Pervasive prophage recombination occurs during evolution of spore-forming *Bacilli*

Running title: **Evolution of *Bacillus* prophages**

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, 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. 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 indicates that similar diversification of SPβ-like phages takes place in nature. Our work is a stepping stone toward empirical studies on phage evolution, and understanding the eco-evolutionary relationships between bacteria and their phages. By capturing the first steps of new phage evolution, we reveal striking relationship between survival strategy of bacteria and evolution of their phages.

Introduction

Bacteriophages are major regulators of bacteria. Each day, approximately half of bacterial biomass is killed by lytic phages, imposing a constant predator-prey evolutionary arms race [1–3]. 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]. Despite their abundance and relevance, we still understand very little about ecological and evolutionary imprint of prophages. This knowledge gap is particularly paramount for non-pathogenic bacteria, such as widely applied biocontrol and probiotic strains, which often (e.g. *Bacillus* genus) can be extremely rich in prophages [6]. Moreover, between-strains differences in prophage elements in certain *Bacillus* species (e.g. *Bacillus subtilis*) leads to social 'incompatibility', which manifests in strong competitive interactions and physical barriers between bacterial swarms [7, 8]. Currently, we do not understand what promotes prophage diversity within closely related strains.

It is believed that phages predominantly evolve through recombination [9–11], but the rate of horizontal gene transfer (HGT) can vary substantially, depending on phage lifestyle of host phylum [12]. Gene gain/loss by HGT seems to be more common in Firmicutes or Bacteroides phages, compared to Cyanophages, and it is more pronounced in phages with temperate lifestyle [12–14]. Recombination between phages or between phages and their hosts can be either homologous [15, 16] or non-homologous [17], relying on phage- or host-encoded recombinases [15, 18]. It was proposed that gene shuffling regularly occurs between (functional or defective) prophages and phages that co-infect the same host bacterium [15]. Prophages of naturally competent bacteria (e.g. *B. subtilis*) can recombine with foreign phage DNA via transformation [19]. Finally, phages can randomly or specifically incorporate fragments of host chromosomes via generalised [20, 21] or specialised [22, 23] transduction, respectively, thereby contributing to the spread of antibiotic resistance [21] or virulence genes

[24]. Although evidence from comparative genomics indicates frequent phage recombination [25–27], there is little experimental research in this topic. Phage recombination has been experimentally studied using few models, predominantly *Salmonella typhimurium* P22 with *Escherichia coli* lambdoid phages [28, 29], or *Lactococcus lactis* phages sampled across decades from diary production line [25]. Therefore, empirical research on prophage evolution is limited.

Interestingly, certain prophages of *Bacilli* undergo genetic rearrangements upon host development, acting as so-called phage regulatory switches (RSs) [30]. RS can switch between integrated and extrachromosomal forms to modulate reproduction and survival of their hosts [30]. Certain *B. subtilis* strains carry an SPβ prophage or its derivative that integrates into an *spsM* gene, encoding for polysaccharide synthase [31, 32]. 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 [31]. 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 [33–35]. Similar mother cell-specific excisions have been observed in other *Bacillus* sp. [31] and in *Clostridium* sp. [36, 37], 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 [38, 39], and the *rapE–phrE* signalling system involved in sporulation initiation [40, 41]. Furthermore, certain prophages can improve or even restore sporulation in *B. subtilis* [42], 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 [43]. Normally, the lytic cycle of SPβ prophages is blocked by the ICEBs1 (Integrative and Conjugative Element of *B. subtilis*) conjugative element [44], but in the evolved strains, phages are released spontaneously, killing or lysogenizing the original host [43].

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 recombine with indigenous prophages. The new prophages modulate host development, most likely through regulatory genes. Bioinformatic comparison of 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.

Materials and Methods

Strains and cultivation conditions. Supplementary Table S3 describes the bacterial strains used in this study and Supplementary Table S4 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 S5. 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 lysogenization of 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. Strains DTUB251 and DTUB254 were obtained by transforming DK1042 with plasmids pDR111_rapX and pDR111_spsX, respectively. Plasmid pDR111 was linearized using oAD47/pAD48 primers, *rapX* and *spsX* were amplified from B410mB using oAD49/oAD50 and pAD71/oAD72 primers, respectively (see Suppl. table 5). Next, the linearized pDR111 was ligated with *rapX* and with *spsX* using Gibson Assembly® Master Mix (New England BioLabs), resulting in pDR111_rapX and pDR111_spsX. All modifications of DK1042 were verified by PCR followed by Sanger sequencing.

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. Sequencing of

168 ancestor (ancestor of B310mA, B410mB and B410wtB) was described in our previous manuscript [43].

Evolved PY79 strains (presented in Suppl. dataset 2) were obtained as previously described [45]. 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 [46, 47]. The alignment was sorted using SAMtools [48, 49], 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 sequencing 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 sequencing data of PY79 strains [45] and available from B. Burton (briana.burton@wisc.edu). Raw sequencing data of 168 cultivated under near- zero growth conditions[50] 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 [51] at 30°C, 220 rpm, and total CFU and spore counts were analysed after 12, 24 and 36 hours. In sporulation assay performed with DTUB251 and DTUB254, the medium was supplemented with 0.2mM IPTG. 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 [52]. Additional assessment of germination dynamics was performed using real-time brightfield microscopy by inducing spores with L-alanine on agarose pads, as described previously [53]. 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

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/l NaCl, 0.96 g/l $MgSO_4$, 6 g/l Tris-HCl, pH 7.5). 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[54] 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 [55].

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 [56, 57] from fully assembled *Bacillus* genomes available at NCBI. All prophages between 80kB-50 kB, integrated between 1.9-2.3 Mb in the chromosome were selected for further analysis. 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. Duplicates and misassembled genomes (NZ_CP032855.1) were removed.

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 match the sequence splitting *spsM* or *kamA*. 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, using FFT-NS-2 algorithm and auto settings [58], phylogenetic tree was build using FastTree, using Jukes-Cantor model and auto settings [59, 60] and visualized in CLC Main Workbench. Phylogenetic

tree of *B. subtilis* host strains was constructed using autoMLST (https://automlst.ziemertlab.com/) [61] 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 annotation platform.

Statistical analysis

Statistical differences between two experimental groups were identified using two-tailed Student's *t*-tests assuming equal variance. Normality of data distribution was tested by Shapiro-Wilk and Kolmogorov-Smirnov tests. No statistical methods were used to predetermine sample size and the experiments were not randomized.

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 spontaneous release of phage particles into the culture medium (Fig. 1A). Spontaneous phage release occurred in *B. subtilis* 168 grown in rich medium under static, biofilm-promoting conditions after roughly 10 passages which involved selection for spores, but not in the absence of such selection in parallel populations [43]. Recently, similar spontaneous phage release was observed for *B. subtilis* NCBI 3610 grown in minimal medium under shaken-culture conditions after 10 passages through dormant spore stage [62]. Some of these phages resemble indigenous SPβ[63], but they facilitate killing of the ancestor strain [43] (Fig. 1B). To further characterise

the phage-releasing evolved strains, three isolates [43] were subjected to genome sequencing (see Materials and Methods). *De novo* sequencing revealed the presence of an exogenous phi3T-like prophage (KY030782.1; 99.98% sequence identity) [64], which is closely related to SPβ (58% sequence identity) (Fig. 1C, Suppl. Fig. 1A, Suppl. Fig. 2A). We named this phage phi3Ts. The only difference between 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). Although, the 's' fragment shared no homology with phi3T or *B. subtilis* 168 chromosome, it could be found within SPβ-like prophages of six *B. subtilis* strains from different regions around the world (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, sequencing coverage within the described prophage regions was increased several-fold, suggesting augmented replication of hybrid phage DNA (Suppl. Fig. 1B, Suppl. dataset 1).

Comparison between phi3Ts and SPβ revealed that the first lacks putative repressor gene *yonR* as well as immunity-encoding gene *yomJ* what could explain spontaneous release of phage particles (Suppl. Fig. 2B).

Strains lysogenic for hybrid prophages produce virulent hybrid phages

To examine which phage variants were released into the medium [43], phage-containing supernatants were spotted onto the lawn on Δ6 strain (derivative of 168 strain deprived of prophage elements [54]), followed by purification of phages from single plaques and sequencing (see Materials and Methods). Notably, each evolved strain produced a mix of turbid and clear plaques (Suppl. Fig. 1C). Turbid plaques are typical for temperate phages (like SPβ

or phi3T), while clear plaques are usually formed by phages that have lost their lysogenic cycle [65]. Sequencing revealed that the spontaneously produced phages were either phi3Ts or phi3Ts-SPβ hybrids (Fig. 1C, Suppl. Fig. 1A, Fig. 2A). All phages obtained from the turbid plaques could also be identified within genomes of the evolved strains (Fig. 1C, Suppl. Fig. 1A, Fig. 2A). In addition, the genome of Hyb1$^{phi3Ts\text{-}SP\beta}$ (released by B310mA) was extended by a ~1.2 kb fragment of the host chromosome (*yozE*, *yokU*, and part of the *kamA* gene), indicating specialised transduction, a process that occurs when a phage picks up a fragment of host chromosomal DNA in the immediate vicinity of its attachment site (Fig. 1C, Fig. 2A). In contrast to the turbid plaque-creating phages, all phages obtained from clear plaques were phi3Ts-SPβ hybrids, which were not present on the chromosomes of their corresponding producers (Fig. 1C, Suppl. Fig. 1A, Fig. 2A) suggesting they were virulent phi3Ts-SPβ variants. We could also identify a variety of extrachromosomal phage DNA (epDNA) fragments ranging from 10.9 to 66 kb (Fig. 2B, Suppl. dataset 1, Suppl. Table 1). The epDNA was dominated by phi3Ts-SPβ recombinants, in which DNA from the two parental phages was joined at the homologous region (Fig. 2B). We identified 16 unique recombination sites (epDNA and hybrid phages combined), 15 of each were at regions of phi3Ts-SPβ homology, mostly focused around the left arm of phi3Ts/initial part of SPβ in *B. subtilis* chromosome (Fig. 2B., Suppl. Table 1). A single nonhomologous recombination site was found between a region following *yosA* gene of SPβ and unknown phi3Ts gene (Fig. 2B., Suppl. Table 1).

