nature microbiology

ARTICLES

https://doi.org/10.1038/s41564-021-01008-5

Dormant phages communicate via arbitrium to control exit from lysogeny

Nitzan Aframian[1,4], Shira Omer Bendori[1,4], Stav Kabel[1], Polina Guler[1], Avigail Stokar-Avihail[2], Erica Manor[1], Kholod Msaeed[1], Valeria Lipsman[1,3], Ilana Grinberg[1], Alaa Mahagna[1] and Avigdor Eldar[1]✉

Temperate bacterial viruses (phages) can transition between lysis—replicating and killing the host—and lysogeny, that is, existing as dormant prophages while keeping the host viable. Recent research showed that on invading a naïve cell, some phages communicate using a peptide signal, termed arbitrium, to control the decision of entering lysogeny. Whether communication can also serve to regulate exit from lysogeny (known as phage induction) is unclear. Here we show that arbitrium-coding prophages continue to communicate from the lysogenic state by secreting and sensing the arbitrium signal. Signalling represses DNA damage-dependent phage induction, enabling prophages to reduce the induction rate when surrounded by other lysogens. We show that in certain phages, DNA damage and communication converge to regulate the expression of the arbitrium-responsive gene *aimX*, while in others integration of DNA damage and communication occurs downstream of *aimX* expression. Additionally, signalling by prophages tilts the decision of nearby infecting phages towards lysogeny. Altogether, we find that phages use small-molecule communication throughout their entire life cycle to sense the abundance of lysogens in the population, thus avoiding lysis when they are likely to encounter established lysogens rather than permissive uninfected hosts.

Temperate phages alternate between two states—a lytic state where they replicate, kill the host and release their progeny into the environment—and a dormant, lysogenic state, during which their genome typically integrates into the bacterial chromosome and replicates along with it. Temperate phages are thus faced with a critical decision between lysis and lysogeny at two points in their full life cycle—a decision of whether to enter lysis or lysogeny on infection and a decision made from the lysogenic state whether to persist as prophages or exit lysogeny through phage induction. In the highly studied phage λ, the lysis–lysogeny decision on infection is controlled by phage density through the detection of coinfection[1,2], while DNA damage to the host cell serves as a trigger for induction[3]. Other factors, such as the physiological state of the cell, also influence the development of phage λ[4].

Recently, phages infecting the genus *Bacillus* were shown to employ an alternative mechanism to regulate the lysis–lysogeny decision on infection, which is based on small-molecule communication. These phages encode for the peptide-based arbitrium communication system[5–7]. On phage entry into the cell, the phage-encoded intracellular receptor AimR and the pre-AimP arbitrium pre-peptide are expressed. Pre-AimP is cleaved on secretion to form the mature arbitrium peptide, AimP, which is transported back into the cell by the general oligopeptide permease system[5]. In the absence of mature AimP, AimR acts as an activator of the aimX regulatory RNA. Intracellular mature AimP interacts with AimR and prevents it from activating the expression of *aimX*[8–11].

Phylogenetic characterization identified ten clades of AimR, with SPβ-like phages encoding clade 2 receptors[6]. In SPβ-like phages, it is unknown how aimX promotes lysis. In contrast, in other arbitrium clades, *aimX* is transcribed in an antisense direction to a putative phage repressor and operates in *cis* to reduce its transcript level[6].

In total, the arbitrium system biases infecting phages towards lysogeny at late rounds of infection, when the arbitrium peptide reaches high concentrations and phage density is high. In this sense, small-molecule communication provides a similar phage density-dependent effect on lysis–lysogeny decision-making to that provided by the coinfection-sensing mechanism employed by phage λ[12,13]. Therefore, it is unclear whether there are critical population-level differences between small-molecule signalling and coinfection sensing. One hypothesis we consider in this study is that small-molecule signalling can continue from the lysogenic state, allowing prophages to inform other phages of their presence, a function that could not be achieved by coinfection sensing[14]. The frequency of lysogens in the population is valuable information from the phage perspective—lysogens are often immune to secondary infections, a phenomenon termed superinfection exclusion[15]. Therefore, the probability of failed infection attempts increases with lysogen frequency and the adaptive value of lysis decreases accordingly.

In this study, we show that prophages possessing the arbitrium system continue to communicate intercellularly. We demonstrate that signalling inhibits the process of phage induction triggered by DNA damaging agents and that signalling by prophages influences infection dynamics. Overall, we show that signalling from the lysogenic state allows phages to avoid lysis when surrounded by lysogens.

Results

Prophages communicate by secreting and sensing AimP. In contrast to coinfection-based mechanisms for sensing phage density, the arbitrium system may enable phages to communicate their presence from the dormant prophage (lysogenic) phase[14]. The activity

[1]Shmunis School of Biomedicine and Cancer Research, Faculty of Life Sciences, Tel-Aviv University, Tel-Aviv, Israel. [2]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. [3]Present address: Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot, Israel. [4]These authors contributed equally: Nitzan Aframian, Shira Omer Bendori. ✉e-mail: avigdor@gmail.com

of the arbitrium system during lysogeny and its possible impact have not been previously characterized, although global transcriptome analysis of the *Bacillus subtilis* lab strain 168, which hosts the arbitrium-carrying SPβ prophage, indicates that *aimR*, *aimP* and *aimX* are transcriptionally active[16].

To investigate the possibility of communication during lysogeny, we first asked whether prophages can sense the arbitrium peptide. To this aim, we introduced into an ectopic chromosomal locus an *aimX*-yellow fluorescent protein (YFP) transcriptional reporter containing the *aimX* promoter and part of the aimX transcript in an operon with three consecutive YFP genes (Methods). Comparison to reporter fluorescence in non-lysogens indicated that *aimX* is expressed in lysogens during growth in lysogeny broth (LB) medium (Fig. 1a). Expression was reduced on addition of the mature SPβ AimP arbitrium peptide (GMPRGA peptide, SPβ AimP) at saturating levels but not on addition of the mature AimP of phage ϕ3T (SAIRGA peptide, ϕ3T AimP), which does not affect the SPβ arbitrium system[5]. This suggests that AimR continually activates *aimX* during lysogeny and responds to AimP. Indeed, deletion of *aimR* from the prophage completely abolished *aimX* expression.

To test whether the arbitrium peptide is also produced by lysogens, we measured *aimX* expression in Δ*aimP* lysogens grown in LB (Fig. 1a). Deletion of *aimP* led to a substantial increase in *aimX*-YFP expression, implying that AimP is continuously produced during lysogeny and inhibits *aimX* activation by AimR. To show that AimP is secreted, we cocultured differently marked wild-type (WT) and Δ*aimP* lysogens (Fig. 1a,b). Expression levels of the *aimX*-YFP reporter in Δ*aimP* lysogens were reduced to levels of the WT, as expected from a secreted peptide. To further explore the sensitivity of the arbitrium system in lysogeny, we monitored *aimX*-YFP expression in response to a different concentration of arbitrium peptide in minimal medium (Extended Data Fig. 1b). We found that, at very low cell densities, addition of 1–10 pM of signal to a Δ*aimP* mutant was sufficient to partially repress the *aimX*-YFP transcriptional reporter. Indeed, *aimX*-YFP repression in a WT phage was observed starting at an optical density (OD) of approximately 10^{-3} (Fig. 1b; $f_{WT} = 100\%$). Overall, these results indicate that prophages communicate from lysogeny by a process of quorum sensing—the arbitrium peptide is secreted and sensed by surrounding lysogens.

Communication controls prophage induction. Next, we set out to explore the phenotypic response controlled by prophage communication. DNA damage triggers induction in many distantly related phages, including SPβ-like phages[17,18]. This strategy may be adaptive to the phage since DNA damage strongly reduces host viability and increases the relative benefit of horizontal transfer[19]. DNA damage-dependent induction may become maladaptive when a lysogen is surrounded by other lysogens, as indicated by high concentrations of AimP. This is especially true when free phages can adsorb to lysogens but cannot complete productive infections. A comparison between free phage adsorption of lysogens and non-lysogens suggests that SPβ lysogens can act as sinks of SPβ phages (Extended Data Fig. 1c).

To understand the interplay between arbitrium signalling and DNA damage in SPβ induction, we generated DNA damage by adding mitomycin C (MMC) to lysogens grown in LB. Phage induction was monitored by tracking both bacterial density (Fig. 1c and Extended Data Fig. 2) and free phage production (Fig. 1d). Non-lysogens exposed to MMC showed arrested growth and a slow reduction in OD, as expected in the face of DNA damage[20], while lysogens experienced a sharp population collapse due to phage induction (Fig. 1c and Extended Data Fig. 2). We note that *B. subtilis* lab strains contain an additional lytic element, the degenerate phage-derived bacteriocin PBSX[21]. Therefore our measurements were done in a genetic background where PBSX is inactivated by mutations (Methods).

We found that addition of ectopic SPβ AimP along with MMC to SPβ lysogens reduced phage production and prevented population collapse (Fig. 1c,d), while addition of the orthogonal ϕ3T AimP had no discernible effect (Fig. 1c). This indicates that high concentrations of the arbitrium signal repress the activation of induction in response to DNA damage. As expected from this regulation, the addition of MMC to Δ*aimR* lysogens did not lead to massive cell lysis and population collapse (Fig. 1c).

Since we established that AimP inhibits prophage induction, we expected induction of Δ*aimP* prophages to be even more rapid compared to WT. Induction rates in LB did not differ between these two strains (Fig. 1d and Extended Data Fig. 2). One possible explanation is that the expression levels of *aimX* in WT lysogens grown in LB may not be sufficiently low to prevent induction. While lysogens express *aimX* at mid-log phase in LB (Fig. 1a), in minimal medium *aimX* expression is shut down by this growth stage (Fig. 1b; $f_{WT} = 100\%$). This may be due to the competition the arbitrium peptide faces over the oligopeptide permease transporter with other peptides in LB, as shown with other peptide-based quorum-sensing systems[22]. Additional physiological factors may also contribute to the effect of the medium on phage decision-making, as described for other phages[23]. In agreement with *aimX* expression, the addition of MMC to Δ*aimP* lysogens led to substantially more induction compared to WT lysogens in minimal medium (Fig. 1d). Adding ectopic AimP along with MMC to Δ*aimP* lysogens reduced induction rates to WT levels. As expected given the complete suppression of *aimX* expression in minimal medium (Fig. 1b), induction of WT prophages was not affected by supplementing minimal medium with ectopic AimP. Therefore, our results show that native concentrations of secreted peptide are sufficient to dominantly suppress induction in minimal but not rich medium.

Since the frequency of lysogens in the population is a crucial variable for the probability of successful future infections, we expected the expression patterns of *aimX* to depend on the frequency of signal producers. To test this, we cocultured differentially marked WT and Δ*aimP* lysogens in minimal medium and tracked *aimX*-YFP reporter expression in WT lysogens (Fig. 1b). We found that under these conditions, approximately 3% of WT lysogens in the population were sufficient to partially suppress *aimX* expression by mid-log phase, while a fraction of approximately 30% led to strong suppression. In contrast, when WT lysogens make up <1% of the population, high levels of *aimX* expression are maintained even at high densities. To conclude, we find that prophages communicate during lysogeny to inhibit induction when they sense that they are surrounded by other lysogens, as indicated by high concentrations of the arbitrium peptide (Fig. 1e). This strategy may allow phages to avoid induction under conditions where their progenies are likely to encounter established lysogens rather than permissive bacteria.

Communication inhibits induction through different mechanisms in different phages. To gain further insight into the inhibition of induction by arbitrium signalling, we used PCR with reverse transcription (RT–PCR) to measure the transcript levels of several bacterial and SPβ genes (Fig. 2a). Transcription levels of the bacterial SOS genes *recA* and *yneA*[17] rose just 20 min after MMC addition and were not strongly affected by the presence of the arbitrium signal (Extended Data Fig. 3a). Transcription levels of *sprB*, which controls prophage excision[24,25], rose only later (45 min after MMC addition) in the presence of DNA damage but only when the medium was not supplemented with the arbitrium peptide. The transcription of genes coding for phage RNA polymerase (*yonO*), putative phage capsid (*yonH*) and phage lysin (*cwlP*)[26,27] were similarly activated by DNA damage and inhibited by adding the signal. Overall, these results suggest that the arbitrium signal prevents the expression of the phage excision module without shutting down the general SOS response regulon.

