prophages[7]. Moreover, prophage cargo divergence in certain *Bacillus* species (e.g. *Bacillus subtilis*) leads to social 'incompatibility', which manifests in strong competitive interactions and physical barriers between bacterial swarms[8,9]. Currently, we do not understand what promotes diversity within the prophage cargo of closely related strains.

Based on striking mosaicism of prophage genomes, we believe that they predominantly evolve through recombination[10–12]. Through exchange of functional groups of genes (modules), phages can rapidly gain or lose functions[11,13]. Recombination between phages or between phages and their hosts can be either homologous[14,15] or non-homologous[16], relying on phage- or host-encoded recombinases[14,17]. It was proposed that gene shuffling regularly occurs between (functional or defective) prophages and phages that co-infect the same host bacterium[14]. It was also shown that prophages of naturally competent bacteria (e.g.

B. subtilis) can recombine with foreign phage DNA via transformation[18]. Finally, phages can randomly or specifically incorporate fragments of host chromosomes via generalised[19,20] or specialised[21,22] transduction, respectively, thereby contributing to the spread of antibiotic resistance[20] or virulence genes[23]. Although evidence from comparative phage genomics indicates frequent recombination into new phage variants (so-called gene shuffling)[13,24,25], this is not reflected in experimental research. Phage recombination has been experimentally studied using limited models, predominantly *Salmonella typhimurium* P22 with *Escherichia coli* lambdoid phages[26,27]. Therefore, despite our knowledge of pronounced genomic mosaicism, empirical research on prophage evolution is relatively limited. Combining such research with broader comparisons of available host genomes may prove key to understanding the ecology and evolution of bacteria and phages, including whether prophages serve as a major source of bacterial within-species diversity, or as regulators.

Interestingly, certain prophages of *Bacilli* undergo genetic rearrangements upon host development, acting as so-called phage regulatory switches (RSs)[28]. RS phages can switch between integrated and extrachromosomal forms to modulate reproduction and survival of their hosts, through processes that differ from classical lysogeny or lysis[28]. To date, most documented RS phages have been detected in Firmicutes as regulators of the sporulation process[29–32], associated with vegetative cells transforming into partially dehydrated dormant spores, and related to resistance to extreme conditions, including starvation for millions of years[33,34]. Certain *B. subtilis* biocontrol strains carry an SPβ prophage or its derivative that integrates into the polysaccharide-related gene *spsM*, and this genetic interruption prevents robust submerged biofilm formation in the host[29,35]. In addition, in SPβ prophage-harbouring strains, immediately prior to sporulation the prophage undergoes precise excision and circularisation, allowing *spsM* reconstitution and expression in the sporulating mother cell. The resulting *spsM*-related polysaccharide eventually becomes part of the spore coat, contributing

to spore dispersability[29]. Besides SPβ, another prophage-like element named *skin* also undergoes excision in the mother cell, allowing reconstitution of *sigK* encoding a late sporulation sigma factor that is necessary for completing the sporulation process[32,36,37]. Similar mother cell-specific excisions have been observed in other *Bacillus* sp.[29] and in *Clostridium* sp.[31,38], but we do not understand what drives such a distinctive relationship between spore-forming hosts and their phages, nor what eco-evolutionary consequences this has. Interestingly, SPβ and *skin* both encompass genes relevant for sporulation, including *sspC* that is crucial for spore DNA protection and repair[39,40], and the *rapE–phrE* signalling system involved in sporulation initiation[41,42]. Furthermore, certain prophages can improve or even restore sporulation in *B. subtilis*[43], suggesting the possibility of a cooperative relationship between certain phages and spore-formers.

We recently demonstrated that under a repeated imposed sporulation selection regime, SPβ prophages of *B. subtilis* undergo major genetic rearrangements, giving rise to new hybrid phages[44]. Normally, the lytic cycle of SPβ prophages is blocked by the ICEBs1 (Integrative and Conjugative Element of *B. subtilis*) conjugative element[45], and new phage variants are released spontaneously, killing or infecting the original host[44]. Therefore, it is important to reveal the genetic changes that lead to prophage awakening, and determine whether similar prophage evolution pathways occur outside the laboratory.

Herein, we investigated the triggering cause of diversification and spontaneous release of SPβ prophages, and sought evidence for similar diversification of SPβ taking place in nature. Using experimental evolution, *de novo* genome sequencing and testing, we showed that barely detectable, low copy number phage DNA residing in certain *B. subtilis* strains can propagate under an appropriate selection regime, and hybridise with indigenous prophages. These new prophage elements modulate host development, most likely through regulatory genes. Bioinformatic comparison of prophage elements within available *Bacillus* sp. genomes

demonstrated that similar recombination may frequently occur in nature between SPβ and related phages. Our work shows how diversification of prophages through recombination can drive early diversification of bacterial populations.

RESULTS

Strains evolved under a sporulation selection regime carry hybrid prophages

We previously showed that several passages of *B. subtilis* through the dormant spore stage leads to genome rearrangements within prophage elements and release of phage particles into the culture medium (Fig. 1A). Some of these phages resemble indigenous SPβ[46], but unlike SPβ prophages they are produced spontaneously and facilitate killing of the original SPβ lysogenic strain[44] (Fig. 1B). To further characterise the genetic changes within the prophage regions of these evolved strains, three isolates (B310mA, B410mB and B410wtB)[44] were subjected to long-read genome sequencing using the PacBio platform (see Materials and Methods). *De novo* sequencing revealed the presence of an exogenous SPβ-like prophage (58% sequence identity), which was nearly identical to *Bacillus subtilis* phage phi3T (KY030782.1; 99.98% sequence identity)[47] in all three strains (Fig. 1C, Suppl. Fig. 1A). We named these extrachromosomal phage elements phi3Ts. The only difference between previously sequenced phi3T and phi3Ts was a 725 bp fragment (labelled 's' for sporulation-derived) within phi3Ts, replacing the 1265 bp fragment of phi3T (nucleotides 101,429–102,694; Fig. 1D). Strikingly, the 's' fragment shares no homology with phi3T or *B. subtilis* 168 chromosomes, but it could be found within SPβ-like prophages of six *B. subtilis* strains isolated in different regions around the world (see Materials and Methods; Fig. 1D). In the evolved strains, the phi3Ts prophage either disrupted the *kamA* gene located ~11 kb from SPβ, or it created a hybrid with SPβ with a ~11 kb fragment deleted between *kamA* and SPβ (Fig. 1C, Suppl. Fig 1A). In addition,

146 sequencing coverage within the described prophage regions was increased several-fold,
147 suggesting augmented replication of hybrid phage DNA (Suppl. Fig. 1B, Suppl. dataset 1).

148

149 **Hybrid lysogens produce virulent hybrid phages**

150 In view of the presence of phi3Ts, SPβ and phage hybrids on the chromosomes of the evolved
151 strains (Fig. 1C, Suppl. Fig 1A), we were curious which phages are spontaneously released into
152 the medium[44]. Therefore, phages released by the evolved strains were purified from single
153 plaques and subjected to genome sequencing (see Materials and Methods). Notably, each
154 evolved strain produced a mix of turbid and clear plaques, but at different relative frequencies
155 (Suppl. Fig. 1C). Turbid plaques are typical for temperate phages (like SPβ or phi3T), while
156 clear plaques are usually formed by phage variants that have lost their ability to enter the
157 lysogenic cycle[48]. Phage sequencing revealed that the spontaneously produced phages were
158 either phi3Ts or phi3Ts-SPβ hybrids (Fig. 1C, Suppl. Fig. 1A, Suppl. Fig. 2). Sequences of all
159 phages obtained from the turbid plaques matched prophage sequences within the evolved
160 strains (Fig. 1C, Suppl. Fig. 1A, Suppl. Fig. 2). In addition, the genome of Hyb1$^{phi3Ts\text{-}SP\beta}$
161 (released by B310mA) was extended by a ~1.2 kb fragment of the host chromosome (*yozE*,
162 *yokU*, and part of the *kamA* gene), indicating specialised transduction, a process that occurs
163 when a phage picks up a fragment of host chromosomal DNA in the immediate vicinity of its
164 attachment site (Fig. 1C, Suppl. Fig. 2). In contrast to the turbid plaque-creating phages, all
165 phages obtained from clear plaques were phi3Ts-SPβ hybrids, which were not present on the
166 chromosomes of their corresponding producers (Fig. 1C, Suppl. Fig. 1A, Suppl. Fig. 2). This
167 suggests that phi3Ts-SPβ recombination not only gave rise to hybrid prophages, but also to a
168 range of virulent phages. In addition to chromosomal DNA, in strains B410mB and B410wtB
169 we identified a variety of extrachromosomal phage DNA (epDNA) fragments ranging from

10.9 to 66 kb (Fig. 2, Suppl. dataset 1). The epDNA was dominated by phi3Ts-SPβ recombinants, in which DNA from the two parental phages was joined at the homologous region (Fig. 2). None of the hybrid epDNA was identical to sequences of hybrid phages released by the corresponding strains (B410mB and B410wtB; Fig. 2, Suppl. Fig. 2). Finally, we also noticed that some epDNA fragments contained parts of the bacterial chromosome adjacent to the phi3Ts integration site, again pointing towards specialised transduction (Fig. 2).

A sporulation selection regime promotes foreign phage invasion

Next, we aimed to identify the source of phi3Ts DNA in the evolved host genomes, and to determine whether this DNA was already present in the ancestor *B. subtilis* 168 stock or acquired as contamination during the evolution experiment. First, we repeated the mapping of raw sequencing reads from the *B. subtilis* 168 ancestral genome onto selected unique phi3T regions lacking homology with SPβ. Indeed, phi3Ts DNA was present in the ancestor strain at a very low but detectable level, rather than as an extrachromosomal form (Fig. 3, Suppl. Fig. 3A), hence that only subset of cells contained the plasmid-like form of phi3Ts. On the other hand, phi3T could be clearly detected by mapping of resequencing reads of the evolved strains (Suppl. Fig. 3A).

The presence of phi3Ts DNA fragments in the ancestor was additionally confirmed by PCR using a series of primer sets specific for unique phi3Ts fragments, phi3Ts-*kamA* attachment sites, and the *kamA* gene (Suppl. Fig. 4). PCR performed on genomic DNA of *B. subtilis* 168 resulted in a strong band from the intact *kamA* gene, a weak band from selected fragments of phi3Ts, and very faint bands indicating *kamA* integration (Suppl. Fig. 4A). Conversely, PCR on genomic DNA extracted from evolved strains showed the presence of very strong bands for both phi3Ts fragments, and left and right integration sides (except for

194 B310mA in which the right part of *kamA* was absent. As expected, strains B310mA and
195 B410wtB1 were negative for intact *kamA*, while B410mB gave a weak product, which could
196 be explained by incorporation of *kamA* into its epDNA (Fig. 2, Suppl. Fig. 4A).

197 The above analysis indicates that low copy number phi3Ts was present in the *B.*
198 *subtilis* 168 stock from the start. Since *B. subtilis* 168 has been shared among research labs
199 around the world, low copy number phi3Ts could also be 'hiding' in culture stocks of other
200 research labs. Accidental detection of such low copy number phage DNA is nearly impossible,
201 because (i) re-sequencing reads matching phi3Ts would be filtered out during standard
202 mapping pipelines, and (ii) phi3Ts appears to only multiply and manifest itself under specific
203 selection regimes. To check for possible contamination of other *B. subtilis* stocks with phi3Ts,
204 we mapped raw re-sequencing data available in the NCBI database to the phi3T genome
205 (KY030782.1). Analysis of five *B. subtilis* 168 genomes from different laboratories showed
206 no evidence of phi3Ts contamination, since re-sequencing reads matched fragments with high
207 homology to phi3T-*B. subtilis* 168 (Suppl. Fig 3B).