Although none of the hybrid epDNA was identical to sequences of hybrid phages released by the corresponding strains, certain homologous recombination sites in Hyb2$^{phi3Ts\text{-}SP\beta}$ and Hyb3 $^{phi3Ts\text{-}SP\beta}$ were overlapping with these found in epDNA (Fig. 2, Suppl. Table 1). 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. 2B, Suppl. Table 1).

A sporulation selection regime promotes foreign phage invasion

Next, we aimed to identify the source of phi3Ts DNA in the evolved host genomes. First, we repeated the mapping of raw sequencing reads from the *B. subtilis* 168 ancestral genome onto selected unique phi3T regions. Indeed, phi3Ts DNA was present in the ancestor strain at a very low but detectable level (Fig. 3, Suppl. Fig. 3A). On the other hand, phi3T could be clearly detected by mapping of sequencing reads of the evolved strains (Suppl. Fig. 3A).

This result was additionally confirmed by PCR (Suppl. Fig. 4). As expected, the evolved strains B310mA and B410wtB, where integration of phi3Ts into *kamA* locus was confirmed by genome sequencing (Fig. 1C, Suppl. Fig. 1A), were negative for intact *kamA*, while B410mB gave a weak product, despite similar integration of phi3Ts (Suppl. Fig. 1A), what could be explained by incorporation of *kamA* into its epDNA (Fig. 2B, Suppl. Fig. 4A).

The above analysis indicates that low copy number phi3Ts was present in the *B. subtilis* 168 stock from the start. Since *B. subtilis* 168 has been shared among research labs around the world, phi3Ts could also be 'hiding' in culture stocks of other research labs. Accidental detection of such low copy number phage DNA is nearly impossible, because (i) sequencing reads matching phi3Ts would be filtered out during standard mapping pipelines, and (ii) phi3Ts appears to only multiply and manifest itself under specific selection regimes. To check for possible contamination of other *B. subtilis* stocks with phi3Ts, we mapped raw sequencing data available in the NCBI database to the phi3T genome (KY030782.1), but found no evidence of phi3Ts contamination (Suppl. Fig 3B). We also PCR-screened a larger collection of *B. subtilis* 168 stocks from different labs around the world[66] for the presence of phi3Ts and obtained a very strong positive result for *B. subtilis* 168 'Newcastle', suggesting that this strain was lysogenized by phi3Ts, or a very similar prophage (Suppl. Fig. 4B). We

also confirmed that, the Newcastle 168 strain contained the ‘s’ fragment, a unique sequence allowing the phi3Ts phage to be distinguished from the previously sequenced phi3T (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 [45, 50] . A group from the University of Wisconsin-Madison, agreed to share raw sequencing data obtained from 12 evolved single isolates. We did not find any mutations within prophage regions (Suppl. dataset 2) nor phi3T-specific DNA fragments in the raw sequencing data (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) [50]. 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 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 [43] 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

sequencing (Suppl. Fig. 3D, Suppl. Fig. 6A) or by PCR (Suppl. Fig. 6B). To our surprise, lytic activity (Suppl. Fig. 7A) and the release of phages (Suppl. Fig. 7) 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). This experiment again demonstrated unsupervised outspread of phi3Ts, taking place solely in populations subjected to sporulation/spore revival selection regime. The latter either facilitates contamination by phi3Ts or propagation of this phage from levels that are beyond PCR or sequencing detection limits.

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 (http://subtiwiki.uni-goettingen.de), we hypothesised that *kamA* may encode a product that is metabolically costly and/or toxic for the bacterium, hence strains lysogenized by phi3Ts/phi3Ts-SPβ hybrid prophage ~~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).

Next, we examined the sporulation and spore revival dynamics of *B. subtilis* 3610 lyzogenized by phi3T, 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 results suggest that the spread of phi3Ts might be partly driven by host benefits from the regulatory arsenal associated with phi3Ts and its hybrids.

Notably, sporulation regulators have been previously linked to mobile genetic elements (MGEs) in this species [67–69]. 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. 2A, Suppl. Table 2) sharing high amino acid sequence identify with RapA (unique for phi3Ts) that is known to modulate sporulation timing [70]. 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. 2A). Indeed, overexpression of *spsX* from ectopic locus resulted in small but significant increase in percent of early spores as well as faster germination of *B. subtilis* NCBI 3610 (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) [71]. Therefore, phi3T/phi3Ts may encode proteins that influence the *B. subtilis* life cycle during sporulation and spore revival.

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 analysed prophage elements within the *B. subtilis* clade including *B. cereus* for comparison of more distant species (see Materials and Methods) (Suppl. dataset 3). We immediately identified a cluster of rather large (>100 kb) prophages integrated

close to the replication terminus, just like phi3Ts, SPβ or phi3Ts-SPβ hybrids. These large prophages were found in 78 out of a total of 320 genomes, mainly (86%; 67/78) within representatives of *B. subtilis* (34/78), *B. velezensis* (17/78). *B. amyloliquefaciens* (10/78) and *B. licheniformis* (6/78) species (Fig. 5A, Suppl. dataset 3, Suppl. Fig. 12).

We further identified 23 strains in which large prophages 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 at exactly like observed in the our evolved strains lysogenized by hybrid phages (Suppl. dataset 3). In the remaining Bacillus strains, the large prophages were mostly integrated close to sporulation-related genes (Suppl. dataset 3). 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 access the natural diversity of large SPβ-like prophages, we collected *Bacillus* sp. genomes carrying a large prophage splitting *spsM* or *kamA*. Notably, such prophage was present in roughly 39% of all unique *B. subtilis* genomes available (28 out of 72). We compared the phylogenetic tree obtained for these host 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, all residing within closely related *B. subtilis* strains (Fig. 6AB). The 'hybrid' clade comprised phi3T, phi3Ts and all phi3Ts-SPβ hybrids that evolved in the above described experiments (Fig. 6B). Within the 'diverse' clade the prophage relatedness did not match the phylogenetic relatedness of the hosts (Fig. 6AB). Here, we found

that among *B. subtilis* isolates from the same soil sample below the mushroom [72], 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). Notably, all members of the 'conservative' clade carried an intact copy of ICEBs1 that was shown to block the SPβ lytic cycle [44], while this element is missing in all members of the 'diverse' clade (Fig. 6B). Like phi3Ts and phi3Ts-SPβ hybrid phages, several representatives of the 'diverse' cluster lacked the putative repressor gene *yonR*, which was present in all members of the 'conservative' cluster (Fig. 6B, Suppl. dataset 3). 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 and facilitated by loss of repressor gene through homologous recombination.

Discussion

The importance of phages in the ecology and evolution of bacteria is indisputable. Genome comparison suggests that prophage elements undergo pervasive domestication within their hosts that gradually lose the ability to reproduce via the lytic cycle [73]. Our current work demonstrates an opposite scenario, where after a prolonged sporulation/spore revival selection regime, a latent prophage of *B. subtilis* (SPβ) [43, 44] regains its lytic reproductive cycle via recombination with 'foreign' phage DNA (phi3Ts). Prophage recombination did not occur in the absence of sporulation/spore revival selection regime, as evident from our previous experimental evolution studies [74, 75], as well as this work (Suppl. Fig. 7).

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 [76, 77]. 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 [76]. Likewise, lytic variants of P22 spontaneously form upon purine starvation of the *Salmonella typhimurium* host [77]. It was later realized that P22 coexists with its host in so called carrier state, when extrachromosomal phage is asymmetrically inherited by only one of the sibling cells after cell division, resulting in coexistence of resistant and susceptible host populations [78, 79]. Similar mechanism could maintain low amount of phi3Ts within *B. subtilis* population.

Identification of phi3Ts and phi3Ts-SPβ hybrids within *B. subtilis* populations evolved under sporulation/spore-revival selection regime, sheds new light on our previous work, that changes within SPβ regions specifically promote the spread of biofilm-deficient strains in the population [43]. Present study reveals that while biofilm-proficient strain (B410wtB) still carries intact phi3Ts and SPβ at different loci, biofilm-deficient strains (B310mA and and B410mB) are lysogenized by phi3Ts-SPβ hybrids, what also results in genome deletion (Fig. 1C, Suppl. Fig. 1A), offering a potential explanation for fitness advantage of these strains. It remains to be tested whether biofilm-deficiency is linked to the likelihood of phi3Ts-SPβ recombination.

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 spread of phi3Ts within the population, infection of phi3Ts-susceptible 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 [80, 81]; (b) since phi3T lysogeny (KY030782.1; 99.98% sequence identity with phi3Ts) [64] 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. We show that *spsX* gene could be one possible candidate modulating sporulation and spore revival timing. Another suspect was *rap-phr* cassette (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 [67–69], however we could not detect significant effects of *rapX* overexpression on sporulation/spore revival timing, contrary to recently hypothesized function of this gene in phi3T [82]. Phi3Ts phage genes (e.g. *sspC*) could modulate the production of resistant and viable spores [71] and/or reduce sporulation failure and premature germination [83]. In addition, the spore revival traits of lysogens may also be indirectly affected by the modulation of sporulation timing [53]. 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 and lysogenizing these species [31, 37, 42, 84–86].