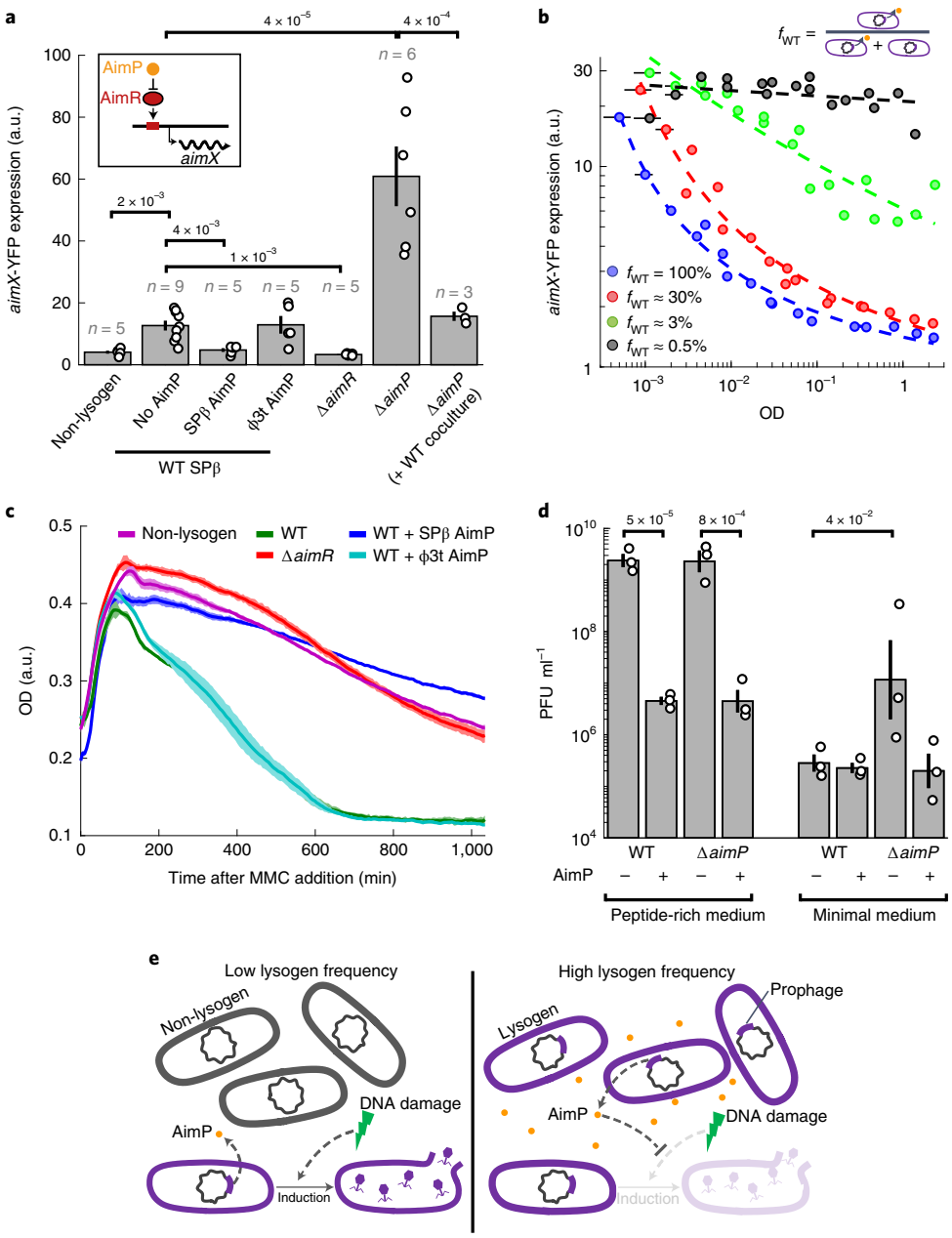


Fig. 1 | Arbitrium signalling is active during lysogeny and represses DNA damage-dependent prophage induction. a, Mean expression of an *aimX*-YFP reporter in LB at $OD_{600} = 0.3$ for the genotypes and conditions described. The empty circles represent the median expression levels obtained by flow cytometry of $>10^4$ cells. The rightmost bar is the expression of a Δ*aimP* lysogen coculture with WT lysogen. AimP was added at a concentration of 10 μM. The error bars represent the s.e.m. The numbers of biological replicates and *P* values of paired two-tailed *t*-tests are indicated in the figure. a.u., arbitrary units. **b**, Expression of a WT lysogen carrying an *aimX*-YFP reporter in coculture with a Δ*aimP* strain as a function of OD and frequency. The colours mark different approximate frequencies of the WT (f_{WT}) as indicated in the legend. The dashed lines show the best fit of data for each frequency ($\log(Y) = a + \frac{b}{\log(n)+c}$, where *n* is cell density). Low ODs (<0.02) were calculated by back extrapolation based on measured growth rate in minimal medium (50 ± 8.5 min; Extended Data Fig. 1a). Errors in the calculation of estimated OD are marked by the horizontal line. **c**, Mean OD (dark colour) and s.e.m. (lighter shade) of three technical replicates as a function of time for the relevant strains and conditions (see legend, Supplementary Table 1). A total of 0.5 μg ml^{-1} of MMC was added at time 0 at $OD_{600} = 0.3$; 10 μM of AimP were added at time 0 when indicated (Extended Data Fig. 2a–c). **d**, PFU for cultures induced with MMC at $OD_{600} = 0.3$ in peptide-rich medium (LB; lysate taken 120 min after MMC addition) or in minimal medium (SMM; lysate taken 150 min after the addition of MMC). The empty circles represent $n = 3$ biological replicates; the geometric means and error bars represent the s.e.m. of the log. The *P* values of a paired two-tailed *t*-test are indicated in the figure. **e**, At low lysogen frequency, DNA damage leads to prophage induction. At high lysogen frequency, the lysogen-secreted arbitrium signal inhibits induction. Lysogens and prophage are marked in purple.

DNA damage and arbitrium signals could be integrated at multiple different stages to yield the final decision of induction. To examine whether DNA damage acts upstream of the arbitrium communication system, we monitored *aimX* expression levels through our *aimX*-YFP reporter. Since cells do not divide under DNA damage, their total fluorescence may rise[20]. We controlled

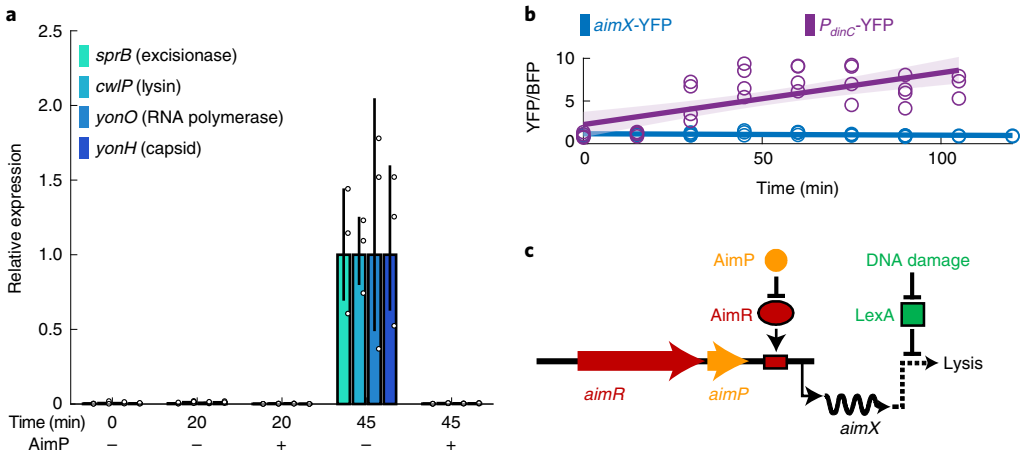


Fig. 2 | In the SPβ phage, DNA damage and communication are integrated downstream of *aimX* expression. a, Relative transcript levels of several SPβ genes (legend) measured using RT–PCR at different times after the addition of 0.5 μg ml^{-1} of MMC in the presence or absence of 10 μM of ectopic SPβ AimP. Data are normalized to 45 min after MMC addition in the absence of peptide. The geometric means of $n=3$ biological replicates are shown. The error bars correspond to the s.e.m. of the log. **b**, Expression of the *aimX*-YFP reporter of the SPβ phage (blue, slope estimate $P=0.08, n=26$) as a function of time after the addition of MMC, taken using flow cytometry. For each measurement, median YFP levels were divided by median BFP levels from a constitutive BFP reporter (Methods). Also shown are the YFP expression levels of the transcriptional reporter of a known LexA-repressed gene, P_{dinC}-YFP[18] divided by the same BFP reporter (purple, slope estimate $P=2\times10^{-5}$, $n=29$). The solid lines and lighter shade represent the linear fit of the corresponding data and its error. **c**, In the SPβ phage, LexA-mediated regulation of phage induction occurs downstream of *aimX* expression[19].

for this effect by dividing the YFP by the blue fluorescent protein (BFP) levels of a constitutive BFP reporter (Methods). We found this relative *aimX*-YFP expression to be independent of DNA damage (Fig. 2b and Extended Data Fig. 3b). In contrast, a reporter for the DNA damage-induced gene *dinC*[17] showed an increased expression in response to DNA damage. *aimX*-YFP insensitivity to DNA damage was further verified by quantitative PCR with reverse transcription (RT–qPCR) of YFP transcript level (Extended Data Fig. 3c) and by analysis of publicly available expression data (Extended Data Fig. 3d)[16]. Therefore, integration of DNA damage and communication in the SPβ phage occurs downstream of *aimX* expression (Fig. 2c).

Next, we wanted to explore the generality of our findings in other arbitrium-coding phages. We initially examined the SPβ-like phage ϕ3T, which also harbours a clade 2 arbitrium system. As in SPβ lysogens, DNA damage caused by the addition of MMC led to population collapse of ϕ3T lysogens, which was prevented specifically when ectopic ϕ3T arbitrium peptide (SAIRGA, ϕ3T AimP) was added to the medium (Extended Data Fig. 2d).

In most arbitrium clades, *aimX* overlaps a putative oppositely directing phage repressor, suggesting that it inhibits repressor expression through *cis* antisense regulation[6]. An arbitrium clade 1-encoding phage with such an architecture is found in *B. subtilis* subsp. *inaquosorum* (strain KCTC13429). This phage is similar to phage ϕ105 in its lytic genes but altered in its lysogenic domain; Extended Data Fig. 4a)[28]; therefore, we termed it ϕ106 (Fig. 3b). While we could not infect any lab strain with this phage, genomic sequencing of DNase-protected DNA in the lysate derived from the above strain identified the existence of the ϕ106 genome, indicating that this phage indeed forms phage particles on induction (Extended Data Fig. 4b and Methods).

We used RT–PCR to monitor the induction of this phage on addition of MMC, either with or without the putative arbitrium AimP signal of this system (DPPVGM). We monitored three genes: a transcription factor (DKG76_14355); a replication organizer homologue (DKG76_14340); and a holin homologue (DKG76_14150). We found that addition of MMC strongly activated these genes. Addition of the AimP signal together with MMC reduced the expression of these genes by a factor of five compared to the addition of MMC alone (Fig. 3a). In contrast, expression of the SOS genes *recA* and *yneA* was induced by MMC but unaffected by AimP (Extended Data Fig. 4c). Altogether, these results suggest that repression of induction by signalling may be a general theme in arbitrium systems.

On inspection of the *aimX* sequence, we noted the existence of a nucleotide sequence with high similarity to the consensus binding site of the DNA damage-responsive SOS repressor LexA (Fig. 3b)[29]. This led us to hypothesize that DNA damage regulation in this phage occurs at the level of *aimX* transcription, through LexA repression of its expression.

To test this hypothesis, we cloned the phage ϕ106 arbitrium system together with part of aimX fused to a YFP gene ($aimRPX_{\phi106}$-YFP; Extended Data Fig. 5a) into a background strain that lacks all other inducible elements and contains a constitutive BFP reporter. We first examined *aimX* expression by monitoring the mean YFP:BFP ratio in the population under different environmental perturbations (Fig. 3c and Extended Data Fig. 5). We found that the addition of MMC increased the relative *aimX* reporter expression. Addition of peptide decreased both basal expression without MMC and MMC-induced expression. Therefore, these results indicate that in this system, in contrast to the SPβ phage, DNA damage is integrated with signalling upstream of *aimX* expression (Fig. 3b, top).

To test the role of LexA in the transcriptional regulation of *aimX*, we first overexpressed a dominant negative non-cleavable allele of LexA, known as LexA(ind-) (Methods)[18]. This allele prevents the expression of LexA-dependent genes under DNA damage conditions (Fig. 3d, top), as verified by examining its effect on a *dinC* reporter (Extended Data Fig. 6). We found that in this strain, *aimX* expression was still repressed by the arbitrium peptide but did not increase in response to MMC (Fig. 3d). A slight decrease in reporter activity can most likely be attributed to the physiological effects of LexA(ind-) (Extended Data Fig. 6). To validate direct repression of *aimX* by LexA, we modified two well-conserved bases in the LexA binding site (Fig. 3b). This mutation is expected to derepress *aimX* expression irrespective of DNA damage (Fig. 3e, top). Indeed, we found that *aimX* expression was high irrespective of the presence of MMC but was still sensitive to the addition of the arbitrium peptide (Fig. 3e). Therefore our results point to a mechanism of integration

