**Arms race coevolution between bacteria with CRISPR-Cas immune systems and a phage**

Dan Morley1,2, Jack Common1,2, Edze Westra1, Stineke van Houte1\*

1ESI and CEC, Biosciences, University of Exeter, Cornwall Campus, Penryn TR10 9EZ, UK

2contributed equally

\* correspondence: C.van-Houte@exeter.ac.uk

**Abstract**

**Introduction**

Clustered Regularly Interspaced Short Palindromic Repeats and their associated *cas* genes (CRISPR-Cas) encode an adaptive immune system that is found in approximately 50% of all bacteria and 90% of archaea (ref). CRISPR-Cas confers immunity to phage infection by incorporating phage-derived sequences into CRISPR loci on the host genome. These loci consist of repeating sequences (“repeats”) which are interspaced by phage-derived sequences (“spacers”), typically around 30nt in length. RNA transcripts of CRISPR loci are processed and form a ribonucleoprotein complex with Cas proteins that can recognise and cleave complementary nucleic acid sequences, preventing future infections by the same phage genotype. CRISPR-Cas systems are highly diverse, and are currently classified into 2 classes, 6 types and 33 subtypes based on their *cas* gene composition, gene synteny and CRISPR repeat sequences, with clear differences in the molecular mechanisms of different variant systems ([5](#_ENREF_5)).

In some natural environments, bacteria with CRISPR-Cas systems and their phages appear to coevolve over long time periods (Banfield; Price? ; Sundberg). However, studying the dynamics of these coevolutionary interactions under controlled laboratory conditions has been difficult due to the scarcity of available model systems. Specifically, while many bacteria encode CRISPR-Cas immune systems, the vast majority of these bacteria do not evolve CRISPR-based immunity upon phage or plasmid infection, or do so at such low frequencies that they are detectable only with deep-sequencing approaches and unlikely to significantly contribute to the reciprocal selection between the host and the parasite (refs). Currently only two bacterial species have been found to naturally evolve high levels of CRISPR immunity under laboratory conditions, namely *Streptococcus thermophilus* strains DGCC7710and LMD-9 (Barrangou 2008; Horvath 2009), and *Pseudomonas aeruginosa* strain UCBPP-PA14 (Westra Curr Biol).

Early studies with *S. thermophilus* demonstrated that phage can overcome CRISPR immunity through the evolution of point mutations in the sequence targeted by the spacers (the so-called protospacer). Point mutations in the protospacer-adjacent motif (PAM) (Deveau JBac 2008), a conserved sequence immediately adjacent to the protospacer that is used by the host to discriminate between self (i.e. CRISPR arrays) and non-self (e.g. phage) DNA, can also confer escape from CRISPR immunity ([Mojica *et al.* 2009](#_ENREF_17); [Semenova *et al.* 2011](#_ENREF_19)) (Westra PLoS Genet 2013). The observation that phage can evolve to ‘escape’ CRISPR immunity suggested a possible scenario for coevolution in free-running systems, where hosts accumulate spacers over time and phages accumulate escape mutations in the corresponding protospacers or PAMs (refs). Consistent with the idea of CRISPR-phage coevolution, it was reported that *S. thermophilus* and phage can coexist for long periods of time, and that for each treatment the single experimental population increased its bacterial spacer repertoire and the frequency of point mutations in phage genomes over time (ref. Paez-Espino 2x, Sun ISME). However, recent modelling and experiments suggest that the observed coexistence of *S. thermophilus* and its phage may be driven by back mutation of hosts with CRISPR immunity to sensitive phenotypes, which would provide a reservoir for phage amplification and therefore allow their persistence (Weissman ISME 2018). Such loss of CRISPR immunity due to mutation has also been observed at high frequencies in *Staphyloccus epidermidis* (Jiang PLoS Genetics 2013), and may more generally be an important mechanism for bacteria-phage coexistence (Levin PLoS Biol in press). Studies with the *P. aeruginosa* PA14 model system also support the idea that a continuous supply of sensitive hosts is required for bacteria-phage coexistence, and in this empirical system phages were unable to coevolve with CRISPR-immune bacteria due to the high levels of CRISPR spacer diversity that bacteria naturally evolve (van Houte Nature; Chabas ProcB; Westra NRI).