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 lysogenized hosts [64, 87] or biosynthetic gene clusters [88, 89], and functional genes related to host dormancy [90, 91]. 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 [92] and in non-homologous end-joining repair [93] . It appears that the absence of other mobile genetic elements (e.g. ICEBs1) constraining the phage lytic cycle [44] may also correlate with a higher level of phage diversification. It is possible that such phage

recombination could be facilitated by regulatory excision of SPβ from the chromosome in the sporulating mother cell [31, 94].

Our work shows how large SPβ-like prophages of *Bacilli* pervasively recombine during sporulation, providing new model system to study bacterial evolution in which phages serve as an evolutionary driving force. These observations shed new light on the interplay between bacteria and their phages; while temperate phages commonly undergo domestication [73], they may easily regain genetic mobility by recombination with other phages, thereby altering the physiology, social interactions and evolution of their hosts. The crucial impact of prophage elements on ecological interactions within closely related strains has already been demonstrated for other species [7, 76, 95, 96]. Herein, we showed that such antagonism emerges during the early steps of phage diversification, which suggesting that speciation of prophage elements may be the first step toward speciation of host bacteria.

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.

Competing interests:

The authors declare that there are no competing financial interests in relation to the work described.

Authors contributions:

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

References:

1. Knowles B, Silveira CB, Bailey BA, Barott K, Cantu VA, Cobián-Güemes AG, et al. Lytic to temperate switching of viral communities. *Nature* 2016; 531:466–470.

2. Koskella B, Brockhurst MA. Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol Rev* 2014; 38:916–931.

3. Azam AH, 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* 2019; 103:2121–2131.

4. Howard-Varona C, Hargreaves KR, Abedon ST, Sullivan MB. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *ISME J* 2017; 11:1511–1520.

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

6. Kim MS, Bae JW. Lysogeny is prevalent and widely distributed in the murine gut microbiota. *ISME J* 2018; 12:1127–1141.

7. Štefanič P, Kraigher B, Lyons NA, Kolter R, Mandić-Mulec I. Kin discrimination between sympatric *Bacillus subtilis* isolates. *Proc Natl Acad Sci USA* 2015; 112:14042–14047.

8. Lyons NA, Kraigher B, Štefanič P, Mandić-Mulec I, Kolter R. A combinatorial kin discrimination system in *Bacillus subtilis*. *Curr Biol* 2016; 26:733–742.

9. Canchaya C, Proux C, Fournous G, Bruttin A, Brüssow H. Prophage genomics. *Microbiol Mol Biol Rev* 2003; 67:238–276.

10. Bérard S, Chateau A, Pompidor N, Guertin P, Bergeron A, Swenson KM. Aligning the unalignable: bacteriophage whole genome alignments. *BMC Bioinformatics* 2016; 17:30.

11. Botstein D. A theory of modular evolution for bacteriophages. *Ann N Y Acad Sci* 1980; 354:484–491.

12. Mavrich TN, Hatfull GF. Bacteriophage evolution differs by host, lifestyle and genome. *Nat Microbiol* 2017; 2:17112.

13. Marston MF, Martiny JBH. Genomic diversification of marine cyanophages into stable ecotypes. *Environ Microbiol* 2016; 18:4240–4253.

14. Marston MF, Amrich CG. Recombination and microdiversity in coastal marine cyanophages. *Environ Microbiol* 2009; 11:2893–2903.

15. De Paepe M, Hutinet G, Son O, Amarir-Bouhram J, Schbath S, Petit MA. Temperate phages acquire DNA from defective prophages by relaxed homologous recombination: The role of Rad52-like recombinases. *PLoS Genet* 2014; 10.

16. Swenson KM, Guertin P, Deschênes H, Bergeron A. Reconstructing the modular recombination history of *Staphylococcus aureus* phages. *BMC Bioinformatics* 2013; 14 Suppl 15:S17.

17. Morris P, Marinelli LJ, Jacobs-Sera D, Hendrix RW, Hatfull GF. Genomic characterization of mycobacteriophage giles: Evidence for phage acquisition of host DNA by illegitimate recombination. *J Bacteriol* 2008; 190:2172–2182.

18. Bobay L-M, Touchon M, Rocha EPC. Manipulating or superseding host recombination functions: A dilemma that shapes phage evolvability. *PLoS Genet* 2013; 9:e1003825.

19. Spancake GA, Hemphill HE, Fink PS. Genome organization of Spbetac2 bacteriophage carrying the *thyP3* gene. *J Bacteriol* 1984; 157:428–34.

20. Fillol-Salom A, Alsaadi A, Sousa JAM de, Zhong L, Foster KR, Rocha EPC, et al. Bacteriophages benefit from generalized transduction. *PLOS Pathog* 2019; 15:e1007888.

21. Uchiyama J, Takemura-Uchiyama I, Sakaguchi Y, Gamoh K, Kato SI, Daibata M, et al. Intragenus generalized transduction in *Staphylococcus spp*. by a novel giant phage. *ISME J* 2014; 8:1949–1952.

22. Morse ML, Lederberg EM, Lederberg J. Transduction in *Escherichia coli* K-12. *Genetics* 1956; 41:142–56.

23. 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* 1976; 73: 3238–3242.

24. Penadés JR, Chen J, Quiles-Puchalt N, Carpena N, Novick RP. Bacteriophage-

mediated spread of bacterial virulence genes. *Curr Opin Microbiol*. 2015; 23: 171–178.

25. Kupczok A, Neve H, Huang KD, Hoeppner MP, Heller KJ, Franz CMAP, et al. Rates of mutation and recombination in *Siphoviridae* phage genome evolution over three decades. *Mol Biol Evol* 2018; 35:1147–1159.

26. Yahara K, Lehours P, Vale FF. Analysis of genetic recombination and the pan-genome of a highly recombinogenic bacteriophage species. *Microb genomics* 2019; 5. e000282.

27. Hatfull GF, Hendrix RW. Bacteriophages and their genomes. *Curr Opin Virol* 2011; 1:298–303.

28. Yamamoto N, Wohlhieter JA, Gemski P, Baron LS. λimmP22dis: A hybrid of coliphage λ with both immunity regions of Salmonella phage P22. *Mol Gen Genet MGG* ; 166:233–243.

29. Botstein D, Herskowitz I. Properties of hybrids between Salmonella phage P22 and coliphage λ. *Nature* 1974; 251:584–589.

30. Feiner R, Argov T, Rabinovich L, Sigal N, Borovok I, Herskovits AA. A new perspective on lysogeny: prophages as active regulatory switches of bacteria. *Nat Rev Microbiol* 2015; 13:641–650.

31. Abe K, Kawano Y, Iwamoto K, Arai K, Maruyama Y, Eichenberger P, et al. Developmentally-regulated excision of the SPβ prophage reconstitutes a gene required for spore envelope maturation in *Bacillus subtilis* . *PLoS Genet* 2014; 10:e1004636.

32. Sanchez-Vizuete P, Le Coq D, Bridier A, Herry J-M, Aymerich S, Briandet R. Identification of *ypqP* as a new *Bacillus subtilis* biofilm determinant that mediates the

protection of *Staphylococcus aureus* against antimicrobial agents in mixed-species communities. *Appl Environ Microbiol* 2015; 81:109–18.

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

34. Stragier P, Kunkel B, Kroos L, Losick R. Chromosomal rearrangement generating a composite gene for a developmental transcription factor. *Science* 1989; 243:507–512.

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

36. Pyne ME, Liu X, Moo-Young M, Chung DA, Chou CP. Genome-directed analysis of prophage excision, host defence systems, and central fermentative metabolism in *Clostridium pasteurianum*. *Sci Rep* 2016; 6:26228.

37. Haraldsen JD, Sonenshein AL. Efficient sporulation in *Clostridium difficile* requires disruption of the σK gene. *Mol Microbiol* 2003; 48:811–821.

38. Sohail A, Hayes CS, Divvela P, Setlow P, Bhagwat AS. Protection of DNA by α/β-type small, acid-soluble proteins from *Bacillus subtilis* spores against cytosine deamination. *Biochemistry* 2002; 41:11325–11330.

39. Ki SL, Bumbaca D, Kosman J, Setlow P, Jedrzejas MJ. Structure of a protein-DNA complex essential for DNA protection in spores of *Bacillus species*. *Proc Natl Acad Sci U S A* 2008; 105:2806–2811.

40. Jiang M, Grau R, Perego M. Differential processing of propeptide inhibitors of rap phosphatases in *Bacillus subtilis*. *J Bacteriol* 2000; 182:303–310.

41. Serra CR, Earl AM, Barbosa TM, Kolter R, Henriques AO. Sporulation during growth in a gut isolate of *Bacillus subtilis*. *J Bacteriol* 2014; 196:4184–4196.

42. Silver-Mysliwiec TH, Bramucci MG. Bacteriophage-enhanced sporulation: Comparison of spore-converting bacteriophages PMB12 and SP10. *J Bacteriol* 1990; 172:1948–1953.

43. Martin M, Dragoš A, Hölscher T, Maróti G, Bálint B, Westermann M, et al. *De novo* evolved interference competition promotes the spread of biofilm defectors. *Nat Commun* 2017; 8:15127.

44. Eleina England by M, Bell Professor of Biology SP. Effects of cell growth and a mobile genetic element on propagation of the phages SP16 and SP-beta in *Bacillus subtilis*. 2014. Massachusetts Institute of Technology, Department of Biology, https://dspace.mit.edu/handle/1721.1/87468

45. Bose B, Reed SE, Besprozvannaya M, Burton BM. Missense mutations allow a sequence-blind mutant of SpoIIIE to successfully translocate chromosomes during sporulation. *PLoS One* 2016; 11:e0148365.

46. Langmead B, Wilks C, Antonescu V, Charles R. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* 2019; 35:421–432.

47. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012; 9:357–359.

48. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; 25:2078–2079.

49. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data.

Bioinformatics 2011; 27:2987–2993.

50. Overkamp W, Ercan O, Herber M, van Maris AJA, Kleerebezem M, Kuipers OP. Physiological and cell morphology adaptation of *Bacillus subtilis* at near-zero specific growth rates: A transcriptome analysis. *Environ Microbiol* 2015; 17:346–363.

51. Branda SS, Gonzalez-Pastor JE, Ben-Yehuda S, Losick R, Kolter R. Fruiting body formation by *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 2001; 98:11621–11626.

52. Harwood CR, Cutting SM. Molecular biological methods for Bacillus. 1990.

Edited by C. R. Harwood, Wiley, Chichester, United Kingdom

53. Mutlu A, Trauth S, Ziesack M, Nagler K, Bergeest JP, Rohr K, et al. Phenotypic memory in *Bacillus subtilis* links dormancy entry and exit by a spore quantity-quality tradeoff. *Nat Commun* 2018; 9:69.

54. Westers H, Dorenbos R, Van Dijl JM, Kabel J, Flanagan T, Devine KM, et al. Genome engineering reveals large dispensable regions in *Bacillus subtilis*. *Mol Biol Evol* 2003; 20:2076–2090.

55. 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* 2016; 37:150–157.

56. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: A fast phage search tool. *Nucleic Acids Res* 2011; 39: 347-352.

57. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 2016; 44:W16–W21.

58. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* 2013; 30:772–780.

59. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009; 26:1641–1650.

60. Price MN, Dehal PS, Arkin AP. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* 2010; 5: e9490.

61. Alanjary M, Steinke K, Ziemert N. AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res* 2019; 47:276–282.

62. Gallegos-Monterrosa R, Christensen MN, Barchewitz T, Köppenhöfer S, Priyadarshini B, Bálint B, et al. Impact of Rap-Phr system abundance on adaptation of *Bacillus subtilis*. *bioRxiv* 2020; 2020.09.01.278184.

63. Warner FD, Kitos GA, Romano MP, D H Ernest Hemphill AN, Kitos M P Romano GA, Hemphill HE. Characterization of SPP: a temperate bacteriophage from *Bacillus subtilis* 168M. *Can J Microbiol* 1977; 23:45-51.

64. Erez Z, Steinberger-Levy I, Shamir M, Doron S, Stokar-Avihail A, Peleg Y, et al. Communication between viruses guides lysis-lysogeny decisions. *Nature* 2017; 541:488–493.

65. Dennehy JJ. Bacteriophage ecology: population growth, evolution, and impact of bacterial viruses. *Advances in Molecular and Cellular Microbiology* 2008; 15. Edited by Stephen T. Abedon. Cambridge and New York: Cambridge University Press

66. Gallegos-Monterrosa R, Mhatre E, Kovács AT. Specific *Bacillus subtilis* 168 variants form biofilms on nutrient-rich medium. *Microbiology* 2016; 162:1922–1932.

67. Omer Bendori S, Pollak S, Hizi D, Eldar A. The RapP-PhrP quorum-sensing system of *Bacillus subtilis* strain NCIB 3610 affects biofilm formation through multiple targets, due to an atypical signal-insensitive allele of RapP. *J Bacteriol* 2015; 197:592–602.

68. Singh PK, Ramachandran G, Ramos-Ruiz R, Peiró-Pastor R, Abia D, Wu LJ, et al. Mobility of the native *Bacillus subtilis* conjugative plasmid pLS20 is regulated by intercellular signaling. *PLoS Genet* 2013; 9:e1003892.

69. Auchtung JM, Lee CA, Monson RE, Lehman AP, Grossman AD. Regulation of a *Bacillus subtilis* mobile genetic element by intercellular signaling and the global DNA damage response. *Proc Natl Acad Sci U S A* 2005; 102:12554–12559.

70. Perego M, Hoch JA. Cell-cell communication regulates the effects of protein aspartate phosphatases on the phosphorelay controlling development in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 1996; 93:1549–53.

71. Tovar-Rojo F, Setlow P. Effects of mutant small, acid-soluble spore proteins from *Bacillus subtilis* on DNA in vivo and in vitro. *J Bacteriol* 1991; 173:4827–4835.

72. Kiesewalter HT, Lozano-Andrade CN, Maróti G, Snyder D, Cooper VS, Jørgensen TS, et al. Complete genome sequences of 13 *Bacillus subtilis* soil isolates for studying secondary metabolite diversity . *Microbiol Resour Announc* 2020; 9: e01406-19.

73. Bobay L-M, Touchon M, Rocha EPC. Pervasive domestication of defective prophages by bacteria. *Proc Natl Acad Sci* 2014; 111:12127–12132.

74. Dragoš A, Martin M, Falcón García C, Kricks L, Pausch P, Heimerl T, et al. Collapse of genetic division of labour and evolution of autonomy in pellicle biofilms. *Nat Microbiol* 2018; 3:1451–1460.

75. Dragoš A, Lakshmanan N, Martin M, Horváth B, Maróti G, García CF, et al.

Evolution of exploitative interactions during diversification in *Bacillus subtilis* biofilms. *FEMS Microbiol Ecol* 2018; 94:fix155.

76. Song S, Guo Y, Kim J-S, Wang X, Wood TK. Phages mediate bacterial self-recognition. *Cell Rep* 2019; 27:737-749.e4.

77. Downs DM, Roth JR. A novel P22 prophage in *Salmonella typhimurium*. *Genetics* 1987; 117:367–80.

78. Cenens W, Mebrhatu MT, Makumi A, Ceyssens P-J, Lavigne R, Van Houdt R, et al. Expression of a novel P22 ORF an gene reveals the phage carrier state in *Salmonella typhimurium*. *PLoS Genet* 2013; 9:e1003269.

79. Cenens W, Makumi A, Govers SK, Lavigne R, Aertsen A. Viral transmission dynamics at single-cell resolution reveal transiently immune subpopulations caused by a carrier state association. *PLOS Genet* 2015; 11:e1005770.

80. Moreno F. On the trapping of phage genomes in spores of *Bacillus subtilis* 168 reciprocal exclusion of phages φ29 and φe during outgrowth of spores. *Virology* 1979; 93:357–368.

81. Sonenshein AL. Trapping of unreplicated phage DNA into spores of *Bacillus subtilis* and its stabilization against damage by 32P decay. *Virology* 1970; 42:488–495.

82. Bernard C, Li Y, Lopez P, Bapteste E. Beyond arbitrium: identification of a second communication system in Bacillus phage phi3T that may regulate host defense mechanisms. *ISME J* 2020; 1–5.

83. Ramírez-Guadiana FH, Meeske AJ, Wang X, Rodrigues CDA, Rudner DZ. The *Bacillus subtilis* germinant receptor GerA triggers premature germination in response to morphological defects during sporulation. *Mol Microbiol* 2017; 105:689–704.

84. Lewis RJ, Brannigan JA, Offen WA, Smith I, Wilkinson AJ. An evolutionary link between sporulation and prophage induction in the structure of a repressor:anti-repressor complex. *J Mol Biol* 1998; 283:907–912.

85. Sonenshein AL. Bacteriophages: How bacterial spores capture and protect phage DNA. *Curr Biol* . 2006; 16:14-16.

86. Castilla-Llorente V, Muñoz-Espín D, Villar L, Salas M, Meijer WJJ. Spo0A, the key transcriptional regulator for entrance into sporulation, is an inhibitor of DNA replication. *EMBO J* 2006; 25:3890–3899.

87. Babel H, Naranjo-Meneses P, Trauth S, Schulmeister S, Malengo G, Sourjik V, et al. Ratiometric population sensing by a pump-probe signaling system in *Bacillus subtilis*. *Nat Commun* 2020; 11:1–13.

88. Paik SH, Chakicherla A, Hansen JN. Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168. *J Biol Chem* 1998; 273:23134–23142.

89. Denham EL, Piersma S, Rinket M, Reilman E, de Goffau MC, van Dijl JM. Differential expression of a prophage-encoded glycocin and its immunity protein suggests a mutualistic strategy of a phage and its host. *Sci Rep* 2019; 9:2845.

90. Lazarevic V, Dusterhoft A, Soldo B, Hilbert H, Mauel C, Karamata D. Nucleotide sequence of the *Bacillus subtilis* temperate bacteriophage SPβc2. *Microbiology* 1999; 145:1055–1067.

91. Moeller R, Setlow P, Reitz G, Nicholson WL. Roles of small, acid-soluble spore proteins and core water content in survival of *Bacillus subtilis* spores exposed to

environmental solar UV radiation. *Appl Environ Microbiol* 2009; 75:5202–8.

92. Schultz D, Wolynes PG, Jacob E Ben, Onuchic JN. Deciding fate in adverse times: Sporulation and competence in *Bacillus subtilis*. *Proc Natl Acad Sci* 2009; 106: 21027-21034.

93. de Vega M. The Minimal *Bacillus subtilis* nonhomologous end joining repair machinery. *PLoS One* 2013; 8:e64232.

94. 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 subtili*s. *Nucleic Acids Res* 2017; 45:6669–6683.

95. Lyons NA, Kraigher B, Štefanič P, Mandic-Muleć I, Kolter R. A combinatorial kin discrimination system in *Bacillus subtilis*. *Curr Biol* 2016; 26:733–42.

96. Dey A, Vassallo CN, Conklin AC, Pathak DT, Troselj V, Wall D. Sibling rivalry in Myxococcus xanthus is mediated by kin recognition and a polyploid prophage. *J Bacteriol* 2016; 198:994–1004.