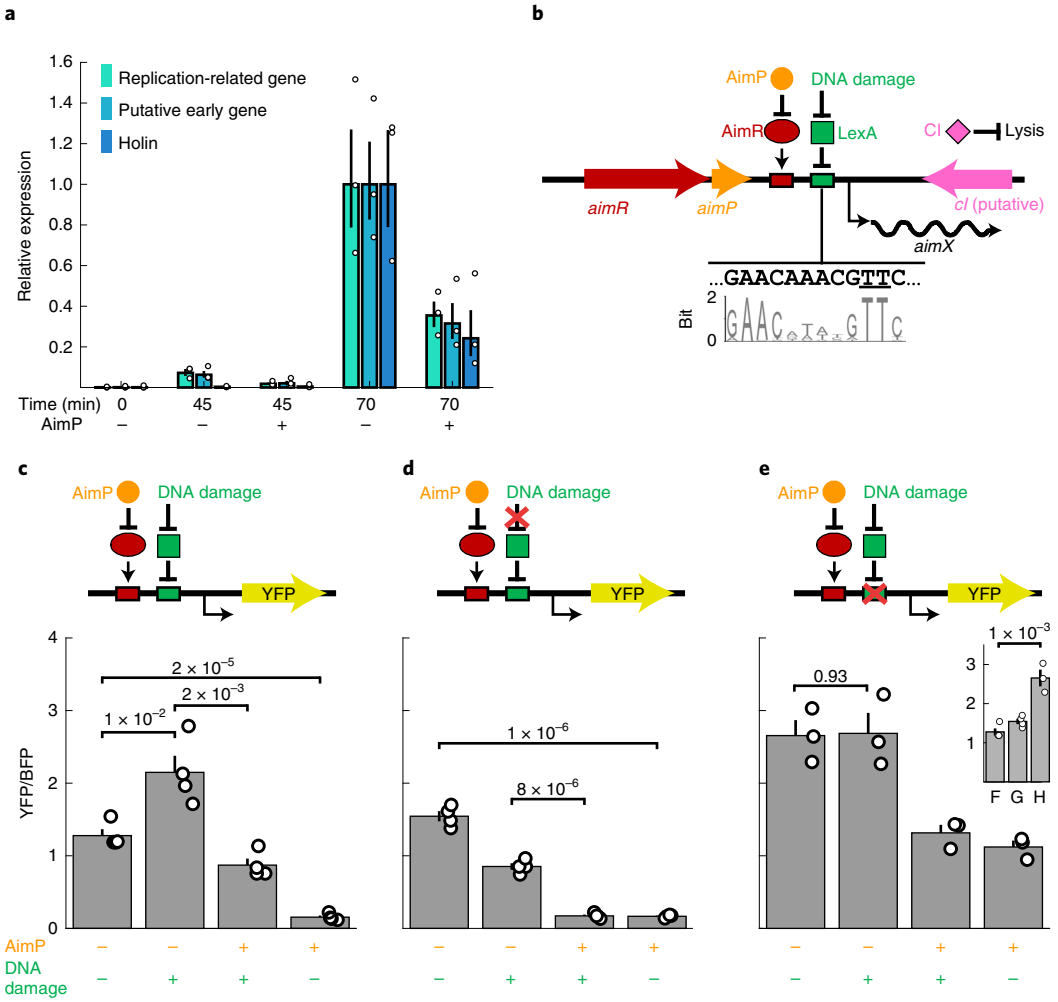


Fig. 3 | In phage ϕ106, DNA damage and communication jointly control *aimX* expression. a, Relative transcript levels of several ϕ106 genes (legend) measured using RT–PCR at different times after the addition of 0.5 μg ml^{-1} of MMC in the presence or absence of 10 μM of ectopic ϕ106 AimP. Data were normalized to 70 min after MMC addition in the absence of peptide. The geometric means of $n = 3$ biological replicates are shown. The error bars correspond to the s.e.m. of the log. **b**, Analysis of the ϕ106 arbitrium module identified a LexA binding site within the *aimX* locus and suggested a direct regulation of *aimX* expression by LexA and putative indirect regulation of repressor activity through the previously suggested sense–antisense mechanism[6]. Shown below the illustration is the identified LexA binding site; below it, the sequence logo of the LexA binding motif is identified[28]. Underlined are two highly conserved thymines that are replaced by adenines in the defective LexA binding mutant. **c**–**e**, Verification of LexA repression of *aimX*. The expression of an *aimRPX*$_{\phi106}$-YFP reporter of the phage ϕ106 arbitrium module, divided by a constitutive BFP reporter is shown. Data were taken 90 min after the conditional addition of ϕ106 AimP and MMC (see Extended Data Fig. 5 for the temporal data). Three different strains were assayed: a strain carrying an unmodified arbitrium module (**c**); a strain carrying an unmodified arbitrium module expressing a non-cleavable dominant negative variant of LexA (**d**)[18]; and a strain carrying an arbitrium module with two thymine base pairs of the LexA binding motif mutated to adenine (**e**) (Methods and Extended Data Fig. 6). Inset: comparison of the first column (no MMC or AimP) in **c**–**e**. The illustrations above each graph show the proposed regulation. A red cross points to the changes in regulation obtained by the modifications in **d**,**e**. **c**,**d**, Means and s.e.m. of $n = 4$ biological replicates. **e**, Means and s.e.m. of $n = 3$ biological replicates. The *P* values obtained by two-sided paired *t*-test are indicated in the figure.

of the arbitrium and DNA damage pathways by combined AimR activation and LexA repression of *aimX* transcription.

Finally, we searched for the clade 1 *aimR* receptor gene in approximately 300 genomes from the *B. subtilis* group of species and identified 81 instances. All these receptors were part of ϕ106-type phages (Methods). Despite variation in both the arbitrium peptide and genes adjacent to the lysogeny module, these phages shared the presence of a putative LexA binding site downstream of the *aimP* gene (Fig. 3b and Extended Data Fig. 7), demonstrating the generality of this regulation.

Prophage signalling alters infection dynamics. Signalling by prophages may not only affect induction but also tip the lysis–lysogeny decision of phages infecting nearby permissive bacteria towards lysogeny[5]. To test whether endogenous levels of peptide secreted by prophages are sufficient to impact the lysis–lysogeny decisions made by infecting phages, we grew cocultures of 90% SPβ lysogens and 10% permissive, non-lysogenic cells and infected them with WT phages (Fig. 4a). Lysogens were either Δ*aimR* mutants (which still produce the signal but are unable to induce; Fig. 1a) or Δ*aimRP* mutants (which neither produce the signal nor induce). Infecting phages, lysogens and non-lysogens were marked with different antibiotic resistance markers to enable identification of newly formed lysogens (Fig. 4a and Methods).

We infected cultures at a total multiplicity of infection (MOI) of 0.1 in minimal medium and shortly after (15 min) performed

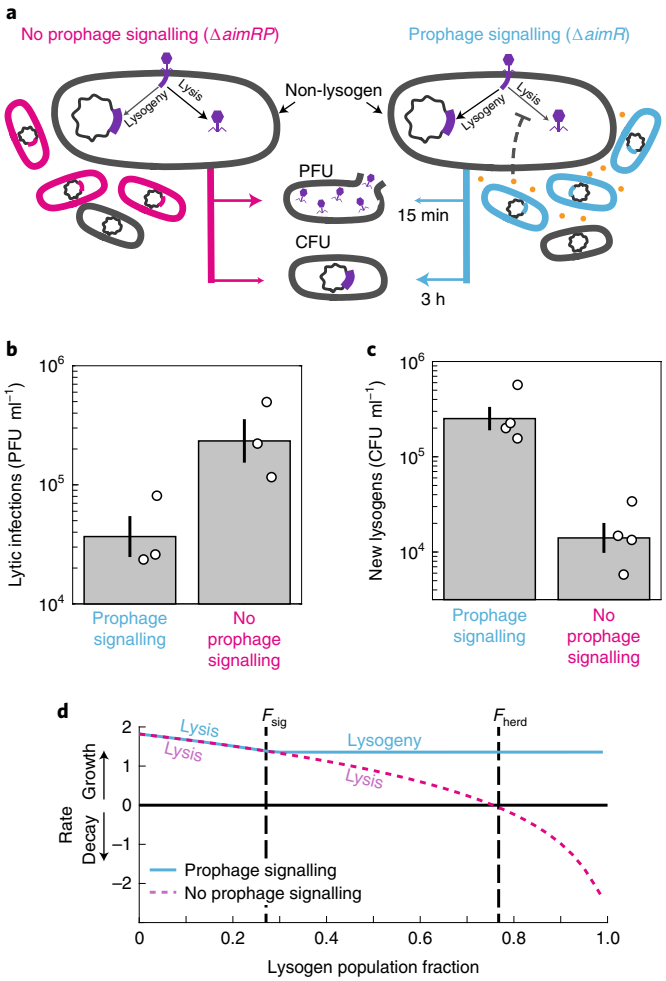


Fig. 4 | Prophage signalling biases infections towards lysogeny. **a**–**c**, Scheme (**a**) and results (**b**,**c**) of the experiment. **a**, A minority (10%) of permissive non-lysogens was incubated with either signalling (Δ*aimR*, cyan) or non-signalling (Δ*aimRP*, magenta) lysogens that do not produce phages. **b**, Cultures were infected by WT phages and monitored for PFU after 15 min. The circles represent $n=3$ biological replicates. The geometric means and error bars representing the s.e.m. of the log ($P=0.01$, two-sided *t*-test for the ratio of PFU between cultures) are shown. **c**, The same cultures were monitored for CFU after 3 h. The empty circles represent $n=4$ biological replicates. The geometric means and error bars representing the s.e.m. of the log ($P=0.02$, two-sided *t*-test for the ratio of CFU between cultures) are shown. **d**, A mathematical model of selection predicts an optimal threshold frequency for response to prophage signalling. The average rate of growth or decay of a phage infecting a permissive host as a function of the initial number of lysogens in the population is shown (Supplementary Fig. 2). Two infection models are considered—without prophage signalling (magenta) or with prophage signalling (cyan). The former is essentially lytic irrespective of the lysogen fraction (Supplementary Fig. 1), while the latter switches from lysis to lysogeny above a signalling threshold corresponding to a population threshold F_{sig}. The optimal threshold, where the switch occurs when lysogeny and lysis have the same growth rate, is shown. Also marked is the herd immunity frequency over which lytic infection results in decay (F_{herd}). Details of the model are presented in the Supplementary Text.

a plaque assay without filtering out bacteria, allowing us to track the first round of lysis–lysogeny decisions (Fig. 4a and Methods). We found that the number of infecting phages that lysed bacteria and formed plaques was reduced by a factor of 6 in cocultures with signal-producing lysogens compared to those with signal-null lysogens (Fig. 4b, $P=0.01$, *t*-test for ratio of plaque-forming units (PFU) between cultures). Conversely, we found that the number of lysogens formed by infecting phages 3 h after infection was 17-fold higher in cocultures containing signal-producing lysogens compared with signal-null lysogens (Fig. 4c, $P=0.02$, *t*-test for ratio of colony-forming units (CFU) between cultures). In total, the results show that under these conditions signalling by prophages indeed pushes infecting phages towards lysogeny and by this protects surrounding bacteria from lysis (Fig. 4c).

Modelling suggests an optimal switching point of lysis–lysogeny in response to prophage signalling. It is clear that sensing lysogens will allow phages to modulate lysis–lysogeny decisions in a way that will increase their fitness, particularly when lysogens are abundant. However, the optimal point at which phages should switch to a strategy of lysogeny is not immediately obvious. To gain insight into this from a theoretical standpoint, we defined the total phage growth as the average growth rate over time of phages in all their forms—free, infecting or lysogenic—per an initial successful infection (that is, phage infection of a permissive cell). Under this framework, the frequency of lysogens at which phages should choose lysogeny over lysis occurs when the growth rate of lysogens is greater than that achieved by proliferation through lysis. Notably, if the lysogen growth rate is positive, this critical point is below the threshold of herd immunity. Importantly, λ-like phages will choose lysis past this point to their detriment since they cannot sense the abundance of lysogens.

This analysis was corroborated by formalizing two mathematical models of phages with a λ-like, coinfection sensing-based mechanism and an arbitrium-like mechanism that responds to the presence of lysogens through prophage signalling (Fig. 4d; see the Supplementary Text for a description of the models). We also used modelling to further investigate the consequences of possible errors in the switching point of arbitrium signalling phages (Supplementary Text and Supplementary Figs. 1 and 2). Altogether, our analysis points to the adaptive value of prophage signalling in the regulation of phage life cycle transitions.

Discussion

Phage-phage interactions are known to affect the lysis–lysogeny decision through coinfection[12,13,30,31]. The newly discovered arbitrium system allows for such interaction to occur between phages infecting different bacteria[5,6]. In this study, we showed that arbitrium intercellular communication between prophages is maintained during lysogeny and controls induction in addition to lysis–lysogeny. In two different phage families, induction was activated by DNA damage and repressed by arbitrium signalling. Prophage signalling also biased infecting phages towards lysogeny. Our results suggest that prophage communication may be an adaptive strategy of phages aimed at reducing the risk of detrimental lysis of a host when lysogens are already prevalent. Thus, prophage signalling, as opposed to communication exclusively during active infections, provides a benefit that cannot be gained by coinfection-based mechanisms.

For the arbitrium systems of clade 1, our data point to a mechanism underlying the control of phage induction by DNA damage and arbitrium communication—*aimX* expression is activated by the arbitrium receptor AimR and repressed by the SOS regulator LexA. aimX in turn regulates the putative phage CI repressor by *cis* antisense regulation[6]. The mechanism connecting aimX to phage induction in SPβ-like phages is to be elucidated. Nevertheless, we have shown that in contrast to clade 1 systems, in these phages, the inputs of DNA damage and arbitrium signalling are integrated downstream of *aimX* expression.

During the final revision of this work, two additional studies exploring the role of the arbitrium system during lysogeny and

prophage induction were published[32,33]. Brady et al.[33] focused on the molecular mechanism of induction in SPβ phages and suggested that *yopR* is the master phage repressor, although the link between phage communication and phage repressor is unknown. Bruce et al.[32] used modelling to explore the eco-evolutionary aspects of phage life cycle regulation by arbitrium. They suggested that in well-mixed conditions induction should be more sensitive to low peptide concentrations compared to the lysis–lysogeny decision and provided experimental evidence to support this[32]. The results presented in both studies fall in line with our main points.

In nature, bacteria are thought to often reside in dense spatially structured communities[34]. Recent analysis suggests that under such conditions, peptide-based quorum-sensing systems report for the local frequency of signal producers rather than their overall density[35]. In a dense structured community, the local frequency of lysogens is a better predictor of successful future infections than the make-up of the population more globally. Therefore, the use of short-range communication may be especially advantageous in a natural setting.

Additionally, the short range of arbitrium communication coupled with the clonal nature of bacterial populations at this local scale may also help relax the conditions under which prophage signalling is beneficial. Modelling suggests that the point at which the trade-off posed by the strategy of lysogen signalling becomes beneficial, depends on several epidemiological parameters that may vary in natural conditions (Supplementary Fig. 2). If phages are likely to encounter communities that at a local scale are completely dominated by either lysogenic or non-lysogenic bacteria, responding to the existence of lysogens is substantially more robust with respect to these parameters (Supplementary Text)[35,36]. Further investigation of the function of phage communication in more natural settings could make an intriguing line of research.

Methods

Strain construction. All bacterial strains, plasmids and primers used in this study are listed in Supplementary Tables 1–3. To construct the new *B. subtilis* strains, standard transformation and SPP1 transduction protocols were used for genomic integration and plasmid transformation[37]. The *B. subtilis* lab strain PY79 contains an additional DNA damage-induced lytic element (the PBSX prophage-derived bacteriocin); therefore it was only used for the gene expression experiments[21]. For the plaque and growth assays, we used a *Δxpf* strain in which PBSX cannot induce[38]. Experiments with the arbitrium system of ϕ106 were conducted on the background of a PY79 strain where PBSX was deleted in its entirety and ϕ3T experiments were conducted on a Δ6 background that lacks all mobile elements[39].