Given the lack of clarity about the role of CRISPR-phage coevolution for bacteria-phage coexistence, we performed highly replicated long-term (30-day) serial transfer experiments with *S. thermophilus* and its lytic phage φD2972. Our phenotypic assays demonstrate that bacteria and phages coevolved in these experiments, and that host resistance and phage infectivity increased over time. Furthermore, our genotypic data show that patterns of resistance and infectivity were explained by bacteria acquiring novel spacers against the phage, and the phage evolving mutations in the regions targeted by the spacers. Collectively, the data show that bacteria-phage coexistence can be explained by coevolution in this system, and that this coevolution is characterized by an arms-race dynamic.

**EXPERIMENTAL METHODS**

**Strains used in the study**

We used the lactic acid bacterium *Streptococcus thermophilus* DGCC7710 (WT) and its lytic phage D2972 as a model system. DGCC7710has four CRISPR-Cas systems, two of which (CRISPR1 and CRISPR3) are active during infection with phage D2972 and both are classified as Type II-A (Carte et al Mol Micro 2014)([Horvath *et al.* 2008](#_ENREF_12)).

**Phage D2972 amplification**

An overnight culture of *S. thermophilus* was transferred 1:10 into fresh LM17 medium (M17 broth supplemented with 0.5% α-Lactose) containing 10mM CaCl2 and incubated shaking at 180 rpm at 42 oC. When the culture reached log phase (OD600 ~ 0.25) approximately 106 plaque forming units (pfus) of φD2972 were added and the culture was incubated under the same conditions for two hours, at which point cells had fully lysed. Lysates were centrifuged and filtered through a 0.22-µm filter, and the resulting phage stocks were stored at 4 oC.

**Coevolution experiment**

Prior to commencing the experiment, *S. thermophilus* was acclimatised in LM17 medium at 42 oC and 180 rpm for 2 days, with a 1:100 transfer into fresh LM17 after 24 hours. To start the coevolution experiment, bacteria were transferred 1:100 into LM17 media supplemented with 10mM CaCl2 and infected with either 109, 108, 107 or 106 pfus of phage D2972, with each phage treatment run in 12 independent replicate experiments, and then incubated at 42 oC shaking at 180 rpm.

Replicates were transferred 1:100 into fresh LM17 supplemented with 10mM CaCl2 every 24 hours and phage titres and bacterial densities were measured every 24 hours for a period of 30 days, or until no phage was detected for four consecutive days. Bacterial densities were measured through OD600, while phage densities were measured by plaque assays. These were performed by mixing phage dilutions with WT bacteria in soft agar overlays (consisting of LM17 + 10mM CaCl2 and 0.5% agar), poured onto hard agar (LM17 + 10mM CaCl2 and 1.5% agar).

**Time-shift assays**

We conducted time-shift assays using phage and bacteria extracted each day from the 108 pfu phage treatment of the coevolution experiment to determine how phage infectivity and host resistance changed over time. Eight of the 12 replicates in this treatment were used for this analysis because phage persisted in these replicates for at least nine days.

Phage extracted from 1, 4 and 9 days post-infection (dpi) were subjected to plaque assays as described above. For each replicate, twelve plaques were randomly picked and amplified in 96 well plates containing LM17 + 10mM CaCl2 in which WT bacteria were inoculated 1:100 from a fresh overnight culture. Bacteria extracted at the same time points were diluted and plated overnight, and twelve colonies from each plate were picked at random and used to make lawns on M17 agar plates supplemented with 1% CaCl2.

To test for infectivity, the 36 phage extracted from each replicate (12 phage x 3 time points) were stamped on the 36 lawns made from their sympatric hosts so that each phage was tested against its sympatric host from the past, present and future (Table 1). Phage were classed as infective against a host if a plaque was visible on the lawn after incubation at 42 oC for 24 hours. If no plaque was visible, the host was classed as resistant.

**STATISTICAL METHODS**

Statistical analyses were carried out in R v3.5.0 ([Team 2014](#_ENREF_20)), and graphics were generated using r-base and the ggplot2 package ([Wickham 2009](#_ENREF_24)). Model selection followed a nested design, and the final models in all analyses were selected based on the reduction of heteroskedacity, χ2 tests, and log-likelihood and AIC comparisons ([Akaike 1973](#_ENREF_3); [Burnham & Anderson 2003](#_ENREF_6), [2004](#_ENREF_7)).

**Sequence analysis**

Twelve clones from each replicate at each of the three time points had both CRISPR loci (CRISPR1 and CRISPR3) analysed using PCR to determine spacer acquisition. Clones identified as carrying spacers then had the locus of interest sent for Sanger sequencing (Source Bioscience, UK). Both loci were sequenced using primers [CR1 primer sequence] and [CR3 primer sequence], and the entire sequenced locus was mapped against the φD2972 genome (Accession: NC 007019.1) using BLAST. Putative spacer sequence hits were approximately 30 bp in length. The genomic location and read direction of each protospacer hit was parsed using Geneious v9.1.8 ([Kearse *et al.* 2012](#_ENREF_14)). False positive differences in hit location between putative spacer sequences generated by the BLAST algorithm were then identified, and the data was cleaned accordingly.