208 In addition to resequencing data analysis, we also PCR-screened a larger
209 collection of *B. subtilis* 168 stocks from different labs around the world[49] for the presence of
210 phi3Ts. Although the vast majority of tested strains lacked phi3Ts sequences (in agreement
211 with sequencing data analysis), a very strong band was observed for *B. subtilis* 168
212 'Newcastle', suggesting that this strain was infected with phi3Ts, or a very similar prophage
213 (Suppl. Fig. 4B). Further PCR analysis confirmed phage integration into the *kamA* gene, but
214 also the presence of an intact *kamA*, indicating that a subpopulation of cells could be
215 pseudolysogenic (Suppl. Fig. 4B). We also confirmed that, similar to the experimentally
216 evolved strains, the Newcastle 168 strain contained the 's' fragment, a unique sequence
217 allowing the phi3Ts phage to be distinguished from the previously sequenced phi3T, hence it
218 is a specific marker for the 'laboratory' phage variant (Fig. 1D).

As phi3Ts multiplies under a prolonged sporulation selection regime, we contacted colleagues who also performed experimental evolution with *B. subtilis* strains imposing the same or similar selection[50,51] . First, we approached a group from the University of Wisconsin-Madison, with whom we had not previously shared strains, because they published a study on the evolution of *B. subtilis* strain PY79 (NC_022898.1) under a prolonged sporulation selection regime[50]. They kindly agreed to share raw sequencing data obtained from 12 evolved single isolates, and we investigated potential changes within prophage regions, and searched for the presence of phi3Ts DNA. We did not find any mutations within prophage regions (Suppl. dataset 2). Furthermore, mapping of raw sequencing reads of evolved PY79 strains to the phi3T genome excluded the presence of phi3T-specific DNA fragments (see Materials and Methods; Suppl. Fig. 5).

We also approached a group from the University of Groningen, who performed experimental evolution of *B. subtilis* 168 under nutrient-limited conditions in which bacteria could neither grow nor complete sporulation (due to *sigF* deletion)[51]. Mapping their raw sequencing reads to the phi3T genome clearly revealed the presence of phi3T-specific reads (Suppl. Fig. 3C). Similar to our case (Fig. 3, Suppl. Fig. 3A), the phage DNA was already present at the start, and it either gradually decreased or increased in two different biological samples (Suppl. Fig. 3C).

The above results strongly suggest that the prophage activation scenario requires not only a sporulation selection regime, but also contamination with low copy number phi3Ts DNA or phage particles. The exchange of strains between Newcastle University (the origin of *B. subtilis* 168 PCR-positive for the phi3Ts-specific fragment) and the University of Groningen, and later between the University of Groningen and our lab, represents a possible transmission route for phi3Ts.

Finally, the evolution experiment performed previously[44] was repeated under the sporulation selection regime using the undomesticated *B. subtilis* NCIB 3610 (hereafter 3610) strain in which the presence of phi3Ts DNA could not be detected during analysis of genome resequencing (Suppl. Fig. 3D, Suppl. Fig. 6A) or by PCR (Suppl. Fig. 6B). This time, alongside classical heat treatment (20 min at 80°C), a chemical spore-selection method (see Materials and Methods) was also employed, along with consecutive testing of lytic activity in the culture supernatant and analysis of the presence of phi3Ts DNA and the integrity of the *kamA* gene at every transfer (Suppl. Fig. 7). To our surprise, lytic activity (Suppl. Fig. 7A) and the release of phages (Suppl. Fig. 7B) were observed as early as the fourth transfer when the sporulation selection regime was applied. Similarly, targeted PCR analysis of host DNA revealed a gradual increase in the phi3T-specific PCR product and a gradual decrease in the PCR product corresponding to intact *kamA* (Suppl. Fig. 8). No lytic activity was observed in a parallel control treatment without the sporulation selection regime (Suppl. Fig. 7A).

Foreign phages modulate sporulation dynamics

We next explored whether propagation of low copy number phi3Ts DNA and its integration into the *kamA* gene has any positive fitness effects on *B. subtilis*. Since expression of *kamA* is dramatically increased upon sporulation entry (*Subti*Wiki website (http://subtiwiki.uni-goettingen.de), we hypothesised that *kamA* may encode a product that is metabolically costly and/or toxic for the bacterium, hence the phi3Ts/phi3Ts-SPβ hybrid lysogen may benefit from inactivation of this gene (Suppl. Figure 9A). However, competition assays between wild-type vs. Δ*kamA* strains with and without sporulation selection revealed no difference in performance between strains (Suppl. Fig. 9B).

We hypothesised that certain genes encoded by phi3Ts may provide benefits to the host under a sporulation/spore revival selection regime. Therefore, we examined the sporulation and spore revival dynamics of *B. subtilis* 3610 deliberately infected with phi3T, a phage stock isolated from the lysogen available from the Bacillus genetic stock center (BGSC). We observed that the phi3T lysogen sporulated prematurely compared with the wild-type strain (Fig. 4). We also observed a general trend indicative of better revival of the phi3T lysogen (Suppl. Fig. 10A), which may include contributions from faster germination (Suppl. Fig. 10B) and/or an altered frequency of premature germination during dormancy (Suppl. Fig. 10C).

These observations indicate the possibility that phi3T/phi3Ts may encode proteins that influence the *B. subtilis* life cycle during sporulation and spore revival. Notably, sporulation regulators have been previously linked to mobile genetic elements (MGEs) in this species[52–54]. Annotation of phi3Ts and phi3Ts-SPβ hybrids (see Materials and Methods) revealed the presence of several genes that could modulate sporulation or spore traits. Specifically, we found a gene (labelled as *rapX*) encoding a putative Rap phosphatase (Suppl. Fig. 11, Suppl. Table 1) sharing high amino acid sequence identify with RapA (unique for phi3Ts) that is known to modulate sporulation timing[55]. We also found that the 's' phi3Ts marker sequence may encode stationary phase survival protein YuiC (100% confidence Phyre prediction = 100% confidence), hence we labelled this sequence *spsX* (Suppl. Fig. 11). In addition, we identified s*spC* that controls spore resistance traits and encodes an acid-soluble protein involved in spore DNA protection (present on both SPβ and phi3Ts)[56]. Notably, we did not find genes that are known to affect spore revival (i.e. germination or spore outgrowth), suggesting that the effects on spore revival may be conferred indirectly (e.g. by modulation of sporulation timing)[57]. Together, these results suggest that the spread of phi3Ts under a prolonged sporulation selection regime might be partly driven by host benefits from the regulatory arsenal associated with phi3Ts and its hybrids.

Recombination between SPβ-like prophages takes place on global and local ecological scales

To understand the ecological relevance of extensive phage recombination observed under a sporulation selection regime, we performed global analysis of prophage elements within the *B. subtilis* clade including *B. cereus* for comparison of more distant species (see Materials and Methods). In a total of 350 fully-assembled genomes, 1365 prophage elements were identified using Phaster software (Suppl. dataset 3). Interestingly, we could immediately identify a cluster of rather large (<100 b) prophages integrated close to the replication terminus, just like phi3Ts, SPβ or phi3Ts-SPβ hybrids. These large prophages were found mainly (86%) within representatives of *B. subtilis*, *B. amyloliquefaciens*, *B. licheniformis* and *B. velezensis* species (Fig. 5A, Suppl. dataset 3). In total, we selected 78 strains carrying a large prophage close to the replication terminus for further analysis (see Materials and Methods; Suppl. Fig. 12A).

Among these strains we identified 23 (including SPβ lysogens *B. subtilis* NCBI 3610 and *B. subtilis* 168) in which large prophages had split the *spsM* gene in a manner identical to SPβ (Suppl. dataset 3), and four *B. subtilis* isolates in which the *kamA* gene was split by a prophage region at exactly the same site, as observed in the hybrid lysogens (Suppl. dataset 3). In the remaining Bacillus strains, the large prophages were mostly integrated close to sporulation-related genes, including a homolog of *fisB* encoding a sporulation-specific membrane fission protein (*B. velezensis* SCDB 291), a homolog of *ymaG* encoding an inner-spore coat protein (*B. atrophaeus* BA59) and a homolog of *cotD* encoding an inner-spore coat protein (*Bacillus amyloliquefaciens H*). Interestingly, 10 strains carried extrachromosomal phage DNA (as predicted by Phaster; Suppl. dataset 3), and in one of them (*B. subtilis* SRCM103612**)** this epDNA was a truncated version of an SPβ-like prophage present within

the chromosome (Fig. 5B). The SRCM103612 prophage contained regions sharing homology to both SPβ and phi3Ts, indicating recombination and an unstable lysogenic cycle within SPβ-like recombinant phages in natural *B. subtilis* isolates (Fig. 5B).

To assess the within-species conservation of large prophages, we performed multiple sequence alignment of all the aforementioned prophage sequences. Prophages clustered according to host species, possibly as a result of phage-host specificity and/or prophage-host coevolution (Suppl. Fig. 12B). To access the natural diversity of large SPβ-like prophages, we collected *Bacillus* sp. genomes carrying a large prophage splitting *spsM* or *kamA* (see Materials and Methods) and compared the phylogenetic tree obtained for these strains (see Materials and Methods) with the phylogenetic tree obtained for their SPβ-like prophages (Fig. 6AB). The strains could be divided into six phylogenetic clades (Fig. 6A), while prophages clustered into three clades ('conservative', 'hybrid' and 'diverse'. The 'conservative' clade comprised prophages that were nearly identical to SPβ, that were also found within closely related *B. subtilis* strains (all were members of the 3610 clade; Fig. 6AB). The 'hybrid' clade comprised phi3T, phi3Ts and all phi3Ts-SPβ hybrids that evolved in the above described experiments under a sporulation selection regime (Fig. 6B). Within the 'diverse' clade the prophage relatedness did not match the phylogenetic relatedness of the host strains (Fig. 6AB). For example, in phylogenetically distinct NCD-2 and WR11, isolated from different sources (Suppl. dataset 3), an identical prophage disrupted the *spsM* gene. By contrast, prophages of closely related strains MB8_B1 and MB8_B10 that were isolated from the same mushroom differed in genetic architecture and in integration site. Indeed, we found that among *B. subtilis* isolates from the same soil sample below the mushroom[58], one strain (MB8_B7) carried an *spsM*-integrated SPβ prophage, one strain (MB8_B1) carried a SPβ-like prophage in *spsM*, and one strain (MB8_B10) carried an SPβ-like prophage in *kamA* (Fig. 6AB). Additionally, we noticed that nearly all members of the 'conservative' clade carried an intact copy of iCEBs1

that was shown to block the SPβ lytic cycle[45], while this element is missing in all members of the 'diverse' clade (Fig. 6B). Finally, we could clearly see modules sharing high homology with SPβ and phi3T in the large prophages (Fig. 6C). These results are consistent with our lab data showing that SPβ-like phages diversify in nature, and this diversification may be constrained by other MGEs present on the host chromosome.

Discussion

The importance of phages in the ecology and evolution of bacteria is indisputable. Interactions between bacteria and temperate phages are especially complex, because the latter can serve as both beneficial genetic cargo and as a constant threat of cell death. Genome comparison suggests that prophage elements undergo pervasive domestication within their hosts that gradually lose the ability to reproduce via the lytic cycle[59]. Our current work demonstrates an opposite scenario, where after a prolonged sporulation/spore revival selection regime, a latent prophage of *B. subtilis* (SPβ)[44,45] regains its lytic reproductive cycle via recombination with 'foreign' phage DNA (phi3Ts). The fact that phi3Ts only manifests itself under specific conditions (a prolonged sporulation/spore revival selection regime) is reminiscent of previously described examples of *Proteobacteria* phages[60,61]. Specifically, the lytic phage SW1 can thrive undetected within *E. coli* populations, but manifests itself in spontaneous plaque formation after overexpression of a putative methylase from an indigenous cryptic prophage[60]. Likewise, lytic variants of P22 spontaneously form upon purine starvation of the *Salmonella typhimurium* host[61]. In our case, an increase in phi3Ts DNA copy number and its integration into the chromosome took place upon application of a sporulation selection regime.