The SPβ strain used in this work was originally described as WT by Erez et al.[5]. Notably, subsequent to the experiments described in this work, we found through sequencing and perturbation experiments that the SPβ phage strain we used harbours mutations that render it heat-inducible (SPβc2 strain)[40]. Importantly, we corroborated the effect of the arbitrium peptide on induction of WT heat-insensitive SPβ, showing that these mutations do not affect our observed phenotype (Supplementary Fig. 3).

aimP, *aimR* and *xpf* were deleted by transformation of kanamycin resistance cassettes from a *B. subtilis* 168 deletion library[41] into the appropriate strains. The kanamycin resistance cassette was excised from the *xpf* deletion using a Cre/*lox* system[42]. For the *aimRP* and PBSX deletions, we used a long flanking homology PCR method to replace genes with a kanamycin resistance cassette[42]. All primers used are listed in Supplementary Table 3.

Plasmid construction. The pAEC1563 plasmid was constructed by amplifying the *aimX* promoter using the POL01 and POL02 primer pair. The PCR product was digested with BamHI-HF and NheI-HF and cloned into pAEC277 digested with NheI-HF and BamHI-HF.

The pAEC1909 plasmid was constructed by amplifying the arbitrium locus (AimR, AimP, AimX) using the stav128 and stav132 primer pair from *B. subtilis* subsp. *inaquosorum* KCTC 13429 (AES7108) genomic DNA. The PCR product was digested with BamHI-HF and MfeI-HF and cloned into pAEC277 digested with EcoRI-HF and BamHI-HF.

The pAEC2079 plasmid was constructed by amplifying the whole pAEC1909 using the stav152 and stav153 primer pair that contains the point mutation in the putative LexA binding site. The original plasmid in the PCR product was degraded using DpnI. The PCR product was closed after adding phosphate using the polynucleotide kinase enzyme, followed by ligation.

The pAEC2081 plasmid was constructed by amplifying LexA(ind-) from AIG246 (ref. [18]) using the EM9 and EM10 primer pair. The PCR product was digested with SphI-HF and NheI-HF and cloned into pAEC1505 digested with NheI-HF and SphI-HF.

The pAEC1919 plasmid was constructed by amplifying the *dinC* promoter using the P200 and P201 primer pair. The PCR product was digested with BamHI-HF and NheI-HF and cloned into pAEC277 digested with NheI-HF and BamHI-HF.

Growth media and conditions. As a rich medium in this study, we used LB: 1% tryptone (Difco), 0.5% yeast extract (Difco) and 0.5% NaCl. For the experiments in minimal medium, we used Spizizen minimal medium (SMM; 2 g l^{-1} of $(NH_4)_2SO_4$, 14 g l^{-1} of K_2HPO_4, 6 g l^{-1} of KH_2PO_4, 1 g l^{-1} of trisodium citrate, 0.2 g l^{-1} of $MgSO_4{\bullet}7H_2O$), supplemented with trace elements (125 mg l^{-1} of $MgCl_2{\bullet}6H_2O$, 5.5 mg l^{-1} of $CaCl_2$, 13.5 mg l^{-1} of $FeCl_2{\bullet}6H_2O$, 1 mg l^{-1} of $MnCl_2{\bullet}4H_2O$, 1.7 mg l^{-1} of $ZnCl_2$, 0.43 mg l^{-1} of $CuCl_2{\bullet}4H_2O$, 0.6 mg l^{-1} of $CoCl_2{\bullet}6H_2O$, 0.6 mg l^{-1} of $Na_2MoO_4{\bullet}2H_2O$); 0.5% glucose served as a carbon source. Liquid cultures were grown with shaking at 220 r.p.m and a temperature of 37 °C. When preparing plates, medium was solidified by adding 2% agar. Antibiotics were added (when necessary) at the following concentrations: spectinomycin, 100 μg ml^{-1}; chloramphenicol, 5 μg ml^{-1}; kanamycin, 10 μg ml^{-1}; macrolide, lincosamide and streptogramin, 3 μg ml^{-1}; erythromycin + 25 μg ml^{-1} lincomycin. For the experiments that involved infection, media were supplemented with 0.1 mM of $MnCl_2$ and 5 mM of $MgCl_2$.

Flow cytometry. Flow cytometry was performed to quantify gene expression at the single-cell level using a Beckman Coulter Gallios flow cytometer equipped with four lasers (405 nm, 488 nm colinear with 561 nm, 638 nm). The emission filters used were: BFP, 450/50; YFP, 525/40; mCherry, 620/30. Constitutive mCherry and mTag2-BFP were used to distinguish between cocultured cells. Median YFP levels were measured relative to a set voltage that was approximately set such that a value of 1.25 was given to the autofluorescence of strain PY79 in SMM medium (Extended Data Fig. 8).

Tracking expression by fluorescent reporters. SPβ phage. To determine the expression levels of *aimX* in SPβ lysogens grown in rich medium, strains harbouring an *aimX*-YFP reporter were grown in LB overnight at 37 °C with shaking at 220 r.p.m. and then diluted by a factor of 1:100 into fresh LB medium containing 10 μM of the SPβ arbitrium peptide (GMPRGA) when indicated. On reaching $OD_{600} = 0.3$, fluorescence was measured using flow cytometry. To measure expression in cocultures, after reaching $OD_{600} = 0.3$, strains were mixed at a 50:50 ratio. Cocultures were grown overnight and subsequently diluted, regrown and measured for fluorescence in the same way as monocultures.

To track gene expression in response to MMC, strains harbouring fluorescent reporters for the transcription of *aimR* or *aimX* were grown in LB overnight at 37 °C with shaking at 220 r.p.m. and then diluted by a factor of 1:100 into fresh LB medium. On reaching $OD_{600} = 0.3$, 0.5 μg ml^{-1} of MMC (catalogue no. M4287; Sigma-Aldrich) was added to the medium and both YFP and BFP were measured every 15 min using flow cytometry.

To determine expression in minimal medium, strains were grown to $OD_{600} = 0.1$ in SMM containing trace elements and glucose at 37 °C with shaking at 220 r.p.m. At this stage, strains were mixed at different frequencies as indicated. Cultures were then diluted by a factor of 10,000–1,000,000 into fresh SMM medium supplemented with 10 μM of the SPβ arbitrium peptide when indicated and grown for about 12 h in exponential phase. OD and gene expression were measured at approximately 1-h intervals using a spectrophotometer and a flow cytometer, respectively. For very low densities, the OD at the time of measurement was estimated by extrapolating backwards from the OD at later times assuming a growth rate of 50 ± 8.5 min supported by growth rate measurements, where we fitted growth data (OD_{600} versus time) to an exponent.

ϕ106 phage. Strains harbouring the arbitrium system of phage ϕ106 upstream of a YFP reporter and a constitutive BFP reporter at a different locus (see plasmid construction and Supplementary Tables 1–3) were used to determine *aimX* expression in this system under different conditions. Cultures were grown in LB overnight at 37 °C with shaking at 220 r.p.m. and then diluted by a factor of 1:1,000 into fresh LB medium containing different concentrations of isopropyl ß-D-1-thiogalactopyranoside. On reaching $OD_{600} = 0.3$, 10 μM of the appropriate arbitrium peptide (DPPVGM) and/or 0.5 μg ml^{-1} of MMC were added when indicated (t = 0). YFP and BFP were measured at several time points.

Plaque forming assay. Samples for PFU measurements were collected from cultures centrifuged for 5 min at 4,000 r.p.m. at room temperature. Next, the supernatant was filtered using a 0.2 μm filter (catalogue no. 14-555-270; Sartorius Stedim Biotech). Then, 100 μl of filtered supernatant at an appropriate dilution was mixed with 200 μl of *B. subtilis* PY79 grown to OD_{600} in MMB (LB supplemented with 0.1 mM of $MnCl_2$ and 5 mM of $MgCl_2$) and left to incubate at room temperature for 5 min. Three millilitres of molten LB-0.5% agar medium (at 60 °C) supplemented with 0.1 mM of $MnCl_2$ and 5 mM of $MgCl_2$ was added, mixed and

then quickly overlaid on LB-agar plates. Plates were then incubated at 37 °C for 1 h and then overnight at room temperature to allow plaques to form.

Induction rate. For induction in rich medium, strains were grown overnight in at 37 °C with shaking at 220 r.p.m. and then diluted by a factor of 10,000 into fresh LB medium containing 10 μM of peptide when indicated. To assess levels of spontaneous induction, samples for PFU measurements were taken at $OD_{600} = 0.3$. To measure induction in the presence of DNA damage, 0.5 μg ml^{-1} of MMC was introduced at $OD_{600} = 0.3$ and samples for PFU measurements were collected after 120 min.

For induction in minimal medium, strains were grown to $OD_{600} = 0.1$ at which point cultures were mixed as indicated and diluted by a factor of 1,000,000 in fresh SMM medium. To assess spontaneous induction, samples for PFU measurements were taken at $OD_{600} = 0.3$. To measure induction in the presence of DNA damage, MMC was introduced at $OD_{600} = 0.3$ and samples for PFU were collected after 150 min.

Induction growth dynamics. To examine growth dynamics during prophage induction, strains were grown overnight in LB at 37 °C with shaking at 220 r.p.m. then diluted by a factor of 1:100 into fresh LB medium. On reaching $OD_{600} = 0.3$, cultures were supplemented with 0.5 μg ml^{-1} of MMC and 10 μM of peptides when indicated. OD measurements at a wavelength of 600 nm were performed in a 96-well plate using a plate reader (VICTOR Nivo Multimode Plate Reader; PerkinElmer).

Infection experiments. Strains were grown at 37 °C with shaking at 220 r.p.m. to $OD_{600} = 0.1$ in SMM containing trace elements and glucose and then mixed for cocultures at the indicated frequencies. Cultures were then diluted by a factor of 1,000,000 into fresh SMM medium containing 0.1 mM of $MnCl_2$ and 5 mM of $MgCl_2$, grown until they reached $OD_{600} = 0.3$ and infected with free SPβ phages harbouring a chloramphenicol resistance marker at an MOI of 0.1. Samples were taken after 15 min without filtering and mixed with a phage-free indicator strain to perform the plaque assay. Additional samples were taken 3 h after the time of infection to measure CFU on plates of agar supplemented with chloramphenicol.

RT–qPCR measurements. Total RNA was extracted from cells using a High Pure RNA Isolation Kit (Roche). To this end, cells were grown in LB at 37 °C with shaking at 220 r.p.m. to $OD_{600} = 0.3$ at which point MMC was added along with the arbitrium peptide when indicated. Samples at different time points were centrifuged for 5 min at 2,000 relative centrifugal force and pellets were flash-frozen in liquid nitrogen. Then, 1 μg of RNA was reverse-transcribed to complementary DNA (cDNA) using a qScript cDNA Synthesis Kit (Quanta BioSciences). RT–qPCR was performed on a Step One Plus Real Time PCR System (Thermo Fisher Scientific) using SYBR Green (Quanta BioSciences). The specificity of primers (listed in Supplementary Table 3) was validated using a melt curve. The efficiency of the primers used was between 95 and 105%, with calibration curves of $r^2 > 0.95$. Thermocycling parameters were as follows: holding stage at 95 °C for 1 min and 40 cycles of 2 steps—a first step at 95 °C for 5 s and a second step at 60 °C for 30 s. Transcript levels were normalized to levels of the reference gene *rpoB*. RNA samples were stored at −80 °C and cDNA was stored at −20 °C. Nucleic acid quantification was performed by NanoDrop 2000c Spectrophotometer (Thermo Fisher Scientific). Results were analysed using the Step One v.2.3 software by the standard ΔΔCt method. Results were analysed using the Step One software by the standard ith calibration curves of resorcinol performed on a Step One Plus Real Time PCR System using SYBR Green (the quantity, temperature and general conditions used were as detailed in the accompanying protocol).

Adsorption experiments. Strains were grown at 37 °C with shaking at 220 r.p.m. to $OD_{600} = 0.1$ in SMM containing trace elements and glucose and then diluted by a factor of 1,000,000 into fresh SMM medium containing 0.1 mM of $MnCl_2$ and 5 mM of $MgCl_2$, grown until they reached $OD_{600} = 0.3$ and infected with approximately 10^6 SPβ phages. Samples were taken after 15 min, centrifuged and filtered using a 0.2 μm filter. Free phages remaining in the medium were enumerated by plaque assay.

Detection of ϕ106-induced DNA. Overnight cultures of *B. subtilis* subsp. *inaquosorum* strain KCTC13429 grown in LB at 30 °C were diluted 1:50 into 2 tubes with 5 ml of fresh LB and grown at 30 °C with shaking at 210 r.p.m. until $OD_{600} = 0.3$. Then, MMC was added to a final concentration of 0.5 μg ml^{-1} and then incubated overnight at 30 °C with shaking at 210 r.p.m. Cell debris was then pelleted by centrifugation for 10 min at 4,000 r.p.m. The supernatant of the 2 samples was combined and filtered through a 0.2 μM filter, which was then stored at 4 °C until DNA extraction; 2 ml of the lysate was concentrated to approximately 500 μl using a 3 kD Amicon by centrifugation at 400 r.p.m. at 4 °C for 45 min.

The flow through was discarded and the concentrated lysate was transferred to a new Eppendorf tube. Then, 20 μg ml^{-1} of DNase I was added to a final concentration of 20 μg ml^{-1} and incubated at 37 °C for 1 h to degrade free bacterial DNA. DNA was then extracted using the QIAGEN DNeasy Blood & Tissue DNA extraction kit, starting from a proteinase K treatment at 56 °C for 1 h and following the rest of the protocol steps (binding to column, wash, elution). DNA concentration was measured using the Qubit dsDNA HS Kit, after which libraries were prepared for sequencing using a modified Nextera protocol[43]. Samples were then subjected to Illumina paired-end sequencing. Reads of proper quality were then aligned to the ϕ106 locus of *B. subtilis* subsp. *inaquosorum* strain KCTC13429 to calculate coverage. Paired reads at a long distance from each other were identified and their position was registered.