Figure legends:

Figure 1. Changes within *B. subtilis* prophage sequence and integration site observed after prolonged sporulation selection regime. A) Experimental evolution with sporulation selection regime leads to spontaneous release of phage particles by the evolved strains[43]. B). Overnight culture of evolved *B. subtilis* strain B410mB (*amyE::mKate,* shown in red) was diluted 100× and spotted on the lawn of undiluted *B. subtilis* ancestor strain (*amyE::gfp*, shown in green), resulting in a clearance zone, and growth of B410mB in that zone. The same experiment was

performed using 100x diluted culture of ancestor strain (*amyE::mKate*) on a lawn of undiluted ancestor (*amyE::gfp*), as control.Scale bar=2.5mm. C) Schematic representation of genome rearrangements in one of the phage-releasing evolved strains (B310mA), compared to the ancestor (Anc). The evolved strains carry a hybrid prophage phi3Ts-SPβ. Fragments of phi3Ts are shown in black, while fragments of SPβ are shown in pink. Below, schematic representation of phage genomes, spontaneously released by B310mA. D) Schematic comparison of phi3Ts genome, with genome of Bacillus phage phi3T (KY030782.1). Fragment 's' which is unique for phi3Ts, can be detected within prophage genomes of 6 *B. subtilis* strains, isolated in different parts of the world, specifically: SRCM103612 (South Korea), MB9_B1 and MB8_B1 (Denmark), JAAA (China), HMNig-2 (Egypt) and SSJ-1 (South Korea).

Figure 2. Hybrid phages and extrachromosomal fragments of phage DNA, detected in the evolved strains. A) Top: Genome comparison of phi3Ts and SPβ (Query cover=58%, Percent Identity=99.73%), where regions of high homology (73.6-100%) are shown in grey, and regions of 99% homology are connected. Segments that are unique for phi3Ts, or SPβ are highlighted in black and pink, respectively. Phage genomes are arranged according to their integration into the host chromosome, which is represented in red. hybrid phi3Ts-SPβ phages spontaneously released by the evolved strains. Below: Each hybrid phage contains DNA segments that are unique to SPβ or phi3Ts. Based on the position of these segments, the simplified schemes of recombinant phages were created (below the hybrid name). $Hyb^{phi3Ts\text{-}SP\beta}$ additionally carries fragment of the host chromosome, adjacent to the left att site. Picture of corresponding plaque is placed on the right site of phage maps. Numbers on phage maps indicate unique recombination sites, which are characterized in Suppl. Table 1. B) Extrachromosomal phage DNA fragments detected during PacBio sequencing, colored according to their homology to phi3Ts, SPβ, or fragments of host chromosome flanking phage integration sites. Fragments are ordered according to sequencing coverage relative to the

chromosomal region, which is represented as bar chart on the left. Numbers on epDNA maps indicate unique recombination sites, which are characterized in Suppl. Table 1.

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

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

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

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

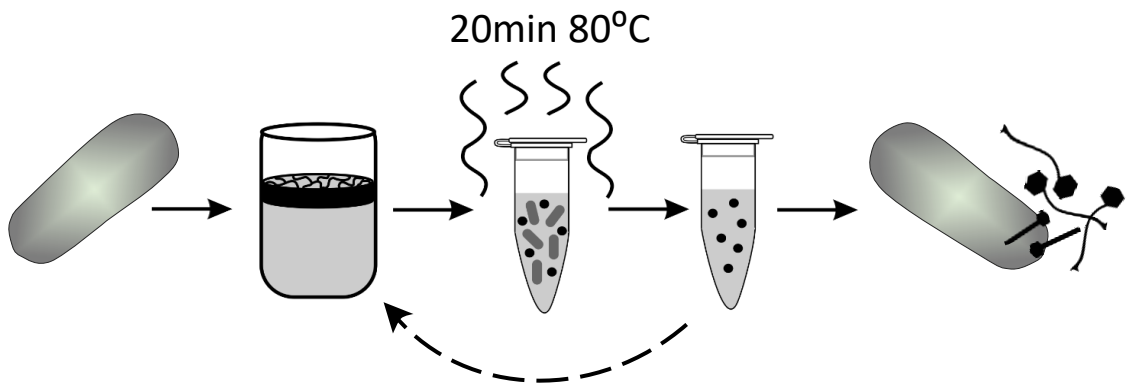
prophage. The tree was arbitrarily divided into 6 clades. B) Phylogenetic tree of SPβ-like prophages hosted by the strains in A). Inner circle shows prophage integration site, while outer circle indicates presence/absence of conjugative element ICEBs1, which blocks SPβ lytic cycle C). Selected prophages 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.

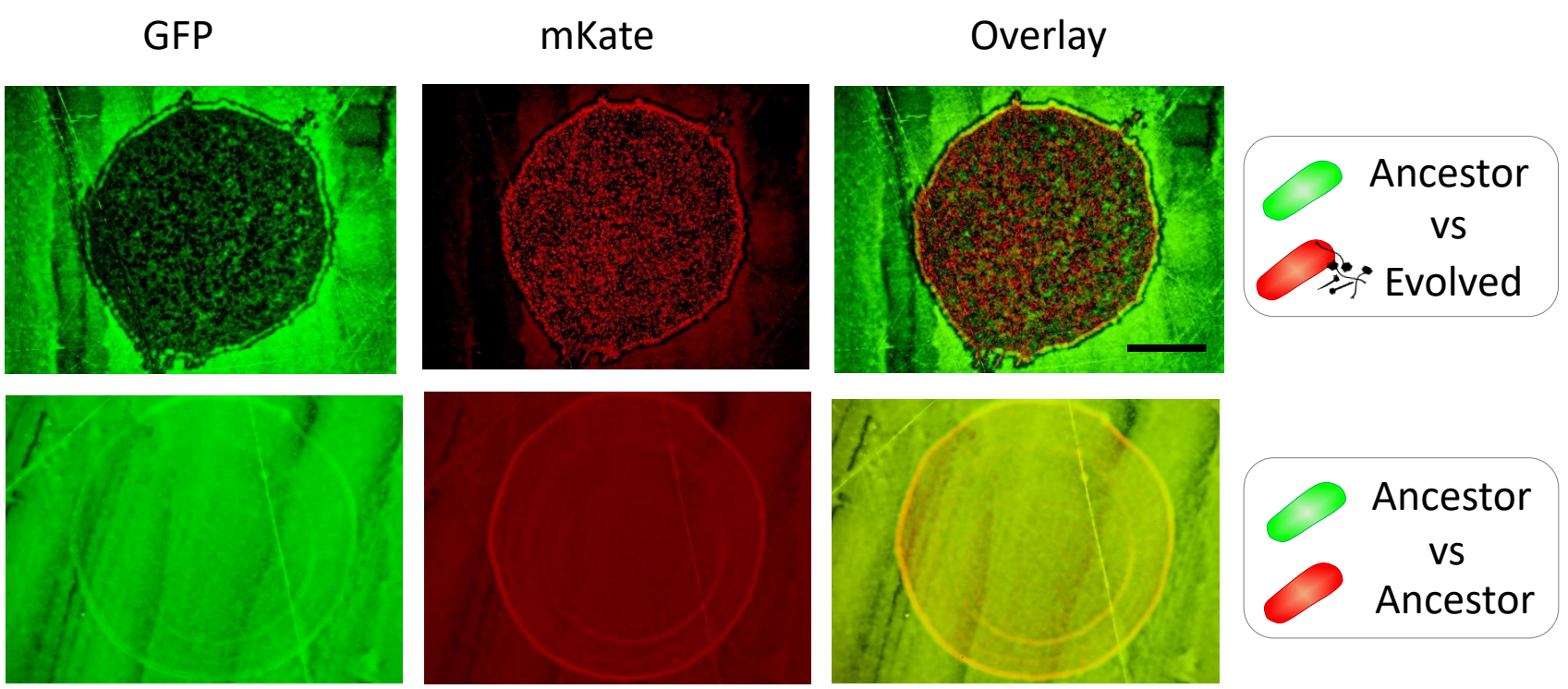
Name of bacterial strain/phage	Data	Accession number
B310mA	Complete genome	CP051860
B410mB	Complete genome	CP053102*
B410wtB	Complete genome	CP052842*
B310mA	Sequencing reads (Illumina)	SRR11561554
B410mB	Sequencing reads (Illumina)	SRR1156151
B410wtB	Sequencing reads (Illumina)	SRR11561552
$168_{ancestor}$	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.