**Phage survival**

Phage survival over the course of the experiment was analysed using a Cox proportional hazards model from the survival package ([Therneau & Lumley 2015](#_ENREF_21)). Hazard ratio coefficients express the relative risk of phage extinction over time.

**Evolution of infectivity and resistance**

To measure the evolution of phage infectivity in terms of host range, we first calculated the proportion of all host genotypes infected by a given phage genotype in each replicate at each time point. To measure the evolution of host resistance, the proportion of all phage genotypes resisted by a given host genotype was also calculated. Infectivity or resistance was analysed in a Generalised Linear Model (GLM) with genotype as a fixed effect and a binomial family with a logit link function. Mean infectivity or resistance was then analysed for each time point in a Generalised Linear Mixed Model (GLMM) using the lme4 package ([Bates *et al.* 2014](#_ENREF_4)), with time point as a fixed effect and replicate as a random effect. Model coefficients and confidence intervals were transformed from logits to probabilities prior to presentation.

**Coevolutionary dynamics**

We used data from the time-shift assays to infer whether coevolution had taken place. Infectivity was measured as the proportion of successful infections, with host resistance necessarily being 1- this value. Infectivity or resistance was analysed in a GLMM with host background (Environment; E) as a fixed effect and the interaction between host background and phage genotype (Genotype X Environment; GxE) as a random effect. Models had a binomial family with a logit link function, with coefficients and confidence intervals transformed to probabilities prior to presentation.

To test for the relative importance of arms race versus fluctuating selection in our experiment, we estimated the strength of the GxE effect on infectivity and resistance following [Hall *et al.* (2011)](#_ENREF_11). Under a simple arms race, phage should be more infective to hosts from their past compared to their present or future, independent of phage genotype. By contrast, under fluctuating selection phage genotypes will differ in their infectivity to hosts from their past, present or future. Differences among phage genotypes are therefore detectable as the proportion of the host environment (E) residual variance explained by the interaction between host environment and phage genotype (GxE). Increasing values of this proportion relate to increasing differences among phage genotypes. We estimated this by calculating the ratio of the mean square (MS) of an E-only model to the MS of a GxE model for each replicate at each time point. Ratios were then analysed in a GLMM with time point as a fixed effect and replicate as a random effect, with a normal family and identity link function. Residuals were square root transformed to fit the assumption of normality.

**RESULTS**

*S. thermophilus* DGCC7710 encodes four CRISPR-Cas systems, each with a single CRISPR locus (CRISPR1-4). Upon phage D2972 exposure, acquisition of novel spacers can be readily observed in CRISPR1 (most active locus) and CRISPR3, whereas CRISPR4 and CRISPR2 are virtually inactive in this context (Paez-Espino Nature Comm, Paez-Espino Mbio; Horvath 2008; Mills et al J Appl Micro 2010). A previous co-culture experiment showed that *S. thermophilus* strain DGCC7710 and its lytic phage D2972 can coexist for approximately 35 days, during which bacteria can acquire novel spacers and phage can acquire point mutations (Paez-Espino MBio). Due to a lack of experimental replication it has remained unclear how representative this observed period is of coexistence and the patterns of bacterial and phage evolution.

We set out to first examine the generality of the previously reported population dynamics following infection of *S. thermophilus* DGCC7710 with a single phage D2972. We therefore infected 12 replicate experimental populations of *S. thermophilus* DGCC7710 with either 106, 107, 108 or 109 plaque forming units (pfu) of phage D2972 (i.e. 48 populations in total), and monitored the bacterial and phage population densities on a daily basis for 30 days. For the first three days following infection, phage titres remained fairly constant in most replicates between 106-108 pfus ml-1, with the exception of the highest phage treatment (109) where phage (and bacteria) went extinct in 11 out of 12 replicates (**Figure 1**). Describe levels of within and between treatment variation here. There was a significant negative relationship between phage titre and host density (*z*= -12.49, DF=594, *p <* 0.0001). While there was a significant interaction effect of initial phage titre on this relationship (*F*(3,591)= 6.07, *p* < 0.001), only the 10-9 phage treatment had a significantly stronger negative slope compared to the others (*z* = -3.03, DF=591, *p* < 0.01). At 16 days post infection (dpi), the phage had gone extinct in 44 / 48 replicates, and phage persisted for the entire 30-day duration of the experiment in only two replicate experiments. For the treatments where bacteria survived, mean time to phage extinction (days) was: 109 – *M(SE*) = 2(0.54); 108 – *M(SE)* = 11.50(1.77); 107 – *M(SE)* = 11.50(2.12); 106 – *M(SE)* = 7.67(1.67).