Exactly how sporulation promotes the spread of 'foreign' phage and its hybrid derivatives requires further molecular studies. There are two, not mutually exclusive,

hypotheses: (a) induction of the phage lytic cycle in a small fraction of sporulating cells leads to rapid amplification of phi3Ts DNA, infection of other sporulating cells, segregation of phage DNA into forespores, and trapping of many of its copies in spores, followed by the release of phages upon germination, as observed previously for lytic *B. subtilis* phages[62,63]; (b) since phi3T lysogeny (KY030782.1; 99.98% sequence identity with phi3Ts)[47] results in earlier sporulation and potentially improved spore quality, integration of this phage into the chromosome may be adaptive for the host. As the functions of most phi3Ts genes are obscure, it is difficult to identify the potential phage-encoded regulatory genes that could affect the host life cycle. One possibility is *rap-phr* cassettes (matching *rap* present in phi3T and phi3Ts) that have been previously found within other MGEs of *B. subtilis*, and have been shown to modulate the timing of sporulation[52–54]. Phi3Ts phage genes (e.g. *sspC*) could modulate the production of resistant and viable spores[56] and/or reduce sporulation failure and premature germination[64]. In addition, the spore revival traits of lysogens may also be indirectly affected by the modulation of sporulation timing[57]. Whatever the exact molecular mechanism and its evolutionary driver, the activation of latent prophage elements upon sporulation/spore revival treatment expands the intriguing connections between sporulation of *Firmicutes* and phages infecting these species[29,31,43,65–67].

Based on comparison of the sizes and integration sites of prophage elements within Bacillus sp., SPβ and phi3T clearly belong to a distinct prophage group. These phages appear extremely large (2–3-fold larger than average prophages), and they possess sophisticated communication systems that are potentially capable of sensing the frequency of infected hosts[47,68] or biosynthetic gene clusters[69,70], and functional genes related to host dormancy[71,72]. All these features, combined with regulatory excision upon sporulation, indicate strong codependence between SPβ-like phages and their hosts. A high level of homology between SPβ and phi3Ts offers extensive regions for homologous recombination, which can

be additionally promoted by the recombination machinery involved in natural competence[73] and in non-homologous end-joining repair[74] . It appears that the absence of other mobile genetic elements (e.g. ICEBs1) constraining the phage lytic cycle[45] may also correlate with a higher level of phage diversification. However, whether sporulation promotes recombination between SPβ and phi3Ts alongside phi3Ts amplification remains to be investigated. It is possible that such phage recombination could be facilitated by regulatory excision of SPβ from the chromosome in the sporulating mother cell[29,30]. It also remains to be confirmed whether all *spsM*-splitting prophage elements, such as SPβ, behave like regulatory switches as previously suggested[29]. The disruption of *kamA* by phi3Ts, SPβ-phi3Ts hybrids, and SPβ-like prophages in *Bacilli* suggests that this gene might also be controlled by regulatory excision. Recombination between SPβ and phi3Ts under a sporulation/spore revival selection regime is an example how new regulatory phage-host relationships may evolve.

In addition to regulatory switch behaviour, *Bacilli* and their large SPβ-like prophages pervasively recombine during sporulation, providing new model systems to study bacterial evolution in which phages serve as an evolutionary driving force. Ecological relevance of prophage recombination observed under lab conditions is well supported by natural diversity within the same group of prophage elements on global and local ecological scales. The crucial role of prophage elements on ecological interactions within closely related strains has already been demonstrated for other species[8,60,75,76]. Herein, we showed that such antagonism emerges during the early steps of phage diversification, which may suggest that speciation of prophage elements may be the first step toward speciation of host bacteria. Finally, our work sheds new light on the interplay between bacteria and their phages; while temperate phages commonly undergo domestication [59], they may easily regain genetic mobility by recombination with other phages, thereby altering the physiology, social interactions and evolution of their host.

Materials and Methods

Strains and cultivation conditions. Supplementary Table S2 describes the bacterial strains used in this study and Supplementary Table S3 lists all phages used in this work. Plasmids and oligonucleotides used for cloning purposes to construct some of the strains used here are listed in Supplementary Table S4. Strains were routinely maintained in lysogeny broth (LB) medium (LB-Lennox, Carl Roth; 10 g/l tryptone, 5 g/l yeast extract, and 5 g/l NaCl).

Strain DTUB200 was obtained by infecting DK1042 (WT NCBI 3610) with a phage phi3T obtained from CU1065. DTUB201 (ΔSPβ) was obtained by transforming DK1042 with gDNA obtained from SPmini and selecting for erythromycin-resistant colonies. Strain DTUB202 (Δ*kamA*) was obtained by transforming DK1042 with gDNA obtained from BKK19690 and selecting for kanamycin-resistant colonies. All modifications of DK1042 were verified by PCR followed by Sanger sequencing. Strain DTUB203 (P_L-*gfp*) was obtained by transforming DK1042 with pDTUB206 (P_L-*gfp*) plasmid and selecting for chloramphenicol-resistant colonies. To obtain this plasmid, P_L promoter was amplified from *B. subtilis* 168 gDNA using oAD10 and oAD11, introducing the EcoRI and NheI restriction sites. The PCR product was then ligated into pre-digested pGFP-rrnB plasmid to obtain pDTUB206. Strains DTUB204 (P_L-gfp^{phi3T}) and DTUB205 (P_L-gfp^{phi3Ts}) were obtain by infecting the DTUB203 with phi3T and phi3Ts phages, respectively.

Genome sequencing and analysis

Phage sequencing was performed by Illumina MiSeq instrument and a 2x250 nt paired-end chemistry (MiSeq Reagent Kit v2 (500-cycles). Primary data analysis (base-calling) was

carried out with Bbcl2fastq^ software (v2.17.1.14, Illumina). In vitro fragment libraries were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina. Reads were quality and length trimmed in CLC Genomics Workbench Tool 11.0 and *de novo* genome assembly was performed using SPAdes-3.13.0-Linux and CLC Genomics Workbench 11.0.

De novo sequencing and assembly of B310mA, B410mB and B410wtB genomes was performed by Functional Genomics Center Zurich, from genomic DNA of exponentially grown cultures, extracted using the EURex Bacterial and Yeast Genomic DNA Kit. Resequencing of 168 ancestor (ancestor of B310mA, B410mB and B410wtB) was described in our previous manuscript[44].

Evolved PY79 strains (presented in Suppl. dataset 2) were obtained as previously described[50]. Samples for whole-genome sequencing were prepared according to the Illumina Multiplexing Sample Preparation Guide, using NEBNext reagents and Illumina's indexed primers. Sequencing was performed by the Bauer Core Facility at Harvard University. Mapping of raw fastq reads to reference PY79 genome (NC_022898.1) was performed using Bowtie2[77,78]. The alignment was sorted using SAMtools[79,80], data filtering and SNP variant calling was performed using the bcftools package. Mapping of raw fastq reads to phi3T genome (KY030782.1) was performed using Bowtie2 in Galaxy platform (https://cpt.tamu.edu/galaxy-pub) and coverage was visualized in the browser using Trackster tool. Mapping of raw SOLiD resequencing reads (168_{anc}) to unique phi3Ts fragments was performed using CLC Genomics Workbench 11.0.1. Short phi3T fragments, to which fastq could be mapped, showed over 90% sequence identity to PY79 genome, as confirmed by BLAST. All bacterial and phage genomes sequenced during this work, where deposited at NCBI database as completed genomes and/or raw sequencing data (Table 1).

Raw re-sequencing data of PY79 strains[50] and available from B. Burton (briana.burton@wisc.edu). Raw re-sequencing data of 168 cultivated under near- zero growth conditions[51] are available from O. Kuipers (o.p.kuipers@rug.nl).

Sporulation and germination assays

To examine sporulation dynamics selected strains were cultivated in MSgg medium[81] at 30°C, 220 rpm, and total CFU and spore counts were analysed after 12, 24 and 36 hours. To access the spore count, cells were incubated at 80°C for 20min, plated on LB-agar (1.5%) and the number of obtained colonies was divided by the number of colonies obtained prior to the heat-treatment. To access the germination, the culture incubation was prolonged to 72h to allow vast majority of cells to sporulate. Next, spores were washed 2× with 0.9% NaCl, and resuspendend in germination solution (0.6g KH2PO4, 1.4g K_2HPO_4, 0.2g $(NH_4)_2SO_4$, 0.1g Na-citrate, 0.02g $MgSO_4 \times 7H_2O$, 0.5g glucose, 3.56g L-alanine resuspended in 100ml of dH_2O) to reach final OD_{600} cca 10. Decline of OD_{600} was measured immediately, indicating germination[82]. Additional assessment of germination dynamics was performed using real-time brightfield microscopy by inducing spores with L-alanine on agarose pads, as described previously[57]. Agarose pads (1.5%, 9 mm diameter, 1 mm height) were inoculated with 2.6 µl spore solution ($3.75*10^5$ spores μl^{-1}) and placed upside down into a 24-well glass-bottom microtiter plate. Germination was induced by adding 5 µl of a 200 mM L-alanine solution to the top of the pad. Germination events were monitored by changes in grey level spore intensity. The fraction of germinated spores at time t was calculated as the number of germinated spores divided by the number of dormant spores before induction (i.e.by excluding pre-germinated spores).

Spore selection experiment with NCBI 3610

Strains were cultivated in 10ml of MSgg medium in 100ml-glass bottles in 30°C with shaking at 220 rpm. Every 48 hours, three alternative transfer methods were applied: direct transfer of untreated cells to fresh medium, transfer of heat-treated cells (80°C for 20 min) and transfer of chemically treated cells (5% NaOH for 2 min, followed by washing in PBS). In each case, fresh cultures were initiated with 1% inoculum. Culture supernatants and cell pellets were collected prior each transfer to monitor phage release and genetic rearrangements, respectively. At each transfer, frozen stocks were preserved, to allow the analysis of subsequent steps of phage recombination in the future.

Isolation of phage particles and phage DNA

All lysogenic strains that were used as source of phages, were producing phage particles spontaneously, therefore treatment with Mitomycin C was not needed to obtain phages from culture supernatants. Lysogens were cultivated in LB medium at 37°C with shaking at 200 rpm for 8h. Culture supernatants were collected, adjusted to pH of 7.0, filter-sterilized and mixed at a 1:4 rate with PEG-8000 solution (PEG-8000 20%, 116 g/l NaCl). After overnight incubation at 4°C, the solutions were centrifuged for 60 min at 12000 rpm to obtain phage precipitates. The pellets were resuspended in 1% of the initial volume in SM buffer (5.8 g/lNaCl, 0.96 g/l $MgSO_4$, 6 g/l Tris-HCl, pH 7.5) to obtain concentrated solution of phage particles. Such phage solutions were visualized by transmission electron microscopy and used as a source of different phage variants, purified from single plaques. In plaque assay and further phage propagation from single plaques, Δ6 strain[83] was used as a host. Specifically, phage solutions were diluted in order to obtain well-separated single plaques. Selected plaques (differing with morphology) were carefully removed from the soft agar using sterile scalpel, resuspended in 200µl of SM buffer and used to infect

exponentially growing phage-free host to allow propagation of selected phage variants. Phages were subsequently propagated in soft agar and liquid host suspension until the titer reached at least 10^9 pfu/ml and then subjected to DNA isolation. Phage DNA was extracted using phenol-chloroform method, as described previously[84].

Transmission electron microscopy

Before use, 400 mesh nickel grids with a 3-4 nm thick carbon film, CF400-Ni-UL EMS Diasum, were hydrophilized by 30 sec of electric glow discharging. Next, 5µl of purified phage solutions were applied onto the grids and allowed to adsorb for 1 minute. The grids were rinsed 3 times on droplets of milliQ water and subjected to staining with 2% uranyl acetate. Specifically, with a help of EM grid-grade tweezers, the grids were placed sequentially on droplets of 2% uranyl acetate solution for 10 sec, 2 sec and 20 sec. Excess uranyl acetate was wicked away using filter paper and the grids were allowed to dry overnight and stored in a desiccator until analysis. Transmission electron microscopy was performed utilizing a FEI Tecnai T12 Biotwin TEM operating at 120 kV located at the Center for Electron Nanoscopy at the Technical University of Denmark, and images were acquired using a Bottom mounted CCD, Gatan Orius SC1000WC.