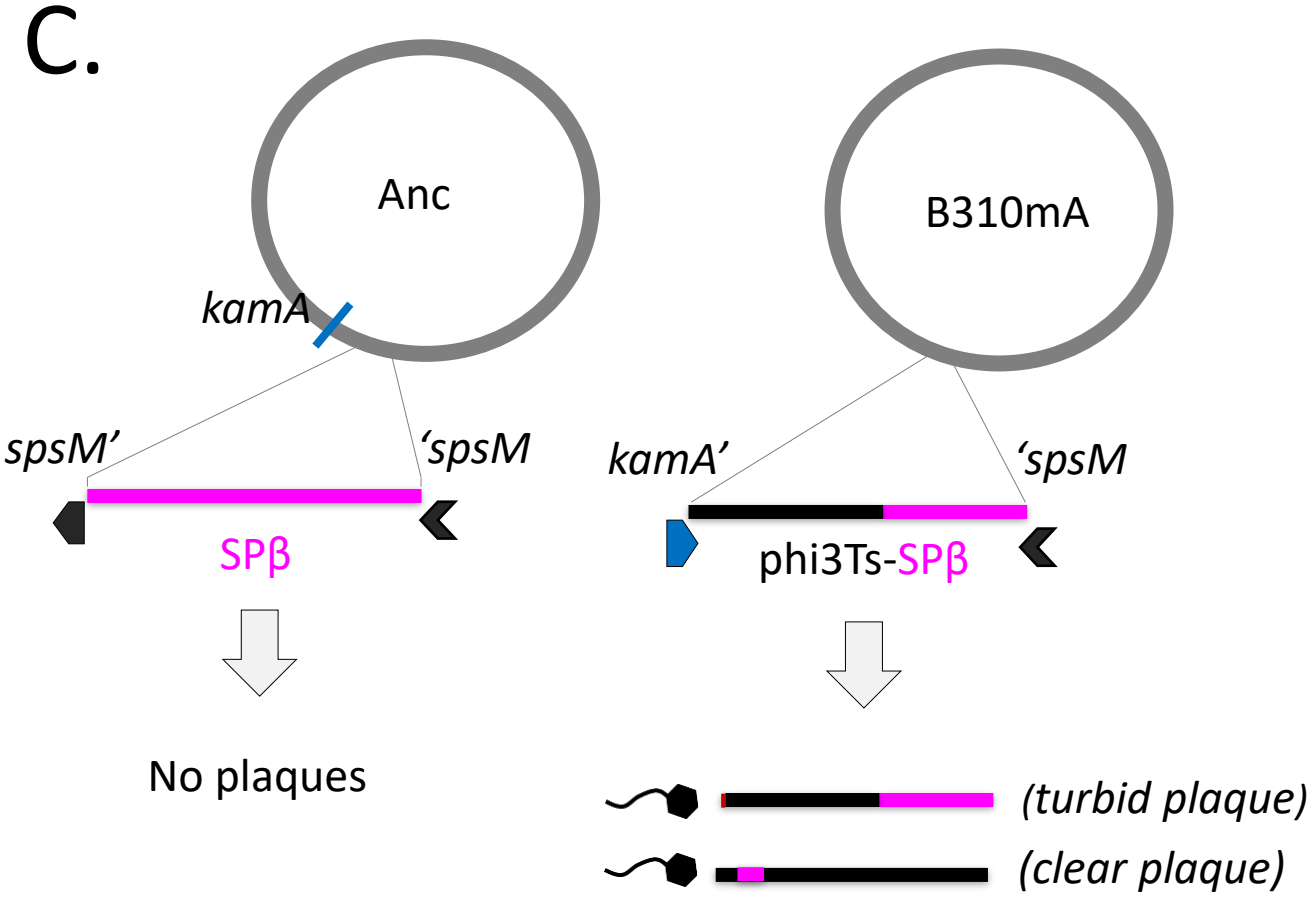
A.


20min 80°C

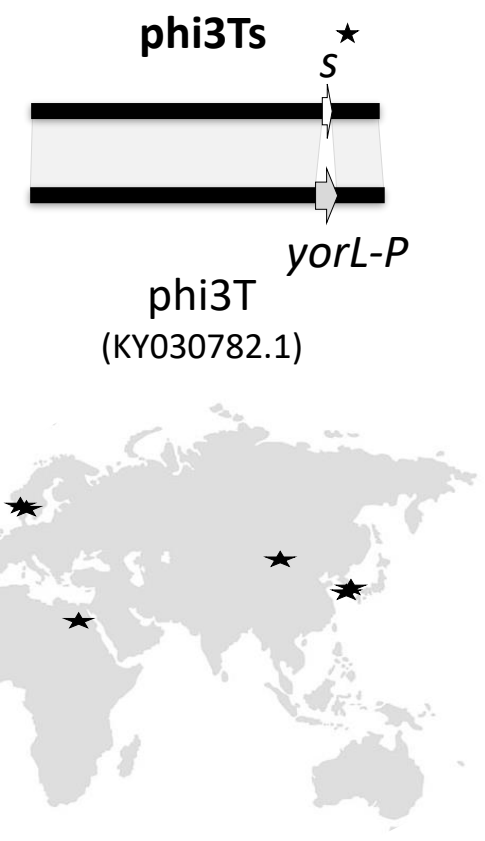
B.


GFP
mKate
Overlay
Ancestor
vs
Evolved
Ancestor
vs
Ancestor

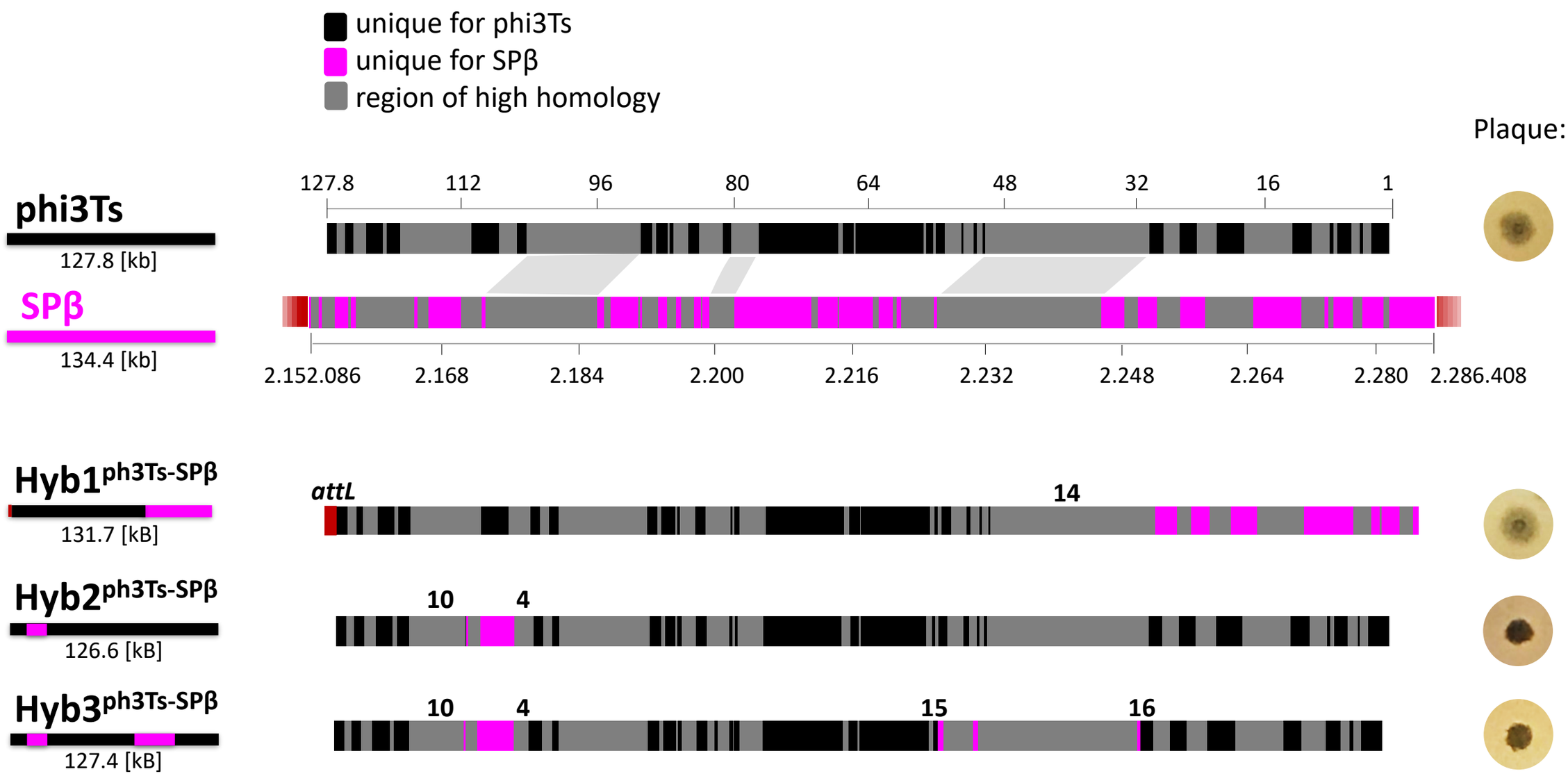
C.


Anc
kamA
spsM'
'spsM
SPβ
No plaques
B310mA
kamA'
'spsM
phi3Ts-SPβ
(turbid plaque)
(clear plaque)

D.