Phylogenetic tree reconstruction, SOS motif discovery and comparison between phages. We collected 297 genomes of *B. subtilis* group genomes. We performed a BlastP search for all AimR homologues in these genomes using a query based on representatives of all nine AimR clades[6], yielding a total of 401 variants. Of these, 81 clustered with AimR of the *B. subtilis* subsp. *inaquosorum* ϕ106 strain and clade 1 AimRs, while the rest belonged to clades 2 and 3. We identified the *aimP* gene and its putative mature AimP signal, as well as other nearby open reading frames for all of these strains. We also verified by basic local alignment search tool (BLAST 2.9.0) the existence of a repressor homologue in an antisense direction to the *aimRP* operon. All 81 variants had a LexA binding site of high similarity to the consensus sequence between *aimR* and the repressor. The LexA binding site was identified using the algorithm provided by the DBTBS release 5 database[29]. Nine representative sequences were chosen, which had a varying AimP peptide or gene organization, including the one from *B. subtilis* subsp. *inaquosorum*. These were used to create Extended Data Fig. 7. The National Center for Biotechnology Information accession numbers for these genes are given in the legend to Extended Data Fig. 7. Comparison between phages ϕ106 and ϕ105 was done by BlastP 2.9.0 for all open reading frames as described in the legend of Extended Data Fig. 4.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data generated in this study are provided as source data with this paper. The data analysed in Extended Data Fig. 3d are available from Nicolas et al.[16].

Received: 6 July 2021; Accepted: 26 October 2021;
Published online: 9 December 2021

References

1. Ptashne, M. *Genetic Switch: Phage Lambda Revisited* 3rd edn (Cold Spring Harbor Laboratory Press, 2004).
2. Kourilsky, P. & Knapp, A. Lysogenization by bacteriophage lambda: III. Multiplicity dependent phenomena occurring upon infection by lambda. *Biochimie* **56**, 1517–1523 (1974).
3. Nanda, A. M., Thormann, K. & Frunzke, J. Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. *J. Bacteriol.* **197**, 410–419 (2015).
4. Węgrzyn, G. & Węgrzyn, A. Genetic switches during bacteriophage λ development. *Prog. Nucleic Acid Res. Mol. Biol.* **79**, 1–48 (2005).
5. Erez, Z. et al. Communication between viruses guides lysis–lysogeny decisions. *Nature* **541**, 488–493 (2017).
6. Stokar-Avihail, A., Tal, N., Erez, Z., Lopatina, A. & Sorek, R. Widespread utilization of peptide communication in phages infecting soil and pathogenic bacteria. *Cell Host Microbe* **25**, 746–755.e5 (2019).
7. Aframian, N. & Eldar, A. A bacterial tower of Babel: quorum-sensing signaling diversity and its evolution. *Annu. Rev. Microbiol.* **74**, 587–606 (2020).
8. Wang, Q. et al. Structural basis of the arbitrium peptide–AimR communication system in the phage lysis–lysogeny decision. *Nat. Microbiol.* **3**, 1266–1273 (2018).
9. Zhen, X. et al. Structural basis of AimP signaling molecule recognition by AimR in Spbeta group of bacteriophages. *Protein Cell* **10**, 131–136 (2019).
10. Trinh, J. T. & Zeng, L. Structure regulates phage lysis–lysogeny decisions. *Trends Microbiol.* **27**, 3–4 (2019).
11. Guan, Z. et al. Structural insights into DNA recognition by AimR of the arbitrium communication system in the SPbeta phage. *Cell Discov.* **5**, 29 (2019).
12. Abedon, S. T. Look who's talking: T-even phage lysis inhibition, the granddaddy of virus-virus intercellular communication research. *Viruses* **11**, 951 (2019).
13. Abedon, S. T. Commentary: communication between viruses guides lysis–lysogeny decisions. *Front. Microbiol.* **8**, 983 (2017).
14. Hynes, A. P. & Moineau, S. Phagebook: the social network. *Mol. Cell* **65**, 963–964 (2017).
15. Abedon, S. T. Bacteriophage secondary infection. *Virol. Sin.* **30**, 3–10 (2015).
16. Nicolas, P. et al. Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science* **335**, 1103–1106 (2012).
17. Au, N. et al. Genetic composition of the *Bacillus subtilis* SOS system. *J. Bacteriol.* **187**, 7655–7666 (2005).

18. Goranov, A. I., Kuester-Schoeck, E., Wang, J. D. & Grossman, A. D. Characterization of the global transcriptional responses to different types of DNA damage and disruption of replication in *Bacillus subtilis*. *J. Bacteriol.* **188**, 5595–5605 (2006).
19. Gandon, S. Why be temperate: lessons from bacteriophage λ. *Trends Microbiol.* **24**, 356–365 (2016).
20. Kawai, Y., Moriya, S. & Ogasawara, N. Identification of a protein, YneA, responsible for cell division suppression during the SOS response in *Bacillus subtilis*. *Mol. Microbiol.* **47**, 1113–1122 (2003).
21. Wood, H. E., Dawson, M. T., Devine, K. M. & McConnell, D. J. Characterization of PBSX, a defective prophage of *Bacillus subtilis*. *J. Bacteriol.* **172**, 2667–2674 (1990).
22. Babel, H. et al. Ratiometric population sensing by a pump-probe signaling system in *Bacillus subtilis*. *Nat. Commun.* **11**, 1176 (2020).
23. Czyż, A., Zielke, R. & Wegrzyn, G. Rapid degradation of bacteriophage λ O protein by ClpP/ClpX protease influences the lysis-versus-lysogenization decision of the phage under certain growth conditions of the host cells. *Arch. Virol.* **146**, 1487–1498 (2001).
24. 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).
25. 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).
26. Sudiarta, I. P., Fukushima, T. & Sekiguchi, J. *Bacillus subtilis* CwlP of the SP-β prophage has two novel peptidoglycan hydrolase domains, muramidase and cross-linkage digesting DD-endopeptidase. *J. Biol. Chem.* **285**, 41232–41243 (2010).
27. Forrest, D., James, K., Yuzenkova, Y. & Zenkin, N. Single-peptide DNA-dependent RNA polymerase homologous to multi-subunit RNA polymerase. *Nat. Commun.* **8**, 15774 (2017).
28. Bose, B., Auchtung, J. M., Lee, C. A. & Grossman, A. D. A conserved anti-repressor controls horizontal gene transfer by proteolysis. *Mol. Microbiol.* **70**, 570–582 (2008).
29. Sierro, N., Makita, Y., de Hoon, M. & Nakai, K. DBTBS: a database of transcriptional regulation in *Bacillus subtilis* containing upstream intergenic conservation information. *Nucleic Acids Res.* **36**, D93–D96 (2008).
30. Doermann, A. H. Lysis and lysis inhibition with *Escherichia coli* bacteriophage. *J. Bacteriol.* **55**, 257–276 (1948).
31. Hays, S. G. & Seed, K. D. Dominant *Vibrio cholerae* phage exhibits lysis inhibition sensitive to disruption by a defensive phage satellite. *eLife* **9**, e53200 (2020).
32. Bruce, J. B., Lion, S., Buckling, A., Westra, E. R. & Gandon, S. Regulation of prophage induction and lysogenization by phage communication systems. *Curr. Biol.* https://doi.org/10.1016/j.cub.2021.08.073 (2021).
33. Brady, A. et al. The arbitrium system controls prophage induction. *Curr. Biol.* https://doi.org/10.1016/j.cub.2021.08.072 (2021).
34. Ladau, J. & Eloe-Fadrosh, E. A. Spatial, temporal, and phylogenetic scales of microbial ecology. *Trends Microbiol.* **27**, 662–669 (2019).
35. van Gestel, J. et al. Short-range quorum sensing controls horizontal gene transfer at micron scale in bacterial communities. *Nat. Commun.* **12**, 2324 (2021).
36. Ben-Zion, I., Pollak, S. & Eldar, A. Clonality and non-linearity drive facultative-cooperation allele diversity. *ISME J.* **13**, 824–835 (2019).
37. Harwood, C. R. & Cutting, S. M. (eds) *Molecular Biological Methods for Bacillus* (Wiley, 1990).
38. McDonnell, G. E., Wood, H., Devine, K. M. & McConnell, D. J. Genetic control of bacterial suicide: regulation of the induction of PBSX in *Bacillus subtilis*. *J. Bacteriol.* **176**, 5820–5830 (1994).
39. Westers, H. et al. Genome engineering reveals large dispensable regions in *Bacillus subtilis*. *Mol. Biol. Evol.* **20**, 2076–2090 (2003).
40. Fink, P. S. & Zahler, S. A. Restriction fragment maps of the genome of *Bacillus subtilis* bacteriophage SPβ. *Gene* **19**, 235–238 (1982).
41. Koo, B.-M. et al. Construction and analysis of two genome-scale deletion libraries for *Bacillus subtilis*. *Cell Syst.* **4**, 291–305.e7 (2017).
42. Yan, X., Yu, H.-J., Hong, Q. & Li, S.-P. Cre/*lox* system and PCR-based genome engineering in *Bacillus subtilis*. *Appl. Environ. Microbiol.* **74**, 5556–5562 (2008).
43. Baym, M. et al. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS ONE* **10**, e0128036 (2015).

Acknowledgements

We thank R. Sorek, S. Pollak, J. Jones, J. van-Gestel and I. Lev for fruitful discussions and comments on the manuscript. The Eldar lab is funded by a European Research Council grant no. 724805. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

N.A., S.O.B., P.G. and A.E. were involved in the conceptualization of the study. N.A., S.O.B., S.K., P.G., A.S.-A. and A.M. performed the experiments. N.A., S.O.B., P.G., S.H. and A.E. analysed the data and formulated the theoretical predictions. N.A. and A.E. wrote the original draft. N.A., S.H., P.G. and A.E. created the figures. N.A. and A.E. edited the manuscript. S.K., E.M., K.M., V.L. and I.G. provided essential resources.

Competing interests

The authors declare no competing interests.