Prophage database construction and phage comparisons

Bacillus prophage database was constructed by finding genomic coordinates using Phaster software[85,86] from fully assembled *Bacillus* genomes available at NCBI, followed by extraction of phage genomes using samtools package. In total, the initial database contained 350 strains, which altogether carried 1365 prophage elements. Out of these prophages, 54 were selected for further analysis according to following criteria: all prophages larger than 80kB (regardless of integration side) and all prophages that are at least 50 kB, integrated between 1.9-2.3 Mb in the chromosome, just like SPβ and phi3Ts-SPβ hybrids identified in the evolved strains. Additional prophages, categorized as SPβ-like, were retrieved the genomes that gave BLAST

hits to phi3T and SPβ, if these hits belonged to a prophage region that was at least 40kB (confirmed by Phaster). All genomes that were re-sequenced copies of *Bacillus subtilis* 168 were removed. In addition, genomes that were starting and finishing with a prophage (likely due to misassembly), were removed (NZ_CP032855.1). Overall, 78 *Bacillus* strains lysogenic for putative SPβ-like prophages were subjected to further analysis.

Interruption of *spsM* and *kamA* in all the selected lysogens was examined by genome BLAST against the sequence of an intact copy of these gene. All strains that carried a split copy of *spsM* and *kamA*, also carried a large prophage between left and right arms of these genes. In such cases, the Phaster-predicted terminal positions of the prophage was corrected to exactly match the sequence splitting *spsM* or *kamA*. Such correction was based on experimentally confirmed sequences of phage DNA. Integration genes of remaining large prophages were determined by extracting and clustering 1000bp-long prophage flanking regions using vsearch at 46% identity. These regions were then compared to well-annotated *B. subtilis* 168, using blastx, to find functional homologs.

The alignment of prophage sequences was performed in MAFFT program[87], phylogenetic tree was build using FastTree[88,89] and visualized in CLC Main Workbench. Phylogenetic tree of *B. subtilis* host strains was constructed using open software autoMLST (https://automlst.ziemertlab.com/)[90] based on 100 shared proteins. Two strains that were not lysogenic for SPβ-like prophage (MB9_B4 and MB9_B6) were included in the analysis to exclude SPβ prophage from the shared pool of proteins in the tree building. The three was visualized in CLC Main Workbench. Prophage annotation was performed using RAST online annotation platform.

Statistical analysis

Statistical differences between two experimental groups were identified using two-tailed Student's *t*-tests assuming equal variance. No statistical methods were used to predetermine sample size and the experiments were not randomized.

Authors contributions:

AD, PB, ZH, CK performed experiments. AD and MLS performed bioinformatics analysis. PK performed electron microscopy, GM performed genome sequencing and analyzed the data, BB and BMB shared sequencing data. AD and ATK designed the study. AD wrote the manuscript. All authors contributed to final version of the manuscript.

Acknowledgements:

The authors thank M. Kilstrup, P. Sazinas, K. Middleboe, D. Castillio and P. Stefanic for their valuable comments. We are profoundly grateful to O. Kuipers, A. de Jong and W. Overkamp from University of Groningen, for sharing their raw sequencing data and all relevant information, which allowed us to finalize the manuscript. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 713683 (H.C. Ørsted COFUND to A.D.), Individual grant from Friedrich Schiller University Jena to support postdoc researchers to A.D., and supported by the Danish National Research Foundation (DNRF137) for the Center for Microbial Secondary Metabolites. Funded in part by NIH R01GM121865 to BMB.

581 **Figure legends:**

582 **Figure 1.** Changes within *B. subtilis* prophage sequence and integration site observed after
583 prolonged sporulation selection regime. A) Experimental evolution with sporulation selection
584 regime leads to spontaneous release of phage particles by the evolved strains[44]. B). Overnight
585 culture of evolved *B. subtilis* strain B410mB (*amyE::mKate*, shown in red) was diluted 100×
586 and spotted on the lawn of undiluted *B. subtilis* ancestor strain (*amyE::gfp*, shown in green),
587 resulting in a clearance zone, and growth of B410mB in that zone. The same experiment was
588 performed using 100x diluted culture of ancestor strain (*amyE::mKate*) on a lawn of undiluted
589 ancestor (*amyE::gfp*), as control.Scale bar=2.5mm. C) Schematic representation of genome
590 rearrangements in one of the phage-releasing evolved strains (B310mA), compared to the
591 ancestor (Anc). The evolved strains carry a hybrid prophage phi3Ts-SPβ. Fragments of phi3Ts
592 are shown in black, while fragments of SPβ are shown in pink. Below, schematic representation
593 of phage genomes, spontaneously released by B310mA. D) Schematic comparison of phi3Ts
594 genome, with genome of Bacillus phage phi3T (KY030782.1). Fragment 's' which is unique
595 for phi3Ts, can be detected within prophage genomes of 6 *B. subtilis* strains, isolated in
596 different parts of the world, specifically: SRCM103612 (South Korea), MB9_B1 and MB8_B1
597 (Denmark), JAAA (China), HMNig-2 (Egypt) and SSJ-1 (South Korea).

598 **Figure 2.** Extrachromosomal fragments of phage DNA, detected in the evolved strains. Top:
599 Genome comparison of phi3Ts and SPβ (Query cover=58%, Percent Identity=99.73%), where
600 regions of high homology (73.6-100%) are shown in grey, and regions of 99% homology are
601 connected. Segments that are unique for phi3Ts, or SPβ are highlighted in black and pink,
602 respectively. Phage genomes are arranged according to their integration into the host
603 chromosome, which is represented in red. Below: extrachromosomal phage DNA fragments
604 detected during PacBio sequencing, colored according to their homology to phi3Ts, SPβ, or
605 fragments of host chromosome flanking phage integration sites. Fragments are ordered

606 according to sequencing coverage relative to the chromosomal region, which is represented as
607 bar chart on the left.

608 **Figure 3**. Detection of phi3Ts DNA in the ancestor strain *B. subtilis* 168 through mapping of
609 raw sequencing reads. Top: Representation of phi3Ts genome according to its homology to
610 SPβ prophage. Fragments of high homology to SPβ (73.6-100%) are shown in grey, while
611 fragments that are unique to phi3Ts are shown in black. Bars 1,2 and 3 correspond to DNA
612 sequences that are unique for phi3Ts and that were used as targets for raw reads mapping (lower
613 part). Green and red bars represent reads obtained from forward and reverse strands,
614 respectively.

615 **Figure 4.** Effect of phi3T infection on *B. subtilis* sporulation and germination dynamics. A)
616 Sporulation dynamics. Percentage of spores compared to total cell count, were examined in *B.*
617 *subtilis* 3610 and the same strain infected with phi3T phage, in 3 different time points of growth
618 in minimal medium (MSgg). Data represent an average from 4 biological replicates, error bars
619 correspond to standard error.

620 **Figure 5.** Overview of prophage elements of natural *Bacillus* sp. isolates. A) Prophage
621 elements were extracted from fully assembled genomes of *Bacillus* sp. and plotted according
622 to size and integration position in the chromosome. Cluster of large prophages, integrated in
623 the area of replication terminus could be detected (black dotted line). B) Schematic
624 representation of SPβ-like prophage found in *B. subtilis* SRCM 103612, isolated from
625 traditional Korean food. The prophage genome was colored according to its homology to
626 phi3Ts and SPβ. Extrachromosomal phage DNA found in this strain is matching left and right
627 arms of the chromosomal prophage.

628 **Figure 6.** Natural diversity of SPβ-like phages. A). Phylogenetic tree of *B. subtilis* strains that
629 carry SPβ-like prophage in *spsM* or *kamA* gene, and two control strains that are free from such

prophage. The tree was arbitrarily divided into 6 clades. B) Phylogenetic tree of SPβ-like prophages hosted by the strains in A). Inner circle shows prophage integration site, while outer circle indicates presence/absence of conjugative element ICEBs1, which blocks SPβ lytic cycle C). Selected prophages of Bacillus sp. colored according to their homology to phi3T and SPβ. The upper 4 sequences integrate either in *kamA* or *spsM* and clearly belong to SPβ-like phages. Bottom four sequences come from other *Bacillus* species, and although they are more distant to phi3T or SPβ, they still carry segments of high homology with these phages. Explanation of ICEBs1 figure legend: intact – intact copy (100% identity to *B. subtilis* 168 or NCBI 3610) of ICEBs1 conjugative element is present; negative – lack of BLAST hits to ICEBs1 sequence; partial – at least 70% of ICEBs1 sequence is present; residual – less than 5% of ICEBs1 sequence is present.

Table 1 | List of bacterial strain and phages subjected to genome sequencing with corresponding NCBI accession numbers.

Name of bacterial strain/phage	Data	Accession number
B310mA	Complete genome	CP051860
B410mB	Complete genome	CP053102*
B410wtB	Complete genome	CP052842*
B310mA	Sequencing reads (Illumina)	SRR11561554
B410mB	Sequencing reads (Illumina)	SRR1156151
B410wtB	Sequencing reads (Illumina)	SRR11561552
$168_{ancestor}$	Sequencing reads (SOLiD)	SRR11559011
NCBI 3610	Sequencing reads (Illumina)	SRR11559035
15.1	Sequencing reads (Illumina)	SRR11566357
16.1	Sequencing reads (Illumina)	SRR11566355

16.2	Sequencing reads (Illumina)	SRR11566354
phi3Ts	Complete genome	MT366945
Hyb1$^{phi3Ts\text{-}SP\beta}$	Complete genome	MT366946
Hyb2$^{phi3Ts\text{-}SP\beta}$	Complete genome	MT366947
Hyb3$^{phi3Ts\text{-}SP\beta}$	Complete genome	MT366948
phi3Ts	Sequencing reads (Illumina)	SRR11587866
Hyb1$^{phi3Ts\text{-}SP\beta}$	Sequencing reads (Illumina)	SRR11587864
Hyb2$^{phi3Ts\text{-}SP\beta}$	Sequencing reads (Illumina)	SRR11587865
Hyb3$^{phi3Ts\text{-}SP\beta}$	Sequencing reads (Illumina)	SRR11587867

*extrachromosomal phage fragments: B410wtB - CP052843; B410mB – supplementary dataset 4.

1. Knowles, B. *et al.* Lytic to temperate switching of viral communities. *Nature* **531**, 466–470 (2016).

2. Koskella, B. & Brockhurst, M. A. Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* **38**, 916–931 (2014).

3. Azam, A. H. & Tanji, Y. Bacteriophage-host arm race: an update on the mechanism of phage resistance in bacteria and revenge of the phage with the perspective for phage therapy. *Appl. Microbiol. Biotechnol.* **103**, 2121–2131 (2019).

4. Howard-Varona, C., Hargreaves, K. R., Abedon, S. T. & Sullivan, M. B. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *ISME J.* **11**, 1511–1520 (2017).

5. Harrison, E. & Brockhurst, M. A. Ecological and evolutionary benefits of temperate phage: what does or doesn't kill you makes you stronger. *BioEssays* **39**, 1700112 (2017).

6. Davies, E. V., Winstanley, C., Fothergill, J. L. & James, C. E. The role of temperate bacteriophages in bacterial infection. *FEMS Microbiol. Lett.* **363**, fnw015 (2016).

7. Kim, M. S. & Bae, J. W. Lysogeny is prevalent and widely distributed in the murine gut microbiota. *ISME J.* **12**, 1127–1141 (2018).

8. Štefanič, P., Kraigher, B., Lyons, N. A., Kolter, R. & Mandić-Mulec, I. Kin discrimination between sympatric *Bacillus subtilis* isolates. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 14042–14047 (2015).

9. Lyons, N. A., Kraigher, B., Štefanič, P., Mandić-Mulec, I. & Kolter, R. A combinatorial kin discrimination system in *Bacillus subtilis*. *Curr. Biol.* **26**, 733–742 (2016).

10. Canchaya, C., Proux, C., Fournous, G., Bruttin, A. & Brüssow, H. Prophage genomics. *Microbiol. Mol. Biol. Rev.* **67**, 238–76, (2003).

11. Bérard, S. *et al.* Aligning the unalignable: bacteriophage whole genome alignments. *BMC Bioinformatics* **17**, 30 (2016).

12. Botstein, D. A theory of modular evolution for bacteriophages. *Ann. N. Y. Acad. Sci.* **354**, 484–491 (1980).

13. Hatfull, G. F. & Hendrix, R. W. Bacteriophages and their genomes. *Curr. Opin. Virol.* **1**, 298–303 (2011).

14. De Paepe, M. *et al.* Temperate phages acquire DNA from defective prophages by relaxed homologous recombination: the role of Rad52-like recombinases. *PLoS Genet.*

10, (2014).