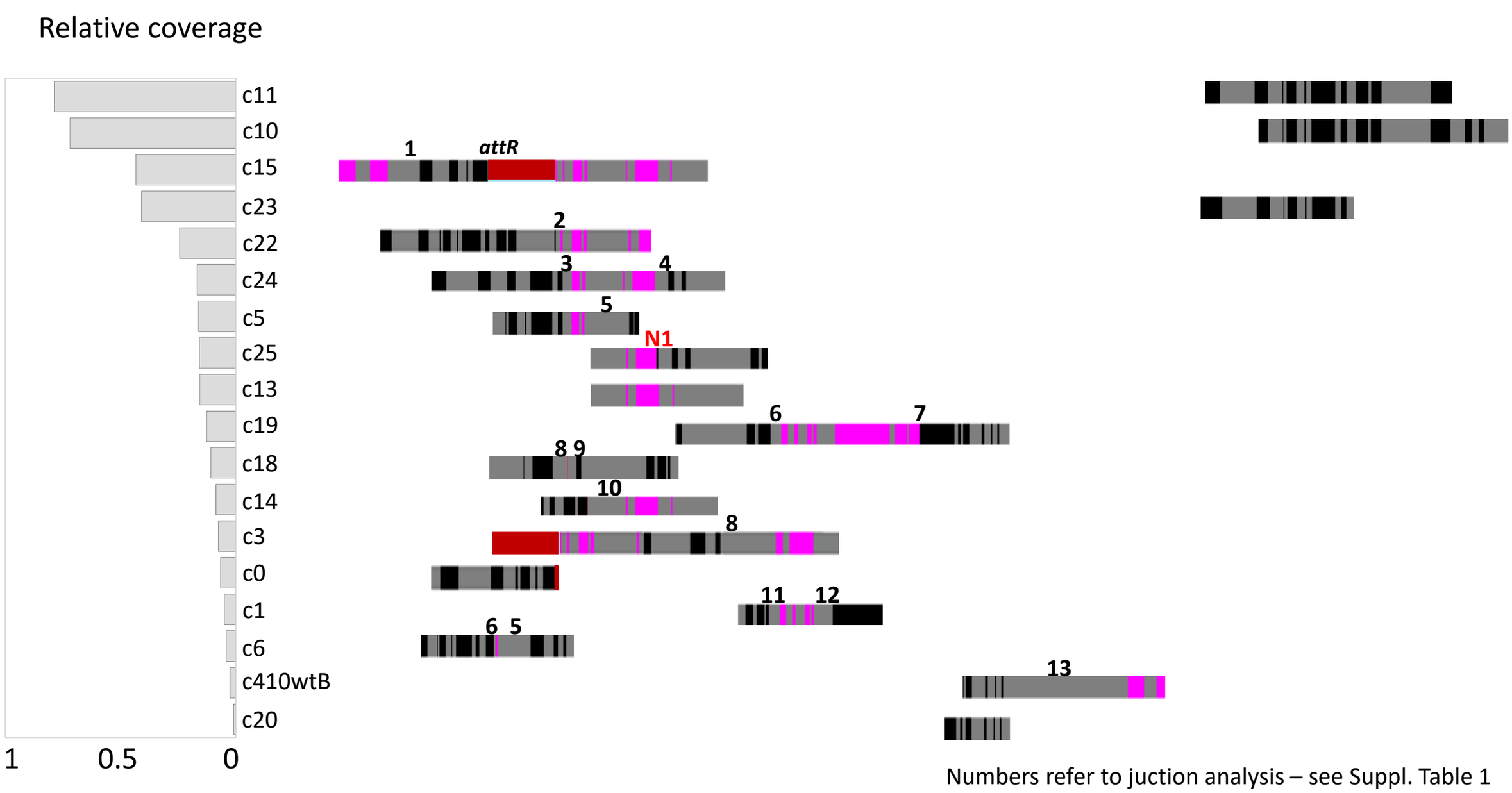

phi3Ts
s
yorL-P
phi3T
(KY030782.1)

A.

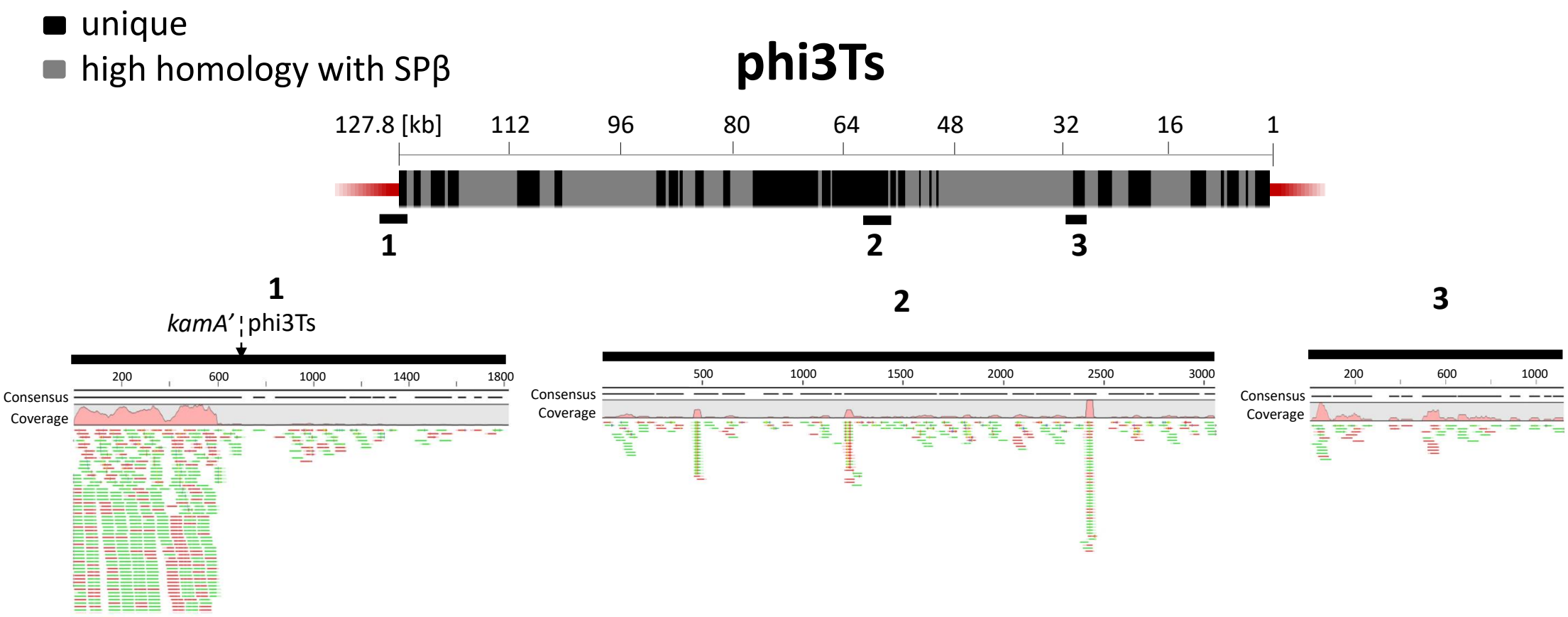


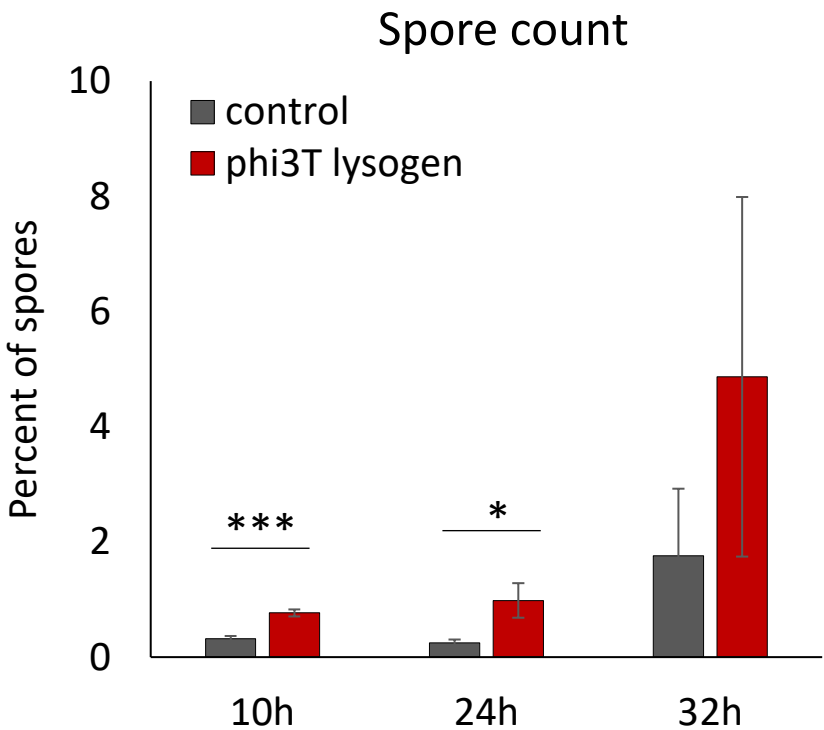
Numbers refer to juction analysis – see Suppl. Table 1

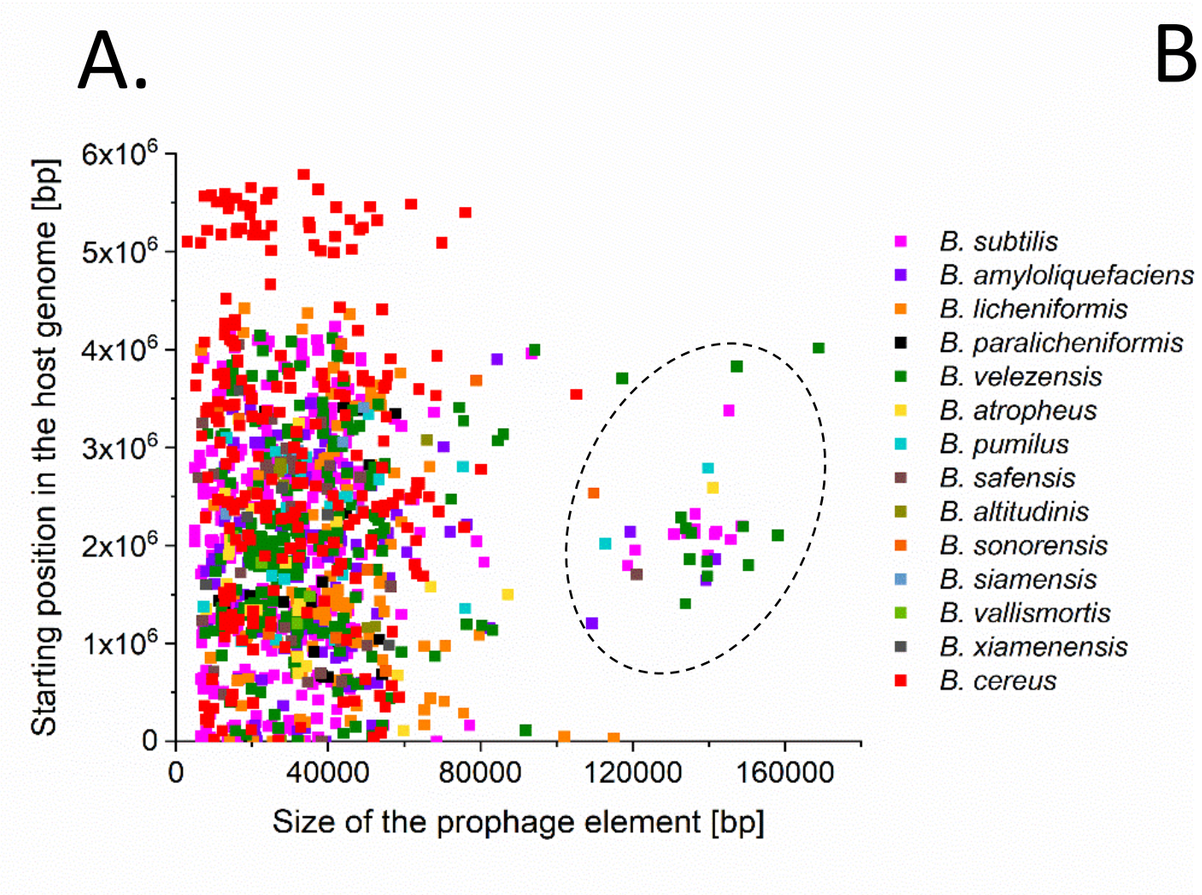
B.