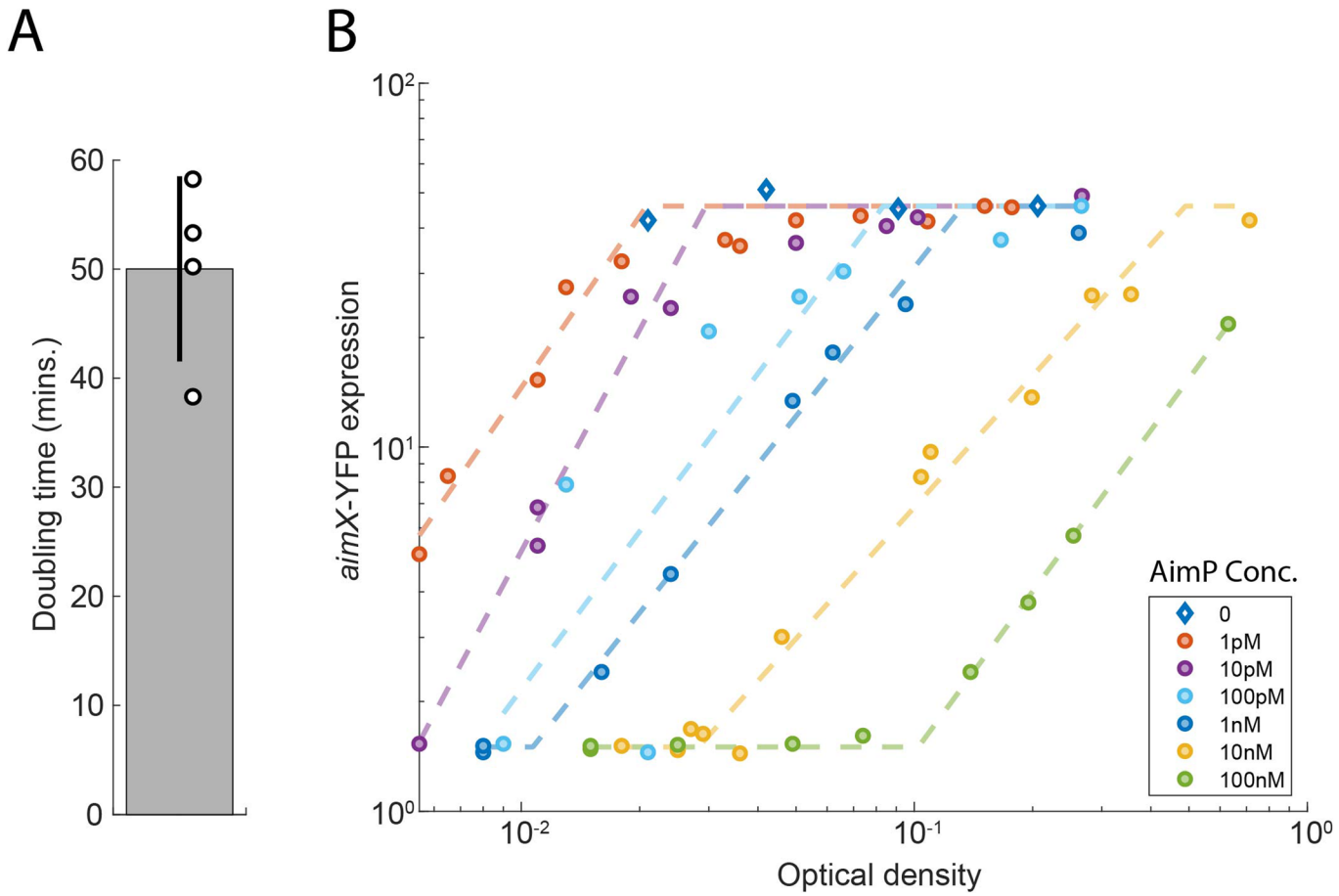
Additional information

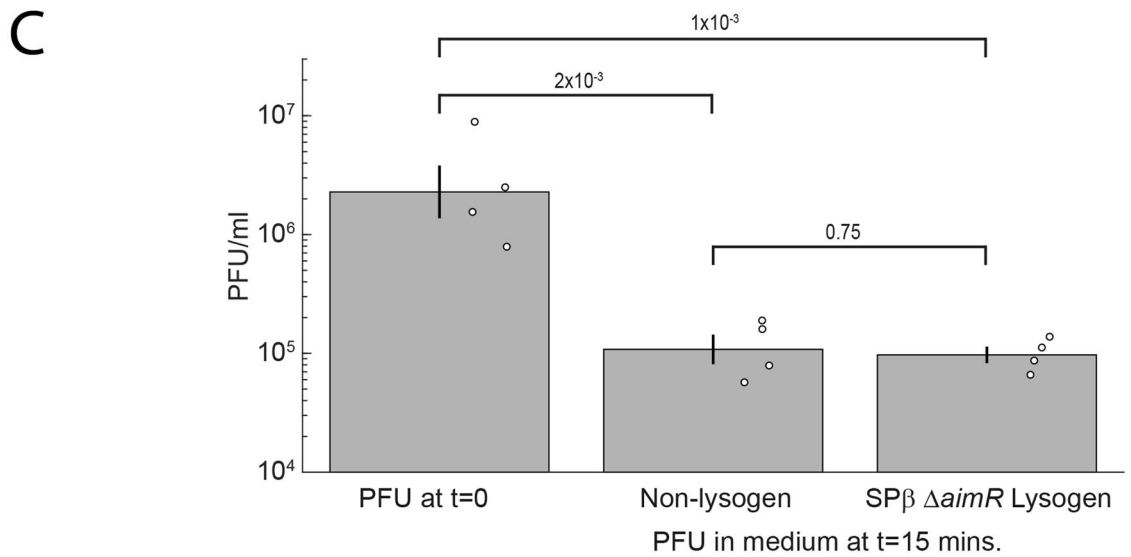
Extended data is available for this paper at https://doi.org/10.1038/s41564-021-01008-5.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-021-01008-5.

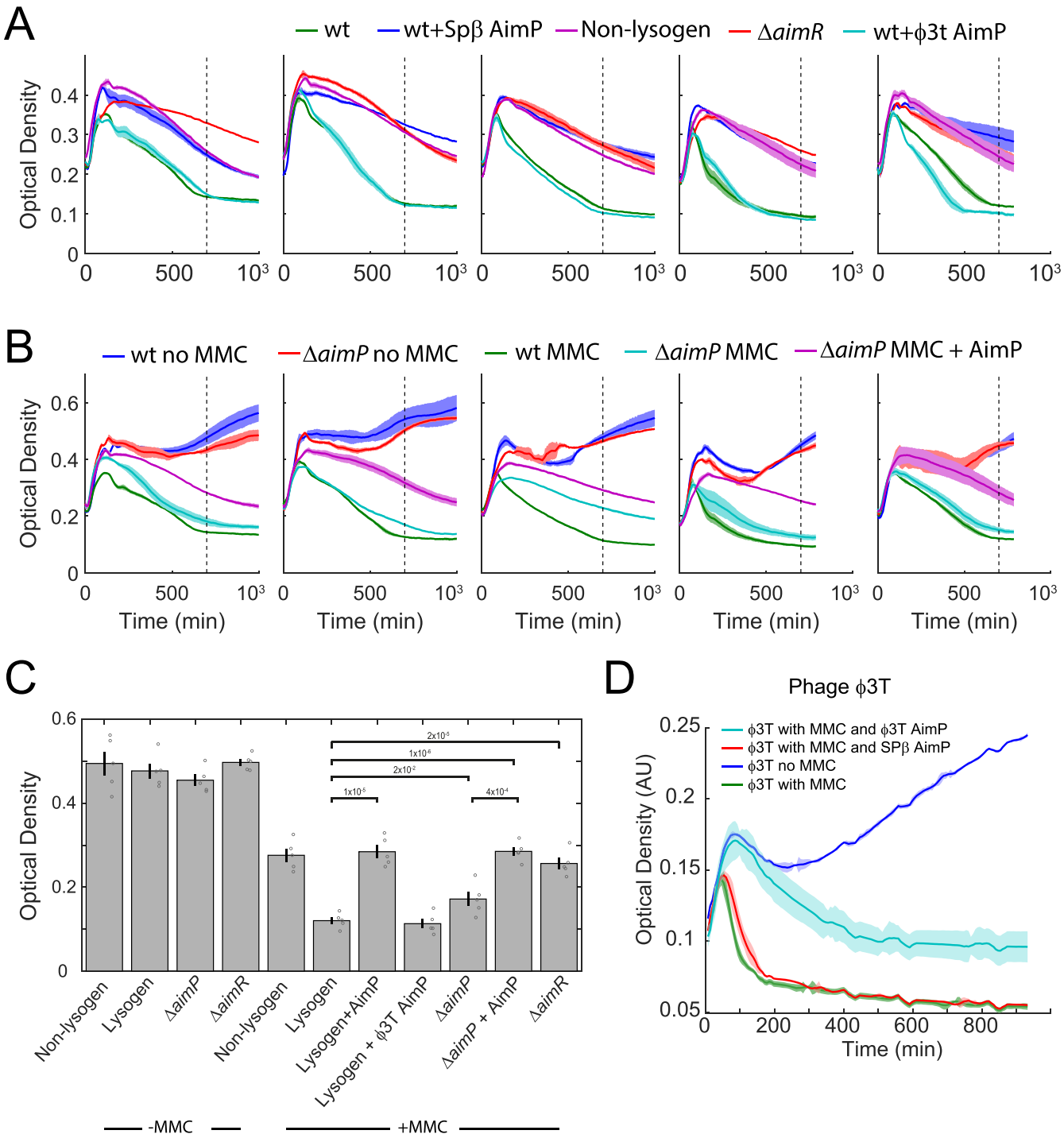
Correspondence and requests for materials should be addressed to Avigdor Eldar.

Peer review information *Nature Microbiology* thanks Joseph Bondy-Denomy and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

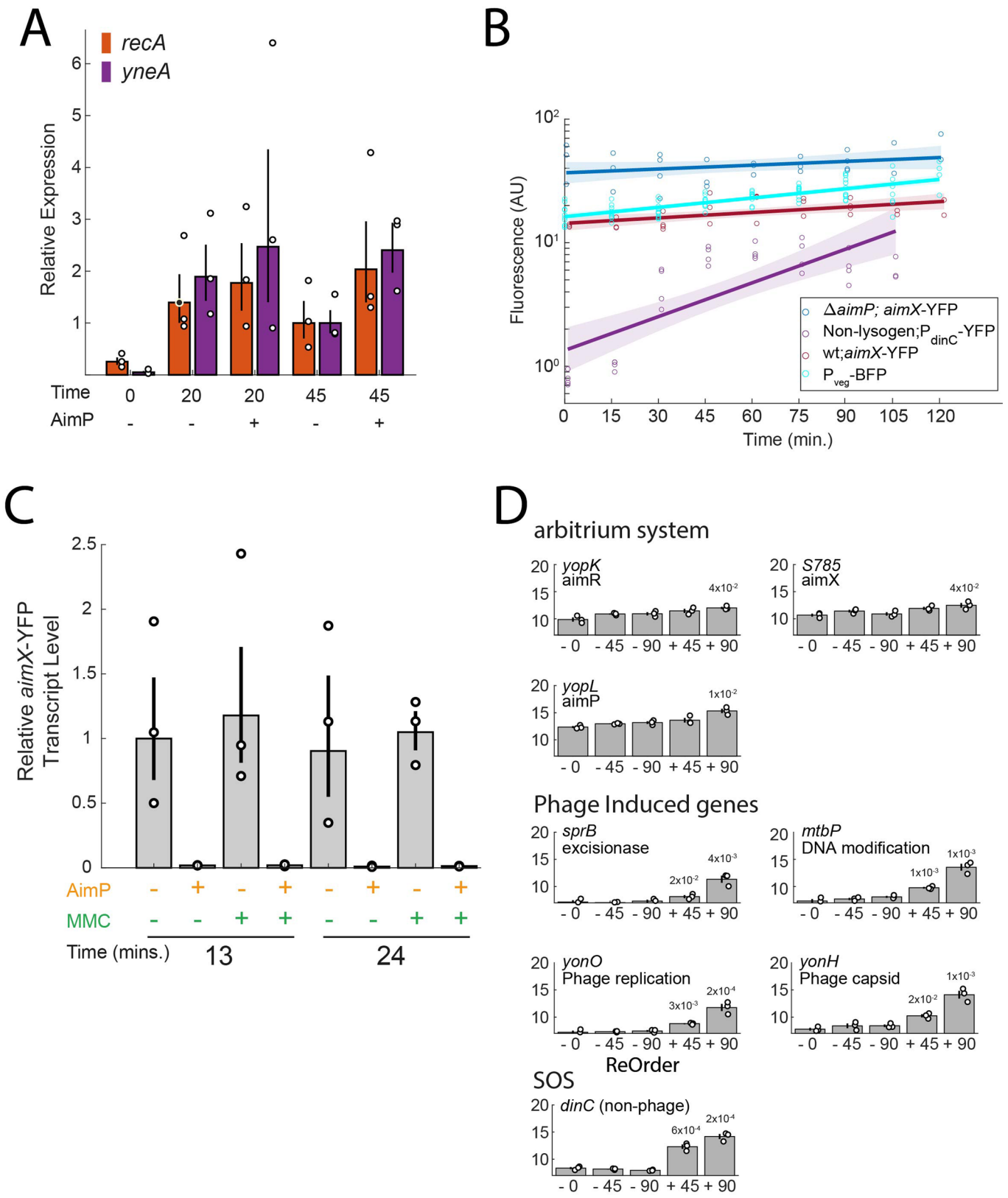




Extended Data Fig. 1 | Sensitivity of the *aimX*-YFP reporter to addition of AimP in SPβ lysogens and the fate of lysogen infection. (**a**) Doubling time in minimal medium was measured by fitting the growth curve to an exponent. Shown is the mean of doubling times for four different experiments. Error bar marks the standard error. (**b**) Lysogens coding for the *aimX*-YFP reporter were diluted 10^6-fold and grown overnight with the indicated levels of externally supplemented AimP peptide. Expression level and optical density were measured at different times. For low optical density, optical density was back-extrapolated from optical densities at later times and from the time passed by using a doubling time of 50 minutes (see A, as in Fig. 1b). Dashed lines show a tri-linear fit on the log-log scale of the data to two horizontal lines and an increasing line. The y-value of the horizontal lines of all fits is equal to the autofluorescence level (bottom line) and to the mean expression of cells with no signal (marked by diamond markers). The slope and intersect of the monotonously increasing lines are subject to fitting. (**c**) SPβ adsorption to lysogens and non-lysogens is equal. Shown are levels of SPβ phages, measured as PFU/ml, in the medium right after phage addition to medium with no cells and 15 minutes after addition of the same phage doses to medium containing either Non-lysogens or Δ*aimR* SPβ lysogens. Shown are geometric means of n = 4 biological repeats and error bars represent standard error of the log. The level of PFU is significantly lower in medium containing cells by a factor of ~30 (P-value of two-tailed paired t-test on the log of the PFU is shown in figure), but is not significantly different between the two cell types (P = 0.75, paired t-test on the log of the PFU).



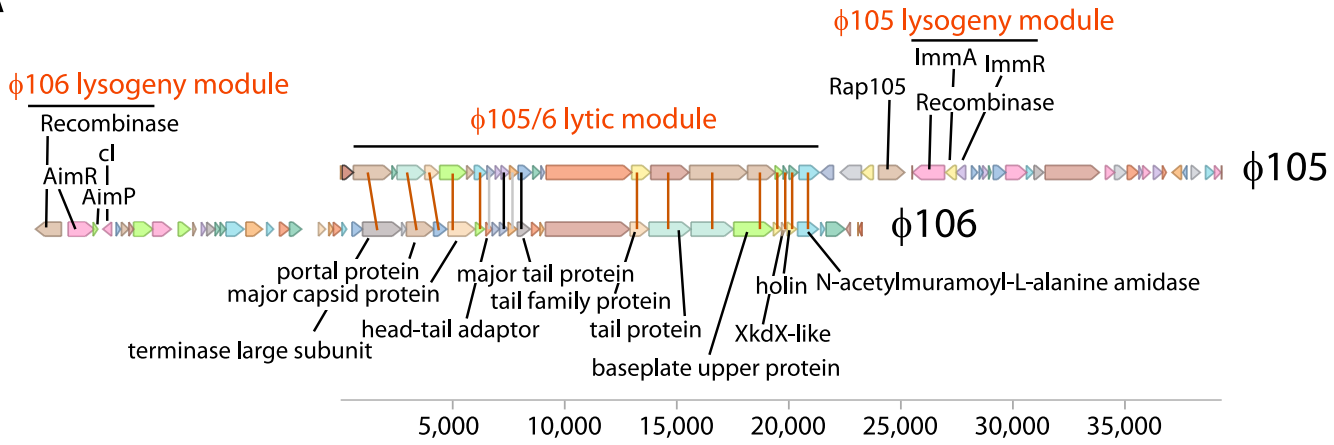
Extended Data Fig. 2 | Impact of signaling on cellular growth curves as measured by plate reader. (**a**) Further examples for plate reader experiments showing the same strains and conditions as shown in Fig. 1c of the main text. The second panel is the same as the panel shown in Fig. 1c. A dashed line marks the time-point at which OD levels were used for the statistical analysis shown in (C). (**B**) Plate reader experiments of additional strains and conditions. Shown are growth curves for the wild-type and Δ*aimP* lysogens with and without MMC, and for the Δ*aimP* lysogen with MMC and AimP. See legend for the color of the curve of each strain. (**c**) Shown are the mean and standard error of 5 biological repeats for each strain and condition, as well as individual measurements of optical density at T = 700 minutes. P-values obtained by two-tailed paired t-test Are shown in the figure. Note that all differences between strains without MMC are non-significant. (**d**) Shown are graphs of optical density as a function of time for four cultures of ϕ3T lysogens grown in LB broth, as described in the legend. Solid lines mark the mean of three replicates done on the same multi-well plate. The region around each line indicates the standard error of the mean. ϕ3T AimP is the ϕ3T arbitrium peptide (SAIRGA), and SPβ AimP is the peptide GMPRGA.