We first aimed to determine whether the coexisting bacteria and phages had evolved during these experiments, with the ultimate aim of determining whether they coevolved. To analyse this, we isolated bacteria and phages from samples taken at 1, 4 and 9 dpi. Specifically, we isolated 12 bacterial clones and 12 phage clones per time point from each replicate experiment where bacteria and phages coexisted for at least nine days. However, since the population dynamics associated with the 106-108 pfus infection regimes was virtually identical, we decided to limit our downstream analyses to the 108 treatment only, resulting in a total of 8 replicate populations that were examined in detail (see **Figure 1**). Using the 288 bacterial clones and 288 phage clones that were isolated from these replicate experiments, we first examined whether the bacteria and phage had evolved increased resistance and infectivity over time. This was done by measuring resistance of each individual bacterial clone against all phage clones derived from the same replicate experiment, and measuring infectivity of each individual phage clone against all bacterial clones from the same replicate experiment. This analysis revealed that mean phage infectivity, i.e. the proportion of all host genotypes that can be infected by a given phage genotype, increased significantly from 0.29 (CI = 0.08, 0.48) at 1 dpi to 0.57 (0.37, 0.74) at 4 dpi, but stayed statistically similar from 4 to 9 dpi (*M*=0.53, CI=0.33, 0.74). Mean host resistance, i.e. the proportion of all phage genotypes resisted by a given host genotype, increased significantly each day, from 0.01 (0.00, 0.05) at 1 dpi to 0.67 (0.18, 0.96) at 4 dpi, and to 0.99 (0.96, 0.99) at 9 dpi (**Figure 2**). Collectively, these data show that bacteria evolved to resist essentially all phage genotypes by 9 dpi, but phage did not evolve high levels of infectivity to match.

Next, we wanted to verify that acquisition of novel spacers formed the mechanistic basis of bacterial resistance evolution during these experiments. We therefore performed PCR analysis on CRISPR1 and CRISPR3 of *S. thermophilus* DGCC7710, since these loci have previously been shown to acquire novel spacers under similar conditions, whereas acquisition of spacers targeting phage D2972 into CRISPR2 and CRISPR4 has not been observed (Horvath 2008; Paez-Espino MBio). This revealed that timepoint had a significant effect on the mean number of spacers acquired by a given clone (χ2(6,1140)=32.9, *p*<0.0001)(Figure 3A), with all clones acquiring at least one spacer by 9 dpi and most acquiring two (M=0.55, CI = 0.45, 0.65)(Figure 3B). Also correlate with resistance data (**Figure 3**). These data demonstrate that all clones that had acquired resistance also had acquired at least one novel spacer in either CRISPR1 or CRISPR3, suggesting that resistance is CRISPR-mediated. Furthermore, Sanger sequencing of all CRISPR amplicons confirmed that all spacers that had been acquired indeed targeted the phage D2972.

Using these sequencing data, we determined the level of spacer diversity that naturally evolved within each replicate, since this is an important determinant of CRISPR -phage coevolution (Childs PLoS One, van Houte Nature). Consistent with deep sequencing analyses of a previous experiment (Paez-Espino Nature Comm; Paez Espino MBio), our data showed that mean pairwise difference (PWD) among spacer sequences was generally low (0.25). Although there was no statistically notable change in mean PWD with timepoint (χ2(2,13)=3.36, *p*=0.18), this was likely due to a lack of sensitivity in our analysis. Additionally, replicate as a random effect explained 63.90% of the model’s conditional *R*2 (Conditional *R*2=0.52, marginal *R*2=0.18), and there was clear qualitative variation in spacer diversity among replicates (Figure 3C). Mean genotype richness was also low (1 dpi = 1, 4 dpi = 1.5, 9 dpi = 2.25)(Supplementary Fig?). The low diversity becomes especially apparent when the spacers are mapped against the phage genome (**Figure 4**), and also shows that the spacer composition between time points can change dramatically, suggestive of selective sweeps.