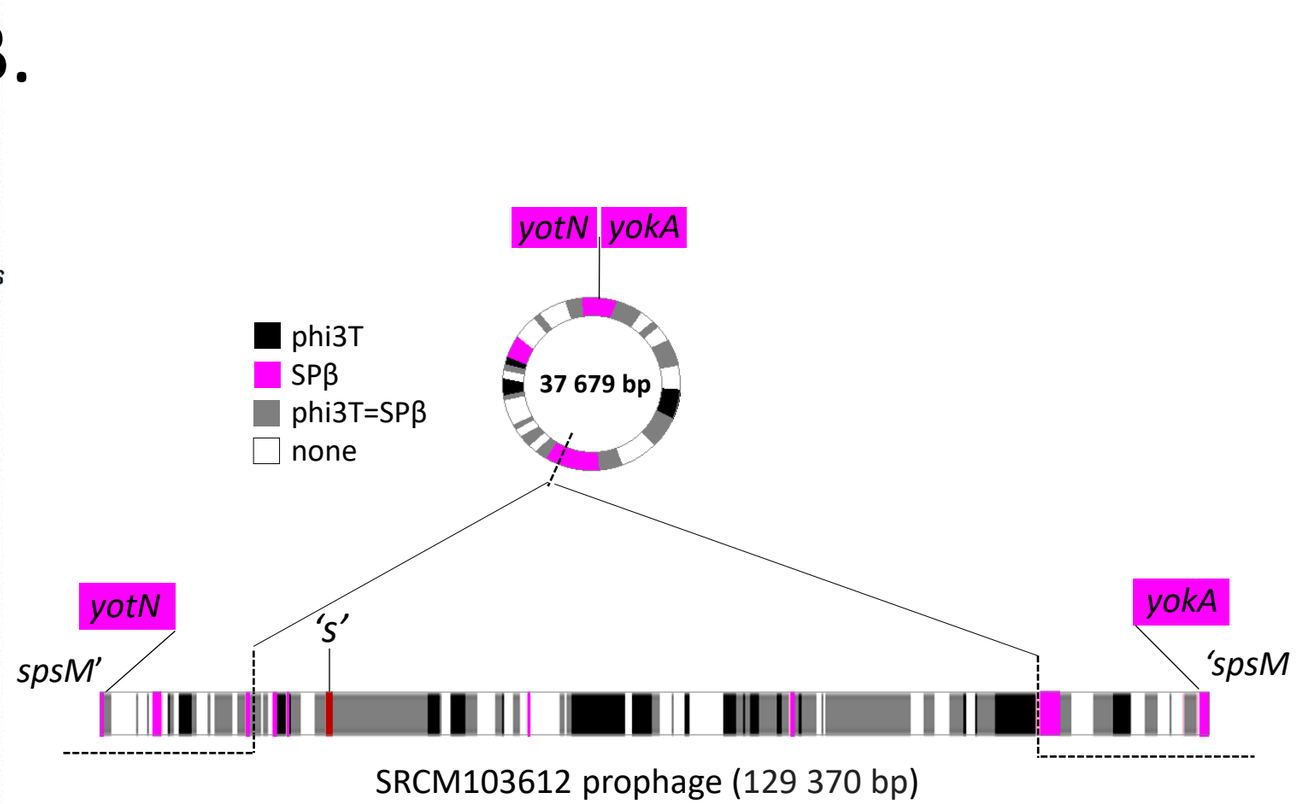


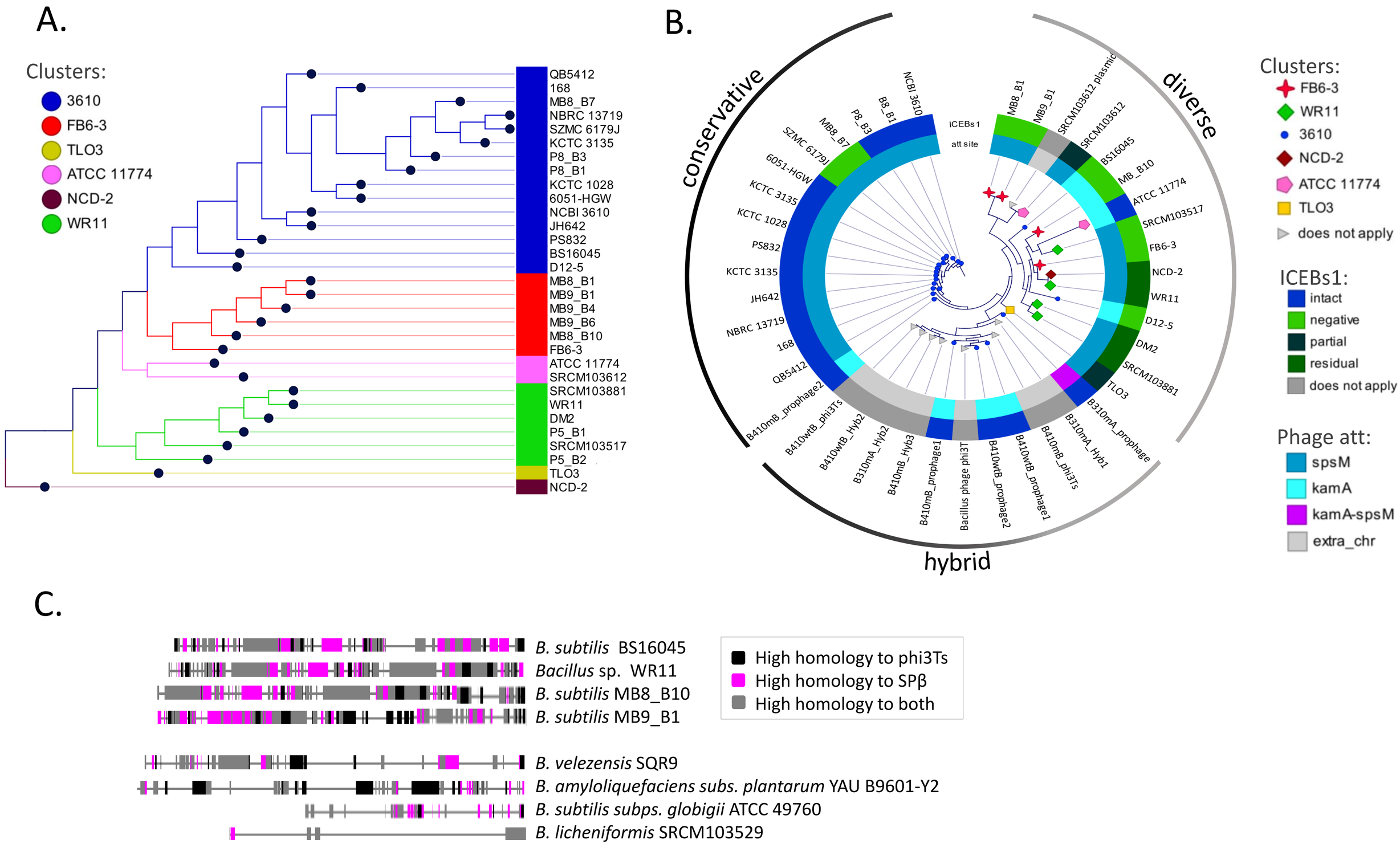
Numbers refer to juction analysis – see Suppl. Table 1


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


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


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
'S'
yokA
spsM'
'spsM
SRCM103612 prophage (129 370 bp)


A.
Clusters:
3610
FB6-3
TLO3
ATCC 11774
NCD-2
WR11
QB5412
168
MB8_B7
NBRC 13719
SZMC 6179J
KCTC 3135
P8_B3
P8_B1
KCTC 1028
6051-HGW
NCBI 3610
JH642
PS832
BS16045
D12-5
MB8_B1
MB9_B1
MB9_B4
MB9_B6
MB8_B10
FB6-3
ATCC 11774
SRCM103612
SRCM103881
WR11
DM2
P5_B1
SRCM103517
P5_B2
TLO3
NCD-2
B.
conservative
diverse
hybrid
ICEBs1
att site
Clusters:
FB6-3
WR11
3610
NCD-2
ATCC 11774
TLO3
does not apply
ICEBs1:
intact
negative
partial
residual
does not apply
Phage att:
spsM
kamA
kamA-spsM
extra_chr
C.
B. subtilis BS16045
Bacillus sp. WR11
B. subtilis MB8_B10
B. subtilis MB9_B1
B. velezensis SQR9
B. amyloliquefaciens subs. plantarum YAU B9601-Y2
B. subtilis subps. globigii ATCC 49760
B. licheniformis SRCM103529
High homology to phi3Ts
High homology to SPβ
High homology to both