15. Swenson, K. M., Guertin, P., Deschênes, H. & Bergeron, A. Reconstructing the modular recombination history of *Staphylococcus aureus* phages. *BMC Bioinformatics* **14**, S17 (2013).

16. Morris, P., Marinelli, L. J., Jacobs-Sera, D., Hendrix, R. W. & Hatfull, G. F. Genomic characterization of mycobacteriophage giles: Evidence for phage acquisition of host DNA by illegitimate recombination. *J. Bacteriol.* **190**, 2172–2182 (2008).

17. Bobay, L.-M., Touchon, M. & Rocha, E. P. C. Manipulating or superseding host recombination functions: A dilemma that shapes phage evolvability. *PLoS Genet.* **9**, e1003825 (2013).

18. Spancake, G. A., Hemphill, H. E. & Fink, P. S. Genome organization of Sp beta c2 bacteriophage carrying the *thyP3* gene. *J. Bacteriol.* **157**, 428–34 (1984).

19. Fillol-Salom, A. *et al.* Bacteriophages benefit from generalized transduction. *PLOS Pathog.* **15**, e1007888 (2019).

20. Uchiyama, J. *et al.* Intragenus generalized transduction in Staphylococcus spp. by a novel giant phage. *ISME J.* **8**, 1949–1952 (2014).

21. Morse, M. L., Lederberg, E. M. & Lederberg, J. Transduction in *Escherichia coli* K-12. *Genetics* **41**, 142–56 (1956).

22. Fukumaki, Y., Shimada, K. & Takagi, Y. Specialized transduction of Colicin E1 DNA in *Escherichia coli* K-12 by phage Lambda. *Proc. Natl. Acad. Sci. U. S. A.* **73**, 3238–3242

23. Penadés, J. R., Chen, J., Quiles-Puchalt, N., Carpena, N. & Novick, R. P. Bacteriophage-mediated spread of bacterial virulence genes. *Current Opinion in*

Microbiology **23**, 171–178 (2015).

24. Kupczok, A. *et al.* Rates of mutation and recombination in *Siphoviridae* phage genome evolution over three decades. *Mol. Biol. Evol.* **35**, 1147–1159 (2018).

25. Yahara, K., Lehours, P. & Vale, F. F. Analysis of genetic recombination and the pan-genome of a highly recombinogenic bacteriophage species. *Microb. genomics* **5**, (2019).

26. Yamamoto, N., Wohlhieter, J. A., Gemski, P. & Baron, L. S. λimmP22dis: A hybrid of coliphage λ with both immunity regions of Salmonella phage P22. *Mol. Gen. Genet. MGG* **166**, 233–243

27. Botstein, D. & Herskowitz, I. Properties of hybrids between Salmonella phage P22 and coliphage λ. *Nature* **251**, 584–589 (1974).

28. Feiner, R. *et al.* A new perspective on lysogeny: prophages as active regulatory switches of bacteria. *Nat. Rev. Microbiol.* **13**, 641–650 (2015).

29. Abe, K. *et al.* Developmentally-regulated excision of the SPβ prophage reconstitutes a gene required for spore envelope maturation in *Bacillus subtilis* . *PLoS Genet.* **10**, e1004636 (2014).

30. Abe, K., Takamatsu, T. & Sato, T. Mechanism of bacterial gene rearrangement: SprA-catalyzed precise DNA recombination and its directionality control by SprB ensure the gene rearrangement and stable expression of *spsM* during sporulation in *Bacillus subtilis*. *Nucleic Acids Res.* **45**, 6669–6683 (2017).

31. Haraldsen, J. D. & Sonenshein, A. L. Efficient sporulation in *Clostridium difficile* requires disruption of the σK gene. *Mol. Microbiol.* **48**, 811–821 (2003).

32. Stragier, P., Kunkel, B., Kroos, L. & Losick, R. Chromosomal rearrangement

726 generating a composite gene for a developmental transcription factor. *Science* **243**,
727 507–512 (1989).

728 33. Wood, J. P. *et al.* Environmental persistence of *Bacillus anthracis* and *Bacillus subtilis*
729 spores. *PLoS One* **10**, e0138083 (2015).

730 34. Cano, R. J. & Borucki, M. K. Revival and identification of bacterial spores in 25- to
731 40-million-year-old Dominican amber. *Science* **268**, 1060–1064 (1995).

732 35. Sanchez-Vizuete, P. *et al.* Identification of *ypqP* as a new *Bacillus subtilis* biofilm
733 determinant that mediates the protection of *Staphylococcus aureus* against
734 antimicrobial agents in mixed-species communities. *Appl. Environ. Microbiol.* **81**,
735 109–18 (2015).

736 36. Kimura, T., Amaya, Y., Kobayashi, K., Ogasawara, N. & Sato, T. Repression of *sigK*
737 intervening (*skin*) element gene expression by the CI-like protein SknR and effect of
738 SknR depletion on growth of *Bacillus subtilis* cells. *J. Bacteriol.* **192**, 6209–6216
739 (2010).

740 37. Kunkel, B., Losick, R. & Stragier, P. The *Bacillus subtilis* gene for the developmental
741 transcription factor σ(K) is generated by excision of a dispensable DNA element
742 containing a sporulation recombinase gene. *Genes Dev.* **4**, 525–535 (1990).

743 38. Pyne, M. E., Liu, X., Moo-Young, M., Chung, D. A. & Chou, C. P. Genome-directed
744 analysis of prophage excision, host defence systems, and central fermentative
745 metabolism in *Clostridium pasteurianum*. *Sci. Rep.* **6**, 26228 (2016).

746 39. Sohail, A., Hayes, C. S., Divvela, P., Setlow, P. & Bhagwat, A. S. Protection of DNA
747 by α/β-type small, acid-soluble proteins from *Bacillus subtilis* spores against cytosine
748 deamination. *Biochemistry* **41**, 11325–11330 (2002).

749 40. Ki, S. L., Bumbaca, D., Kosman, J., Setlow, P. & Jedrzejas, M. J. Structure of a
750 protein-DNA complex essential for DNA protection in spores of Bacillus species.
751 *Proc. Natl. Acad. Sci. U. S. A.* **105**, 2806–2811 (2008).

752 41. Jiang, M., Grau, R. & Perego, M. Differential processing of propeptide inhibitors of
753 rap phosphatases in *Bacillus subtilis*. *J. Bacteriol.* **182**, 303–310 (2000).

754 42. Serra, C. R., Earl, A. M., Barbosa, T. M., Kolter, R. & Henriques, A. O. Sporulation
755 during growth in a gut isolate of *Bacillus subtilis*. *J. Bacteriol.* **196**, 4184–4196 (2014).

756 43. Silver-Mysliwiec, T. H. & Bramucci, M. G. Bacteriophage-enhanced sporulation:
757 Comparison of spore-converting bacteriophages PMB12 and SP10. *J. Bacteriol.* **172**,
758 1948–1953 (1990).

759 44. Martin, M. *et al.* De novo evolved interference competition promotes the spread of
760 biofilm defectors. *Nat. Commun.* **8**, 15127 (2017).

761 45. Eleina England, by M. & Bell Professor of Biology, S. P. Effects of cell growth and a
762 mobile genetic element on propagation of the phages SP16 and SP-beta in *Bacillus*
763 *subtilis* (2014).

764 46. Warner, F. D. *et al.* Characterization of SPP: a temperate bacteriophage from *Bacillus*
765 *subtilis* 168M. *Can J Microbiol* **23**, 45-51 (1976).

766 47. Erez, Z. *et al.* Communication between viruses guides lysis-lysogeny decisions. *Nature*
767 **541**, 488–493 (2017).

768 48. Dennehy, J. J. Bacteriophage Ecology: Population growth, evolution, and impact of
769 bacterial viruses. Part of Advances in Molecular and Cellular Biology - *The Quarterly*
770 *Review of Biology* (2010).

771 49. Gallegos-Monterrosa, R., Mhatre, E. & Kovács, A. T. Specific *Bacillus subtilis* 168

772 variants form biofilms on nutrient- rich medium. Microbiology 2016;162:1922–32.ch
773 medium. *Microbiology* **162**, 1922–1932 (2016).

774 50. Bose, B., Reed, S. E., Besprozvannaya, M. & Burton, B. M. Missense mutations allow
775 a sequence-blind mutant of SpoIIIE to successfully translocate chromosomes during
776 sporulation. *PLoS One* **11**, e0148365 (2016).

777 51. Overkamp, W. *et al.* Physiological and cell morphology adaptation of *Bacillus subtilis*
778 at near-zero specific growth rates: A transcriptome analysis. *Environ. Microbiol.* **17**,
779 346–363 (2015).

780 52. Omer Bendori, S., Pollak, S., Hizi, D. & Eldar, A. The RapP-PhrP quorum-sensing
781 system of *Bacillus subtilis* strain NCIB3610 affects biofilm formation through multiple
782 targets, due to an atypical signal-insensitive allele of RapP. *J. Bacteriol.* **197**, 592–602
783 (2015).

784 53. Singh, P. K. *et al.* Mobility of the native *Bacillus subtilis* conjugative plasmid pLS20
785 is regulated by intercellular signaling. *PLoS Genet.* **9**, e1003892 (2013).

786 54. Auchtung, J. M., Lee, C. A., Monson, R. E., Lehman, A. P. & Grossman, A. D.
787 Regulation of a *Bacillus subtilis* mobile genetic element by intercellular signaling and
788 the global DNA damage response. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12554–12559
789 (2005).

790 55. Perego, M. & Hoch, J. A. Cell-cell communication regulates the effects of protein
791 aspartate phosphatases on the phosphorelay controlling development in *Bacillus*
792 *subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1549–53 (1996).

793 56. Tovar-Rojo, F. & Setlow, P. Effects of mutant small, acid-soluble spore proteins from
794 *Bacillus subtilis* on DNA in vivo and in vitro. *J. Bacteriol.* **173**, 4827–4835 (1991).

795 57. Mutlu, A. *et al.* Phenotypic memory in *Bacillus subtilis* links dormancy entry and exit
796 by a spore quantity-quality tradeoff. *Nat. Commun.* **9**, (2018).

797 58. Kiesewalter, H. T. *et al.* Complete genome sequences of 13 *Bacillus subtilis* soil
798 isolates for studying secondary metabolite diversity . *Microbiol. Resour. Announc.* **9**,
799 (2020).

800 59. Bobay, L.-M., Touchon, M. & Rocha, E. P. C. Pervasive domestication of defective
801 prophages by bacteria. *Proc. Natl. Acad. Sci.* **111**, 12127–12132 (2014).

802 60. Song, S., Guo, Y., Kim, J.-S., Wang, X. & Wood, T. K. Phages mediate bacterial self-
803 recognition. *Cell Rep.* **27**, 737-749.e4 (2019).

804 61. Downs, D. M. & Roth, J. R. A novel P22 prophage in *Salmonella typhimurium*.
805 *Genetics* **117**, 367–80 (1987).

806 62. Moreno, F. On the trapping of phage genomes in spores of *Bacillus subtilis* 168
807 reciprocal exclusion of phages φ29 and φe during outgrowth of spores. *Virology* **93**,
808 357–368 (1979).

809 63. Sonenshein, A. L. Trapping of unreplicated phage DNA into spores of *Bacillus subtilis*
810 and its stabilization against damage by 32P decay. *Virology* **42**, 488–495 (1970).

811 64. Ramírez-Guadiana, F. H., Meeske, A. J., Wang, X., Rodrigues, C. D. A. & Rudner, D.
812 Z. The *Bacillus subtilis* germinant receptor GerA triggers premature germination in
813 response to morphological defects during sporulation. *Mol. Microbiol.* **105**, 689–704
814 (2017).

815 65. Lewis, R. J., Brannigan, J. A., Offen, W. A., Smith, I. & Wilkinson, A. J. An
816 evolutionary link between sporulation and prophage induction in the structure of a
817 repressor:anti-repressor complex. *J. Mol. Biol.* **283**, 907–912 (1998).