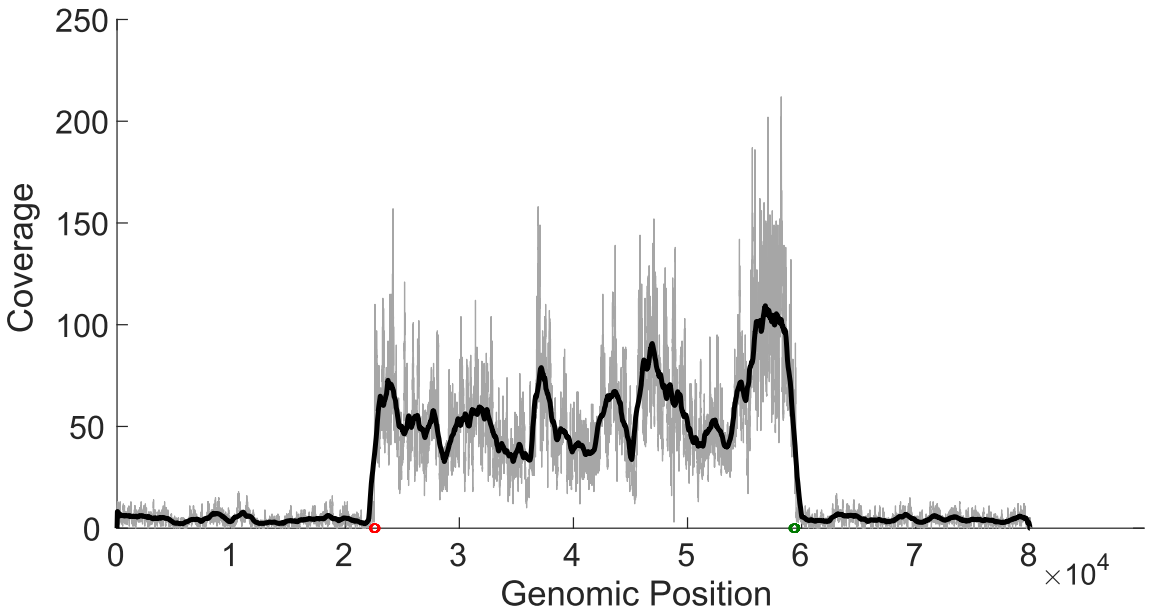
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Gene expression upon addition of MMC to strain PY79 and related strains. (**a**) qRT-PCR results for SOS genes *recA* and *yneA*. Shown are relative expression of the two genes under three different times after the addition of MMC (0,20 and 45 minutes) when 10 μM of AimP are either added or not. Expression is normalized to 45 minutes after MMC addition with no addition of AimP. Shown are geometric means of n = 3 biological repeats. Error bars correspond to standard errors of the log. (**b**) YFP and BFP expression upon addition of MMC. Shown are fluorescence levels (on a log scale) for different reporters and strains as a function of time from addition of MMC. An *aimX*-YFP reporter in wild-type (deep red) and in Δ*aimP* (blue) lysogens, a P_{dinC}-YFP reporter in a non-lysogen strain containing PBSX (purple), and a P_{veg}-BFP constitutive reporter plotted on a single curve for all the strains (cyan). Individual points are biological repeats, solid lines are linear best fit (with y data on a $\log_{10}$ scale) and shaded area is the boundary of error. (**c**) RT-PCR of *aimX*-YFP under different conditions. A Δ*aimP* SPβ lysogen carrying the *aimX*-YFP reporter was assayed for its expression using RT-PCR (methods), either with or without the addition of 0.5 μg/ml MMC and 10μM SPβ AimP. Measurements were taken either 13 or 24 minutes after the addition of MMC/AimP/Mock. Addition of MMC does not significantly alter *aimX*-YFP expression levels. Shown are geometric means of n = 3 biological repeats. Error bars correspond to standard errors of the log. (**d**) Analysis of public RNAseq data of strain 168, after addition of mitomycin C (MMC). Shown are the RNAseq expression data from[16], for selected phage genes (and one non-phage gene, *dinC*) for 5 conditions – t = 0 minutes (−0), t = 45 minutes no MMC (−45), t = 90 minutes no MMC (−90), t = 45 minutes after MMC addition (+45), t = 90 minutes after MMC addition (+90). Genes are divided into flour main categories based on their functional dependence. Gene name in the expression database and its current name or function are annotated. n = 3 biological repeats, P-values obtained by two-sided paired t-test are shown in figure. Error bars indicate standard error of the log. Specifically, note that *aimX*, *aimR* and *aimP* are not significantly different from the control 45 minutes after MMC addition and are marginally higher at 90 minutes. Note that strain 168 carries the mobile element ICEBs1 and contains several additional differences from the PY79 strain background we use in this work.

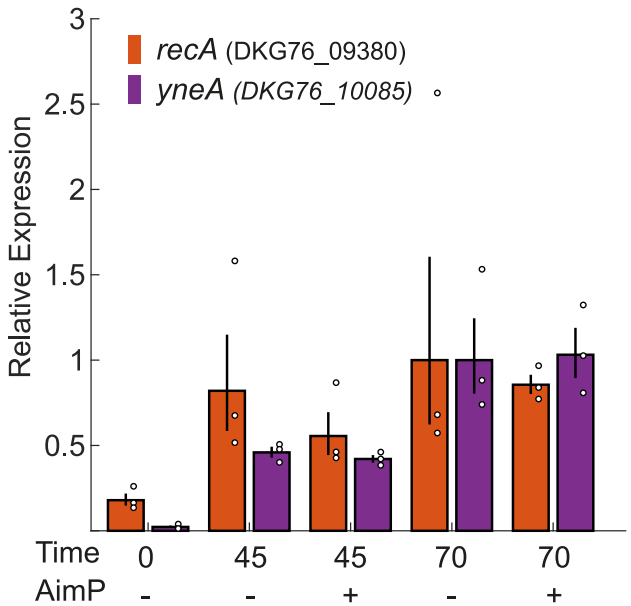
A



B

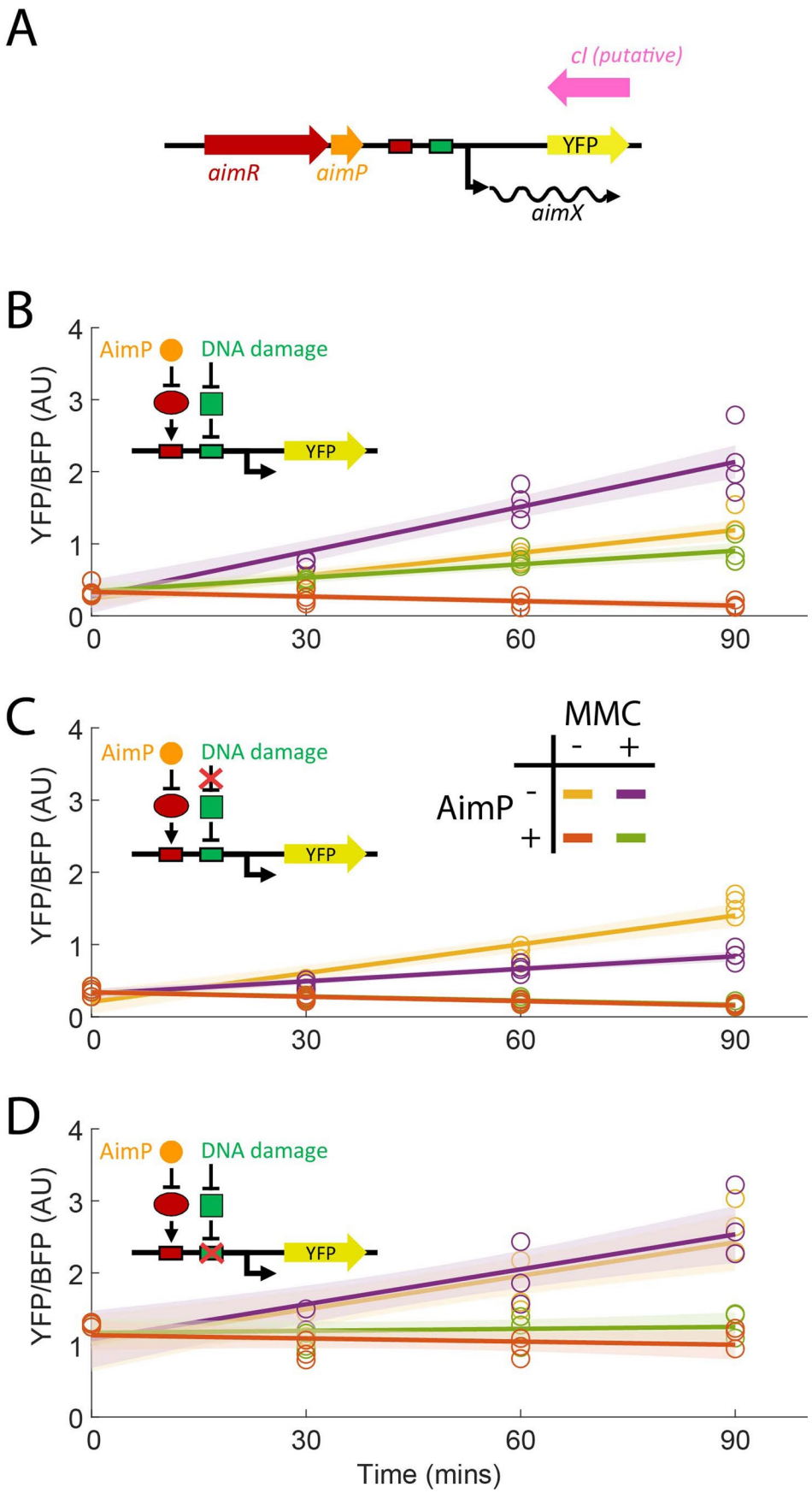


C



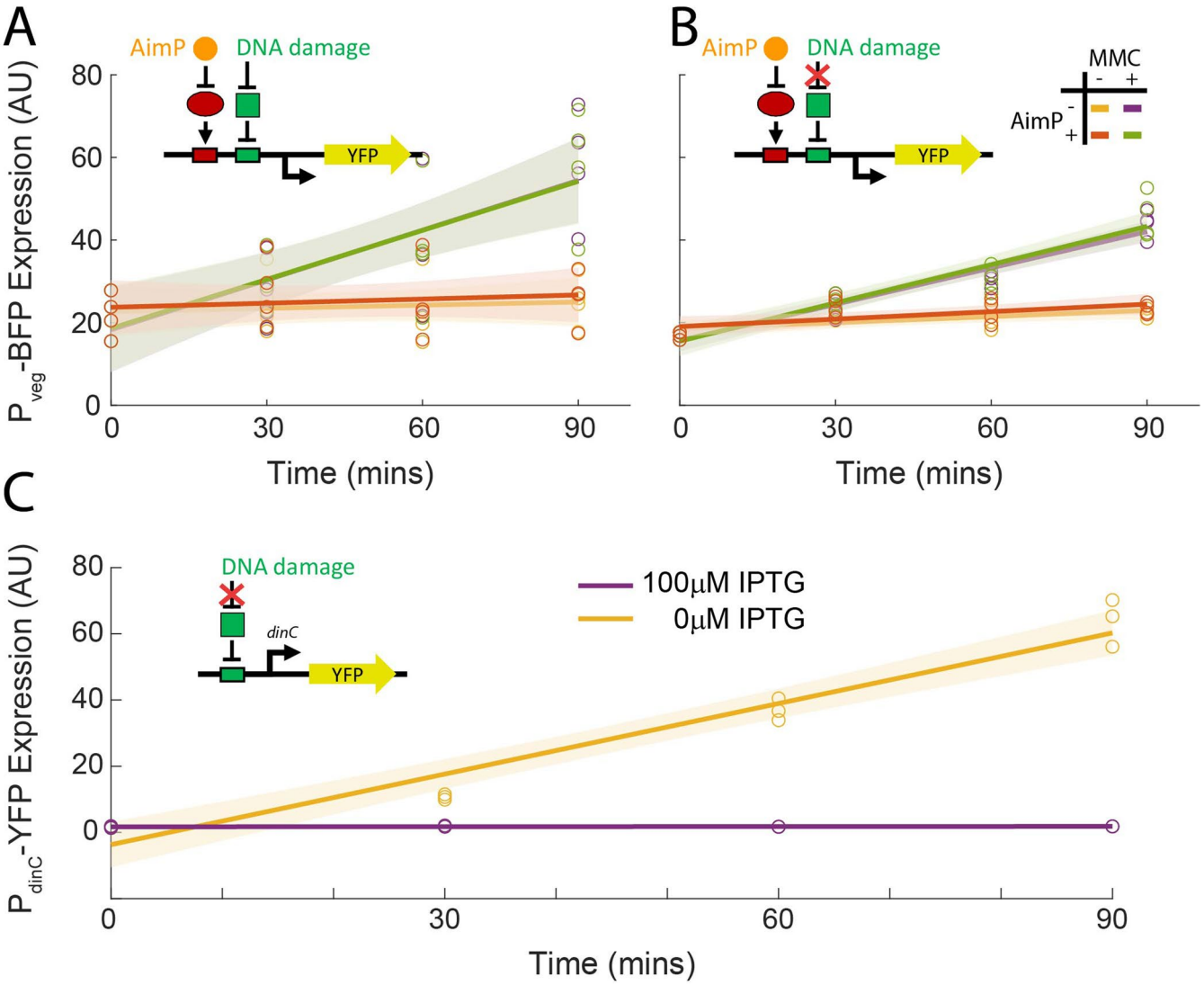
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Characteristics of phage ϕ106 and its host. (**a**) Comparison of phages ϕ105 and ϕ106. We use the following sequences for comparison. ϕ105 is analyzed using its directed sequence file at the NCBI accession number NC_048631. ϕ106 is defined here as the lysogenic segment of *Bacillus subtilis subsp. inaquosorum strain KCTC 13429* (NCBI accession number NZ_CP029465) from locus tag DKG76_14125 to locus tag DKG76_14370. ϕ106 proteins were blasted against ϕ105 proteins. Marked are proteins with high (red, Expectation<10^{-20}), medium (black, 10^{-10} < Expectation<10^{-20}) and few low homology genes within the lytic module (gray, 10^{-5} < Expectation<10^{-10}). There is little or no homology out of the region annotated here as the lytic module. Protein annotation within the lytic module is based on phage ϕ105 annotation. The lysogenic module of phage ϕ105 including the recombinase $ImmRA_{105}$ and Rap gene are shown[28]. Homologs of these genes are absent from ϕ106. Also shown is the putative lysogenic module of ϕ106 analyzed in this work, which is absent from phage ϕ105. (**b**) Coverage statistics of ϕ106 genomic region in paired-end deep sequencing data of DNAase-protected DNA from lysate of *Bacillus subtilis subsp. inaquosorum strain KCTC 13429*. ~34,000 pairs of reads matched into the shown region (positions 2,820,000 to2,900,000 within this strain's genome). A shift in coverage occurs immediately at the beginning and end of the phage. Shown are the coverage (gray) and smoothened coverage over a moving window of 1kbp (black). 72 pairs of reads matched two different sides of the phage, suggesting that the phage genome was excised and circularized to form the attP site. Their locations are shown as green and red dots. (**c**) qRT-PCR results for SOS genes *recA* and *yneA*. Shown are relative expression of the two genes of *B. subtilis subsp. Inaquosorum KCTC 13429* under three different times after the addition of MMC (0,45 and 70 minutes) when 10 μM of AimP are either added or not. Shown are geometric means of n = 3 biological repeats. Error bars correspond to standard errors of the log. Expression is normalized to 70 minutes after MMC addition with no addition of AimP.