818 66. Sonenshein, A. L. Bacteriophages: How bacterial spores capture and protect phage
819 DNA. *Current Biology* **16**, (2006).

820 67. Castilla-Llorente, V., Muñoz-Espín, D., Villar, L., Salas, M. & Meijer, W. J. J. Spo0A,
821 the key transcriptional regulator for entrance into sporulation, is an inhibitor of DNA
822 replication. *EMBO J.* **25**, 3890–3899 (2006).

823 68. Babel, H. *et al.* Ratiometric population sensing by a pump-probe signaling system in
824 *Bacillus subtilis*. *Nat. Commun.* **11**, 1–13 (2020).

825 69. Paik, S. H., Chakicherla, A. & Hansen, J. N. Identification and characterization of the
826 structural and transporter genes for, and the chemical and biological properties of,
827 sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168. *J. Biol. Chem.*
828 **273**, 23134–23142 (1998).

829 70. Denham, E. L. *et al.* Differential expression of a prophage-encoded glycocin and its
830 immunity protein suggests a mutualistic strategy of a phage and its host. *Sci. Rep.* **9**,
831 2845 (2019).

832 71. Lazarevic, V. *et al.* Nucleotide sequence of the *Bacillus subtilis* temperate
833 bacteriophage SPβc2. *Microbiology* **145**, 1055–1067 (1999).

834 72. Moeller, R., Setlow, P., Reitz, G. & Nicholson, W. L. Roles of small, acid-soluble
835 spore proteins and core water content in survival of *Bacillus subtilis* spores exposed to
836 environmental solar UV radiation. *Appl. Environ. Microbiol.* **75**, 5202–8 (2009).

837 73. Schultz, D., Wolynes, P. G., Jacob, E. Ben & Onuchic, J. N. Deciding fate in adverse
838 times: Sporulation and competence in *Bacillus subtilis*.

839 74. de Vega, M. The minimal *Bacillus subtilis* nonhomologous end joining repair
840 machinery. *PLoS One* **8**, e64232 (2013).

841 75. Lyons, N. A., Kraigher, B., Stefanic, P., Mandic-Mulec, I. & Kolter, R. A
842 Combinatorial kin discrimination system in *Bacillus subtilis*. *Curr. Biol.* **26**, 733–42
843 (2016).

844 76. Dey, A. *et al.* Sibling rivalry in *Myxococcus xanthus* is mediated by kin recognition
845 and a polyploid prophage. *J. Bacteriol.* **198**, 994–1004 (2016).

846 77. Langmead, B., Wilks, C., Antonescu, V. & Charles, R. Scaling read aligners to
847 hundreds of threads on general-purpose processors. *Bioinformatics* **35**, 421–432
848 (2019).

849 78. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat.*
850 *Methods* **9**, 357–359 (2012).

851 79. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,
852 2078–2079 (2009).

853 80. Li, H. A statistical framework for SNP calling, mutation discovery, association
854 mapping and population genetical parameter estimation from sequencing data.
855 *Bioinformatics* **27**, 2987–2993 (2011).

856 81. Branda, S. S., Gonzalez-Pastor, J. E., Ben-Yehuda, S., Losick, R. & Kolter, R. Fruiting
857 body formation by *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 11621–11626
858 (2001).

859 82. Harwood, C. R. & Cutting, S. M. Molecular biological methods for Bacillus. (Wiley,
860 1990).

861 83. Westers, H. *et al.* Genome engineering reveals large dispensable regions in *Bacillus*
862 *subtilis*. *Mol. Biol. Evol.* **20**, 2076–2090 (2003).

863 84. Tóth, I., Sváb, D., Bálint, B., Brown-Jaque, M. & Maróti, G. Comparative analysis of

the Shiga toxin converting bacteriophage first detected in *Shigella sonnei*. *Infect. Genet. Evol.* **37**, 150–157 (2016).

85. Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J. & Wishart, D. S. PHAST: A Fast Phage Search Tool. *Nucleic Acids Res.* **39**, (2011).

86. Arndt, D. *et al.* PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* **44**, W16–W21 (2016).

87. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

88. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **26**, 1641–1650 (2009).

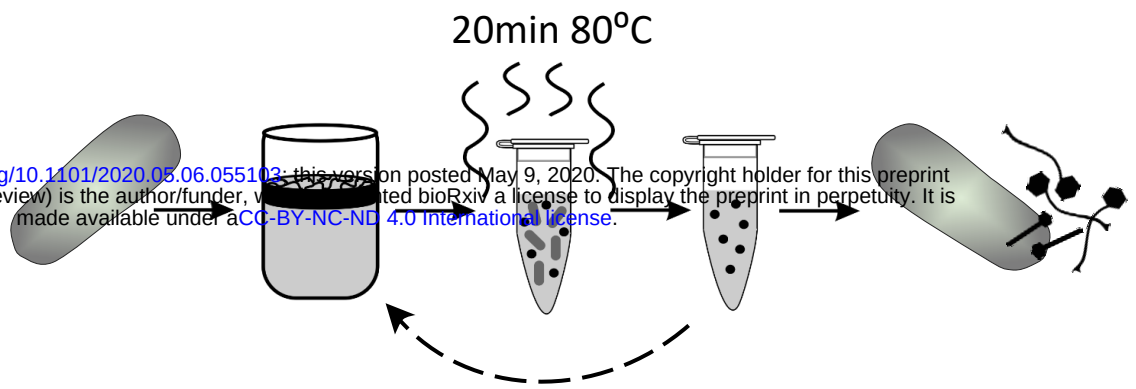
89. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, (2010).

90. Alanjary, M., Steinke, K. & Ziemert, N. AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res.* **47**, W276–W282 (2019).

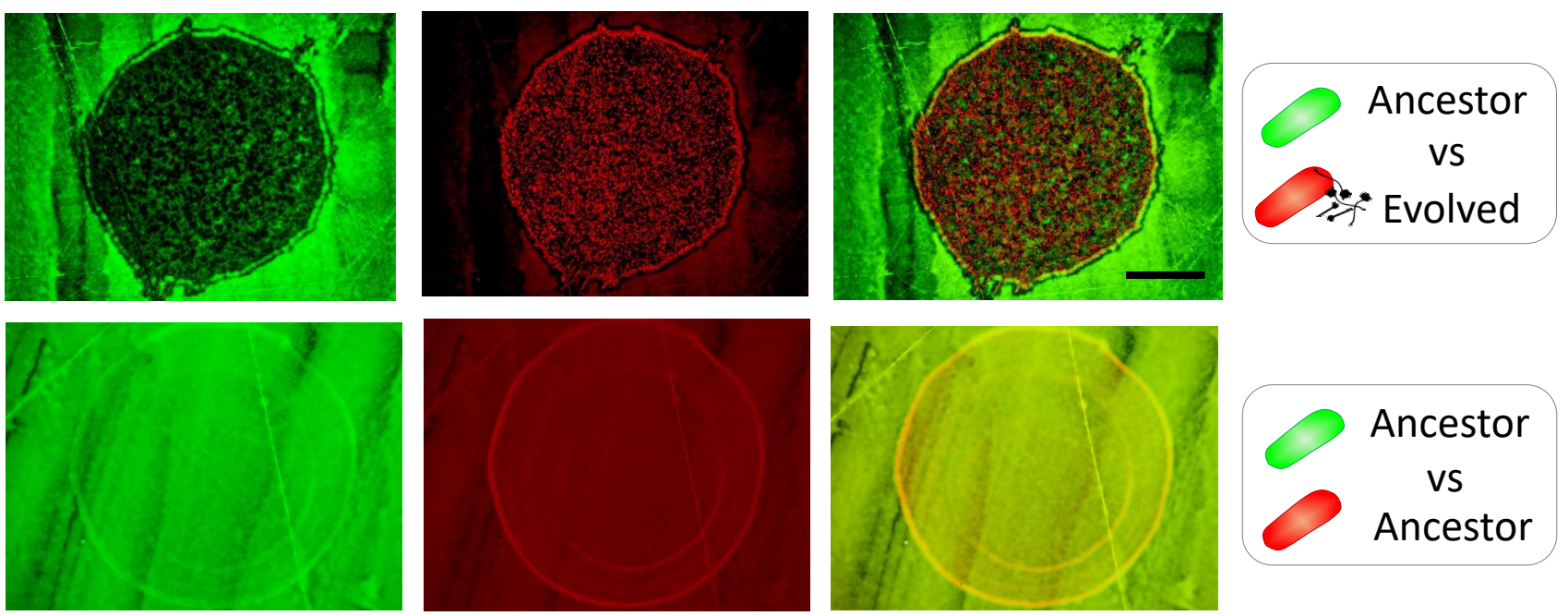
91. Konkol, M. A., Blair, K. M. & Kearns, D. B. Plasmid-encoded ComI inhibits competence in the ancestral 3610 strain of *Bacillus subtilis*. *J. Bacteriol.* **195**, 4085–4093 (2013).

92. Tucker, R. G. Acquisition of thymidylate synthetase activity by a thymine-requiring mutant of *Bacillus subtilis* following infection by the temperate phage ¢3. *J. Gen. Virol* **4**, (1969).

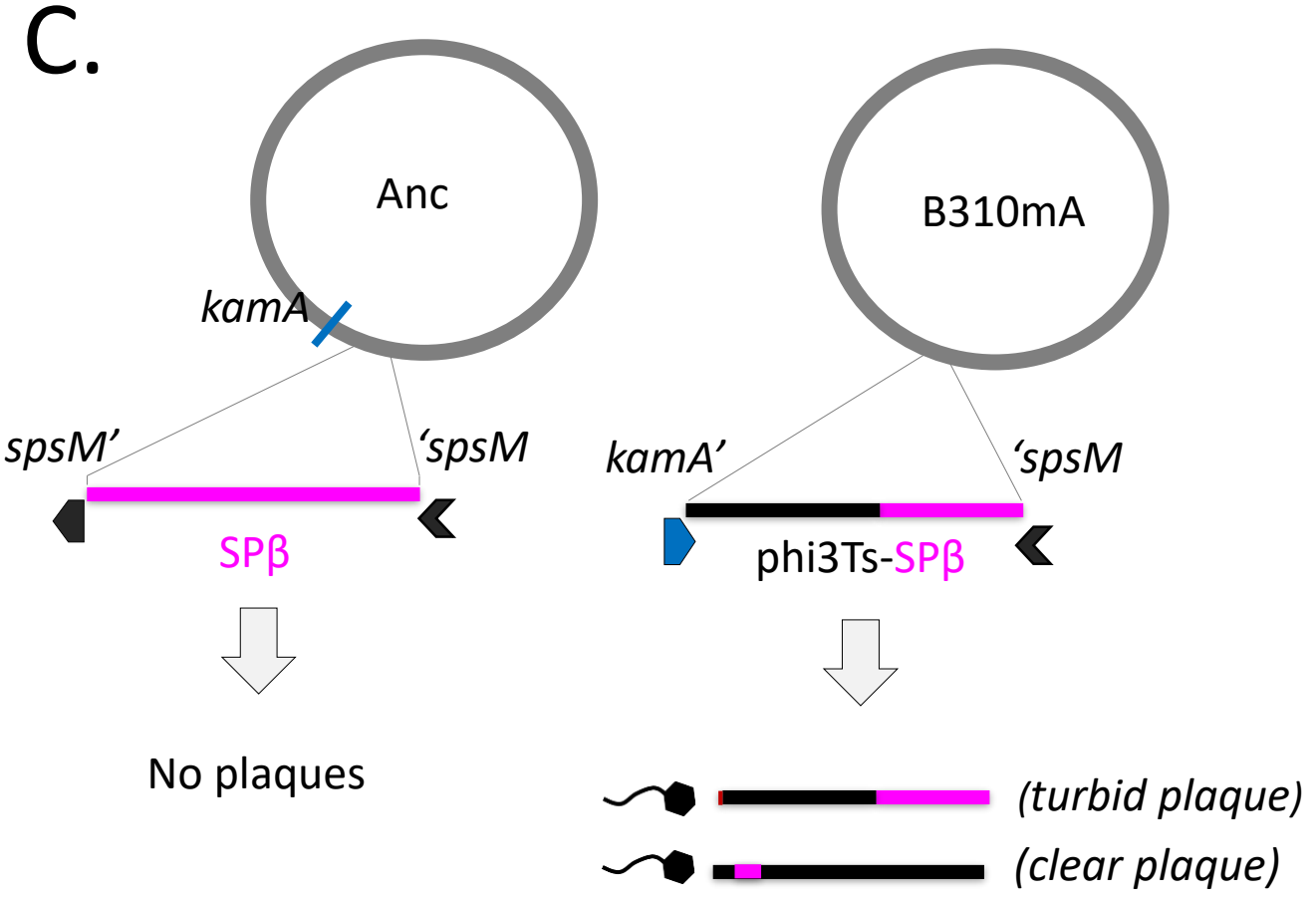
A.