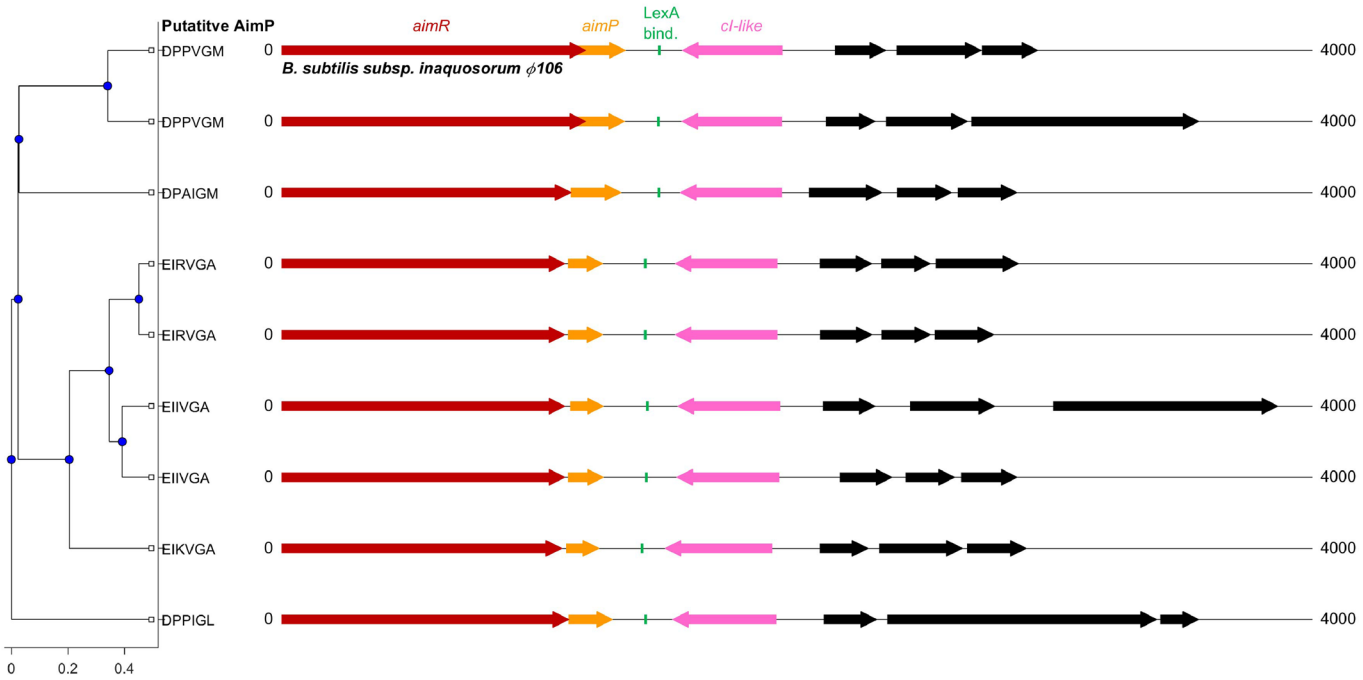


Extended Data Fig. 5 | See next page for caption.

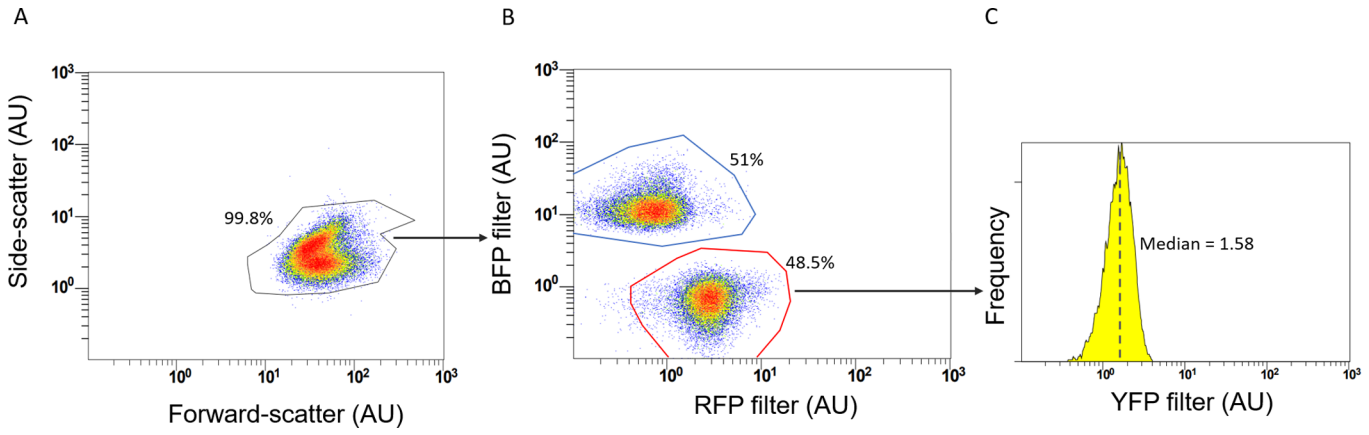
Extended Data Fig. 5 | *aimRPX*$_{\phi106}$-YFP construct and its expression as a function of time in three relevant genetic background. (**a**) schematic structure of the *aimRPX*$_{\phi106}$-YFP reporter. The YFP genes starts where the putative cI gene ends in the native genome of the phage. (**b**-**d**) YFP expression from the *aimRPX*$_{\phi106}$-YFP construct as a function of time after addition of MMC and/or AimP (or mock additions), divided by fluorescence of a P_{veg}-BFP constitutive reporter within the same cells. Circles represent individual measurements (median YFP level from a flow cytometry measurement). Lines of the same color show linear best fit and shaded area show the boundary of error. Each strain was measured with three or four independent time series taken on different days. The legend in (C) is true for all panels. The three panels correspond to the same strains described in Fig. 3c-e of the main work and the illustrations describing them are identical; (B) *aimRPX*$_{\phi106}$-YFP reporter based on the wild-type sequence of phage ϕ106. (C) as in (A) but with overexpression of a non-cleavable defective lexA mutant. (D) a mutant *aimRPX*$_{\phi106}$-YFP reporter with substitution in two base-pairs within the LexA binding site. Figure 3c-e shows the data presented here for 90 minutes.



Extended Data Fig. 6 | P_{veg}-BFP and P_{dinC}-YFP expression upon addition of MMC with and without overexpression of a non-cleavable LexA mutant. (**a**, **b**) Expression as a function of time of a constitutive P_{veg}-BFP reporter for the four conditions shown also in Supplementary Fig. 5. The two strains are the ones presented in panels A, B of Supplementary Fig. 5 correspondingly. Note the change in BFP expression upon addition in MMC between panels A, B, indicating the impact of LexA(ind-) allele expression on growth (**c**) Expression of a P_{dinC}-YFP reporter as a function of time after addition of MMC in a genetic background including the IPTG-inducible P_{hs}-lexA(ind-) allele. Shown are results with 0 IPTG (no induction, orange) and 100μM (full induction, purple). Circles represent individual measurements, lines represent best linear fits and shaded area the boundary of error.



Extended Data Fig. 7 | Consistency and diversity of the ϕ106 family putative lysogeny module. Left: Shown is the phylogenetic tree of 8 different AimR clade 1 proteins (out-grouped by AimR from phage SPβ). Each of the eight corresponding *aimR* genes has an adjacent *aimP* gene coding for the putative mature arbitrium AimP signal shown to the right of the corresponding tree leaf. Right: the genomic organization of the arbitrium locus is shown for each AimR gene. Colored arrows represent genes, according to the annotation at the top. Black arrows correspond to additional genes. Green line marks the position of the putative LexA binding site. The ϕ106 system used in this work is specifically marked (strain *inaquosorum*). The NCBI accession number of the AimR proteins (and the genome GCF number) as shown in the phylogeny, from top to bottom, are: WP_003237457 (003148415), WP_101605414 (002850535), ARW33040 (001747445), WP_049627412 (004119735), WP_047936208 (001023595), WP_003220312 (005218185), API45091 (002982175), WP_017417251 (003860445), WP_181217268 (013620725).



Extended Data Fig. 8 | Raw flow cytometry data and gating procedure. YFP reporters were used to measure *aimX* expression in SPβ (Fig. 1a,b, Extended Data Figs. 2b and 4b) and ϕ106 phages (Figs. 2b, 3c-e, Extended Data Fig. 5b-d). Populations were marked with either BFP or RFP constitutive reporters. Constitutive reporter expression allowed us to distinguish between population in co-cultures (Fig. 1a, b) and to normalize for constitutive reporter expression (Figs. 2b, 3, Extended Data Fig. 5b-d). The following gating scheme was used: **a**. cells were initially gated on their forward and side-scattering to exclude non-canonical elements. **b**. cells were then gated on their BFP or RFP. **c**. finally, YFP expression histograms and medians were calculated for an individual population. For Figs. 2b, 3c-e and Extended Data Fig. 5b-d, median of YFP expression was divided by the median of constitutive BFP expression.

nature research

Corresponding author(s): Avigdor Eldar

Last updated by author(s): Oct 16, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Flow cytometry data were collected with Beckman-Coulter Gallios
Data analysis	Numerical modeling and statistical analysis were performed using Matlab R2019a; RT-qPCR results were analyzed using Step One™ v2.3 software

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data generated during this study are included in this published article (and its supplementary information files).
Data analyzed in extended figure 3D is available at reference 16 in the paper.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	At least 3 biological repeats were used for all experiments.
Data exclusions	No data was excluded
Replication	We performed both technical and biological (at least 3 for every experiments) repeats to ensure replication. Biological repeats were performed at separate dates. All attempts at replication were successful.
Randomization	Randomization is not standard for the experiments performed.
Blinding	Blinding is not standard for the experiments performed.

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*

Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access & import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	*Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Validation	*Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.*

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	*State the source of each cell line used.*
Authentication	*Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.*

Mycoplasma contamination	*Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.*
Commonly misidentified lines (See ICLAC register)	*Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

Palaeontology and Archaeology

Specimen provenance	*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*
Specimen deposition	*Indicate where the specimens have been deposited to permit free access by other researchers.*
Dating methods	*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	*For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☐	☐	Public health
☐	☐	National security
☐	☐	Crops and/or livestock
☐	☐	Ecosystems
☐	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☐	☐	Demonstrate how to render a vaccine ineffective
☐	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☐	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☐	☐	Increase transmissibility of a pathogen
☐	☐	Alter the host range of a pathogen
☐	☐	Enable evasion of diagnostic/detection modalities
☐	☐	Enable the weaponization of a biological agent or toxin
☐	☐	Any other potentially harmful combination of experiments and agents

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	*For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*
Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	*Describe the experimental replicates, specifying number, type and replicate agreement.*
Sequencing depth	*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*
Antibodies	*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Peak calling parameters	*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*
Data quality	*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*
Software	*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Samples of bacteria grown in LB or SMM media were diluted into DDW and then measured by flow cytometry.
Instrument	Beckman-Coulter Gallios flow-cytometer B25062AA
Software	Kaluza for Gallios software
Cell population abundance	The cells were not sorted. Population of at least 10,000 cells were used to calculate medians.
Gating strategy	Cell population was gated using side scatter and forward scatter parameters and non-cell events were excluded. constitutive reporters of BFP and RFP were then used to separate co-cultured genotypes. Genotypes formed clearly distinct populations that were separated by eye, as shown in the exemplifying flow cytometry figure. Median YFP levels were then measured for each sub-population.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring

Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference
(See Eklund et al. 2016)

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.