B.



C.



D.

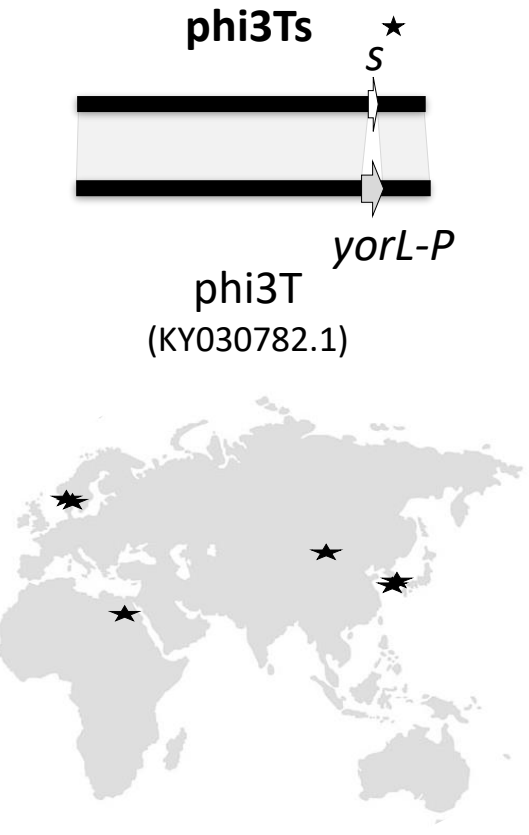


Fig. 2

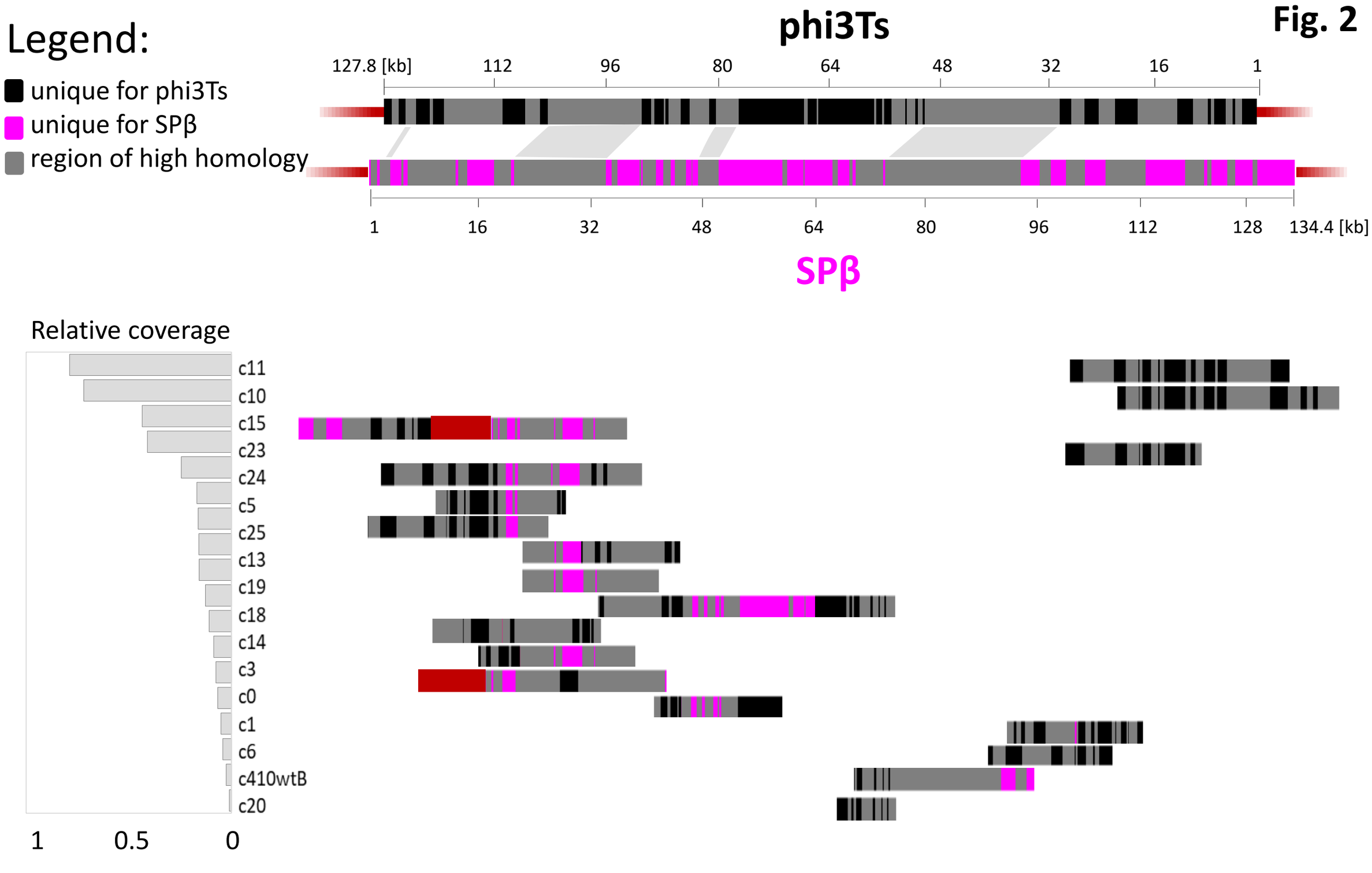

Legend:
unique for phi3Ts
unique for SPβ
region of high homology
phi3Ts
127.8 [kb]
112
96
80
64
48
32
16
1
1
16
32
48
64
80
96
112
128
134.4 [kb]
SPβ
Relative coverage
c11
c10
c15
c23
c24
c5
c25
c13
c19
c18
c14
c3
c0
c1
c6
c410wtB
c20
1
0.5
0

Fig. 3

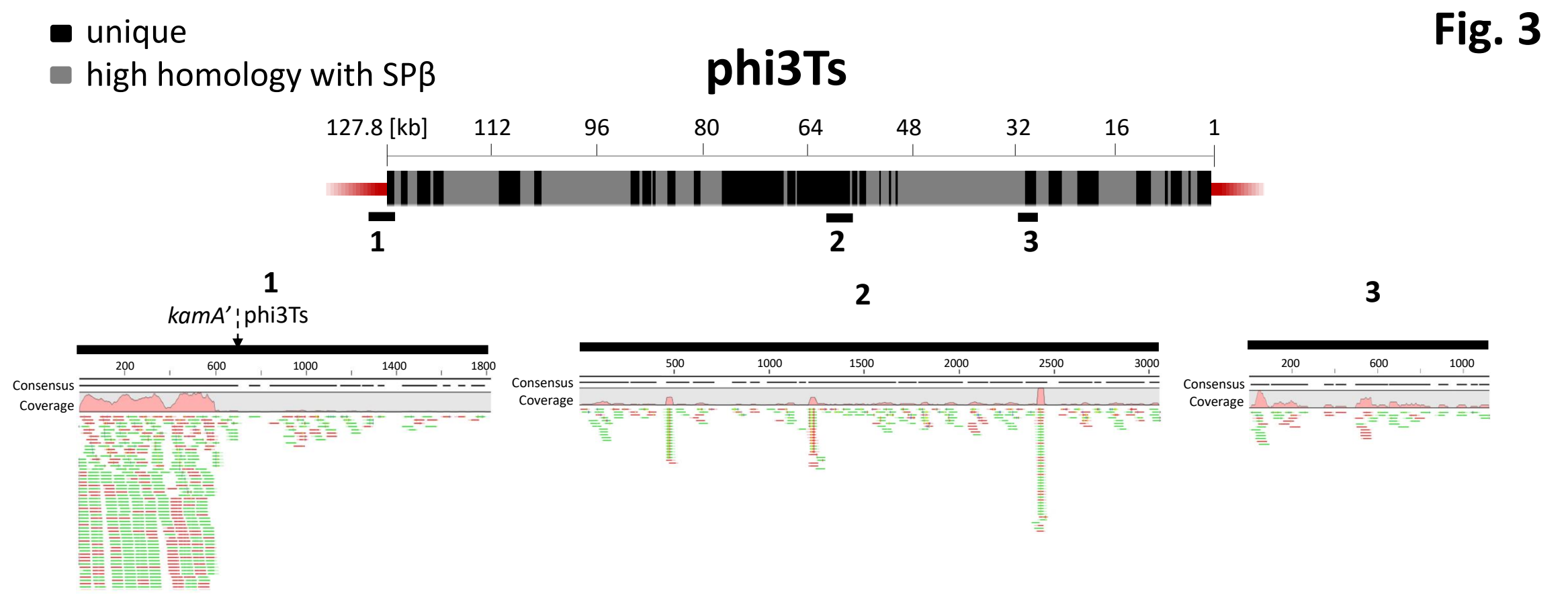

unique
high homology with SPβ
phi3Ts
127.8 [kb]
112
96
80
64
48
32
16
1
1
2
3
1
kamA'
phi3Ts
200
600
1000
1400
1800
Consensus
Coverage
2
500
1000
1500
2000
2500
3000
Consensus
Coverage
3
200
600
1000
Consensus
Coverage

Fig. 4

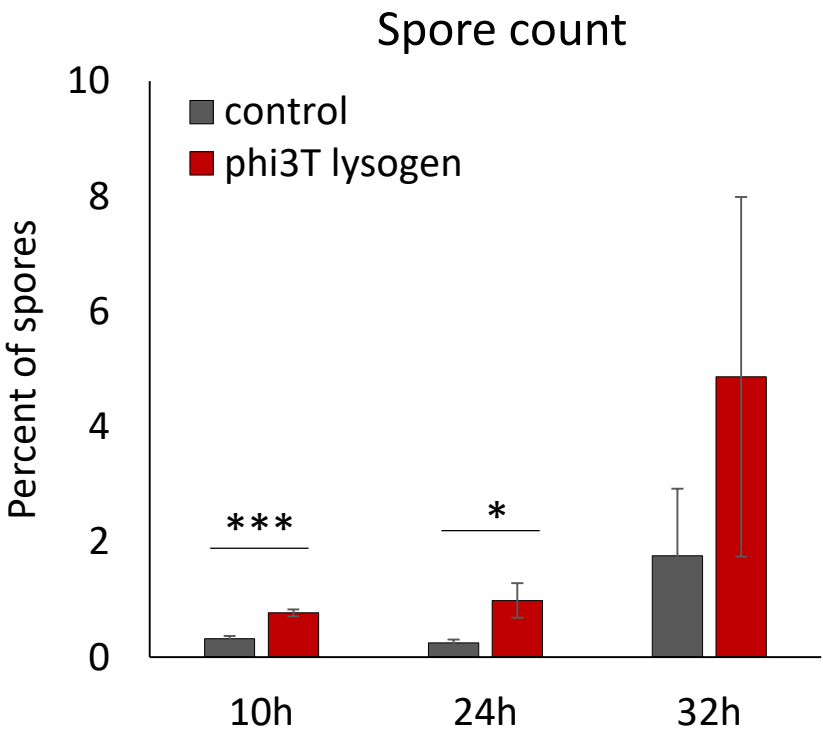
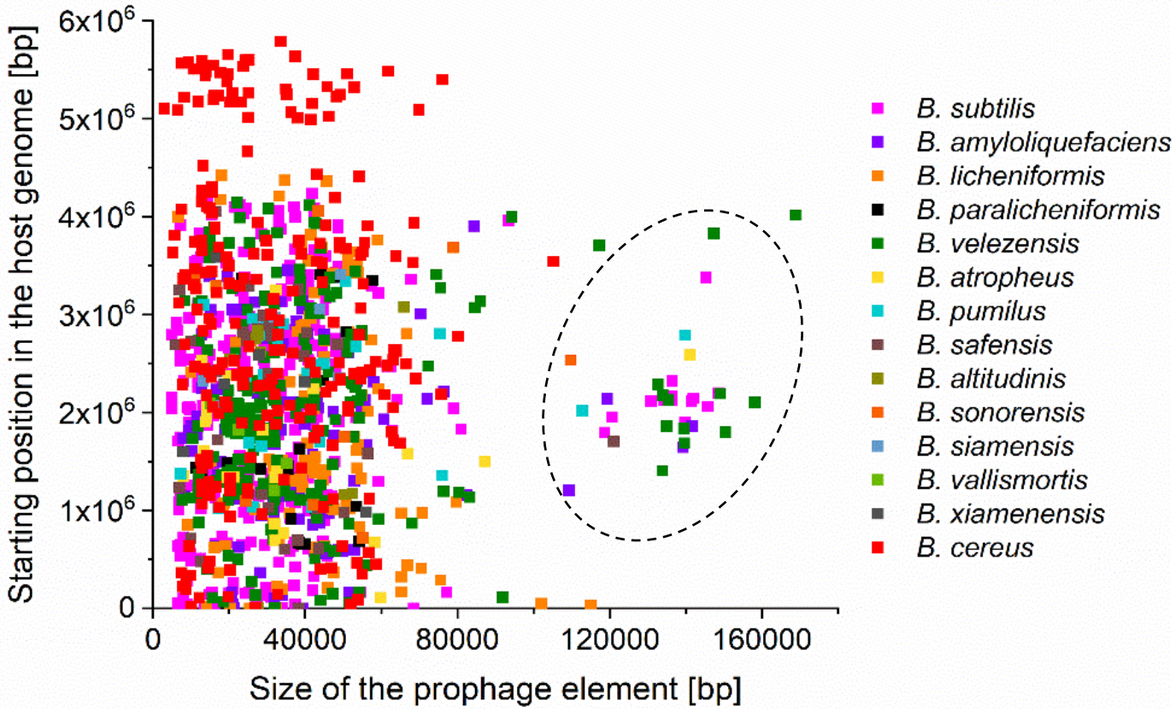
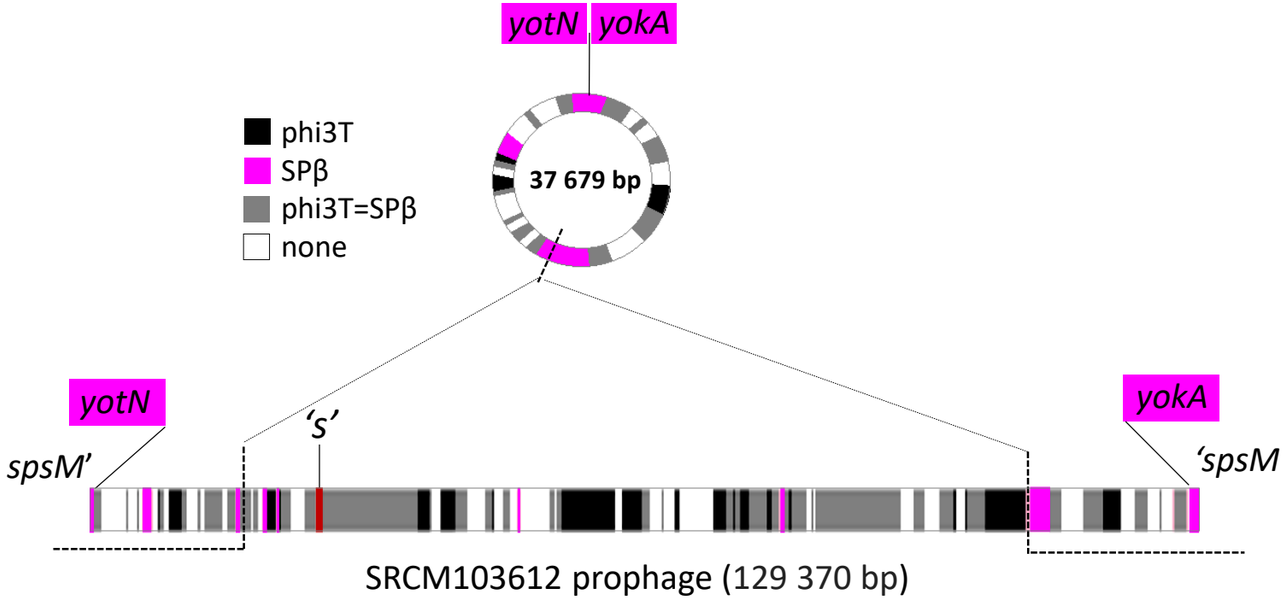

Spore count
control
phi3T lysogen
Percent of spores
10
8
6
4
2
0
***
*
10h
24h
32h

Fig. 5

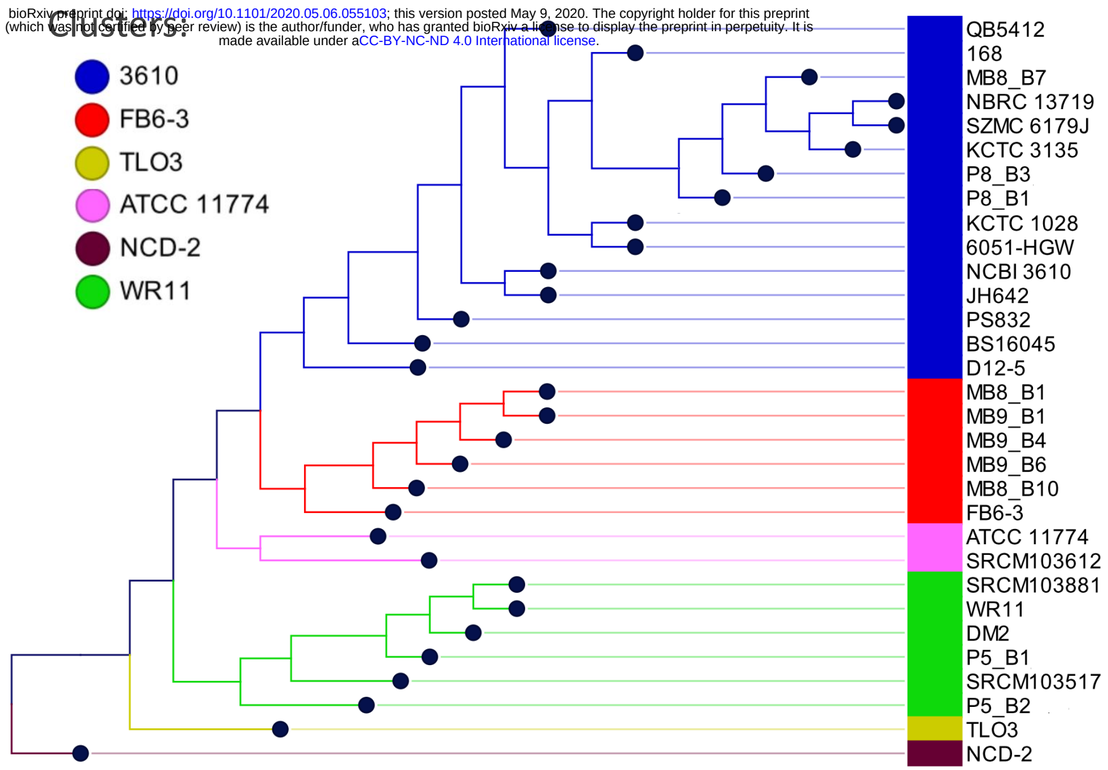
A.


Starting position in the host genome [bp]
6x10^6
5x10^6
4x10^6
3x10^6
2x10^6
1x10^6
0
0
40000
80000
120000
160000
Size of the prophage element [bp]
B. subtilis
B. amyloliquefaciens
B. licheniformis
B. paralicheniformis
B. velezensis
B. atropheus
B. pumilus
B. safensis
B. altitudinis
B. sonorensis
B. siamensis
B. vallismortis
B. xiamenensis
B. cereus

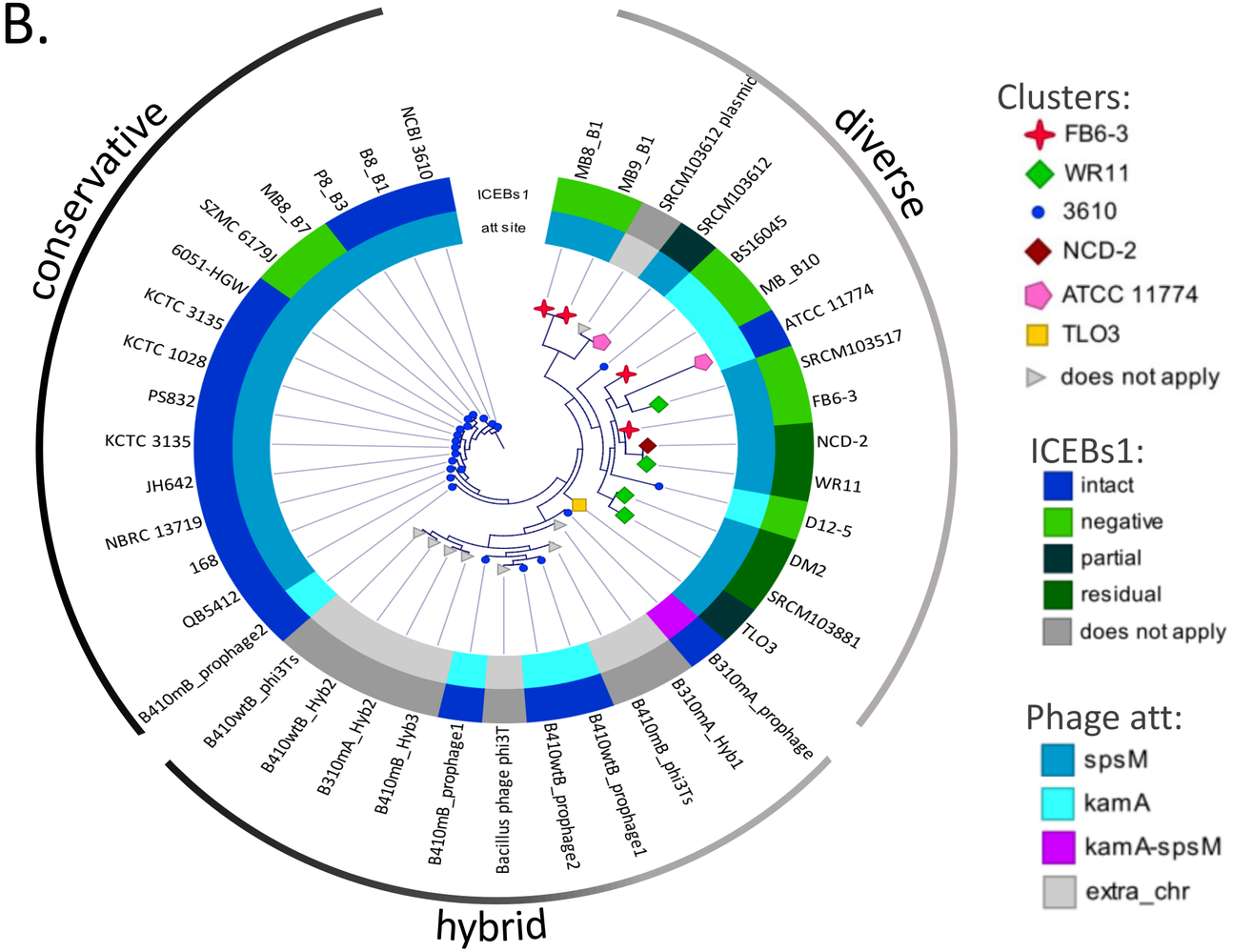
B.


yotN
yokA
phi3T
SPβ
phi3T=SPβ
none
37 679 bp
yotN
'spsM'
's'
yokA
'spsM
SRCM103612 prophage (129 370 bp)

A.



B.



C.