Having established that bacteria evolved CRISPR resistance and phage evolved increased infectivity over time in the replicate experiments where they coexisted, we next wanted to examine if they coevolved (i.e. whether bacteria evolve resistance against evolving phage, and vice versa), and if so, what type of dynamics was associated with this coevolution. A classical phenotypic assay to probe this question are time shift assays (refs Hall, Gandon, …), where one specifically measures the resistance of bacteria against phage from the past, present or future and vice versa, the infectivity of phage against bacteria from the past, present and future. Because individual genotypes may differ in their response to time-shift challenges, generalized linear mixed models (GLMMs) with host or phage genotype as a random effect and phage or host background (see Table 1) as a fixed effect were used to analyse time-shift data. Models of infectivity showed that phage were most infective to hosts from their past, less infective to contemporaneous hosts, and least infective to hosts from their future (Figure 5A) (χ2(4,10044)=5.35, *p*<0.0001)..Turning to host resistance, hosts were overall most resistant to phage from their past, less resistant to contemporaneous phage, and least resistant to future phage (Figure 5B)(χ2(4,10044)=16.82, *p*<0.0001), with similarly identical responses among host genotypes.

While infectivity declined overall from past to future hosts, splitting apart the data by timepoint revealed more subtle differences (Figures 5C-D & Table 2). Notably, phage from 1 and 9 dpi were most infective to contemporaneous hosts, and hosts from 9 dpi were most resistant to contemporaneous phage. Given that these patterns may be consistent with a fluctuating selection dynamic (FSD), we formally tested for the relative importance of arms race versus fluctuating selection in our experiment following Hall *et al*. (2011) by estimating the strength of the genotype X environment (GxE) effect on infectivity and resistance (see Methods & Materials). Stronger GxE effects are consistent with stronger fluctuating selection. This analysis showed that variation among genotypes, while variable, was also weak, consistent with a limited GxE effect (Supplementary Figure ?). The strength of the GxE effect did not change significantly with respect to timepoint for either phage infectivity (χ2(2,24)=1.93, *p*=0.38) or host resistance (χ2(2,24)=1.46, *p*=0.48). Further, host or phage background in our overall model of time-shift data explained 99.5% of the total variance (Conditional *R*2 = 0.377…, marginal *R*2 = 0.375…), suggesting that genotypes were almost identical in their response to time-shift challenges. Collectively, these data demonstrate that *S. thermophilus* DGCC7710 and phage D2972 coevolved under these experimental conditions, and that the dynamics of their coevolution predominantly follows an arms race.

Finally, we wanted to test the hypothesis that the patterns of infectivity evolution in these experiments are caused by the acquisition of point mutations in the target sequences. Such mutations have been observed in a previous co-culture experiment (Paez-Espino MBio), and provides a known mechanism for phage to overcome CRISPR resistance (Deveau JBac). We therefore PCR amplified … protospacer sequences, followed by Sanger sequencing. This showed that …% of infectivity patterns could be explained by point mutations (**Figure X**), suggesting this is the dominant if not exclusive mechanism of CRISPR-phage coevolution in this system.

**DISCUSSION**

*Streptococcus thermophilus* DGCC7710 is one of the best studied models for studying the evolution of CRISPR resistance. The two most active CRISPR loci, CRISPR1 and CRISPR3, are part of type II-A CRISPR-Cas systems and are induced upon infection with D2972 (Young PLoS One). If bacteria carry spacers that target the phage, the crRNA guide Cas9 enzymes to cleave the phage genome within the complementary sequences (protospacers) 3 nt upstream of the PAM (Garneau Nature; Magadan PLoS). If matching spacers are not yet present within the arrays, the bacteria can survive if they capture a new spacer, a process that is Cas9-dependent (Terns ; Marraffini). Such acquisition of novel spacers is a rare event (estimated frequency of 1 in … (see Varble et al)), and thought to occur predominantly in response to defective phages (Hynes & Moineau). Alternatively, if the bacteria encode both a Restriction-Modification system and a CRISPR system, the cleavage by the former may facilitate spacer acquisition by the latter (Hynes Nature Comm).

When the rates of spacer acquisition are low, the levels and evenness of spacer diversity are expected to be relatively low as well. Both theory and data have demonstrated that low spacer diversity levels favour phage persistence through their acquisition of point mutations to overcome host resistance. In the case of the Type I-F system of *Pseudomonas aeruginosa*, high levels of spacer diversity naturally evolve following infection with phage DMS3vir, leading to rapid phage extinction (van Houte). This work and a previous co-culture experiment show that *S. thermophilus* strain DGCC7710 and its lytic phage D2972 can coexist for approximately 35 days (Paez-Espino MBio), but due to lack of phenotypic data to support coevolution it remained unclear whether phage persistence was driven by coevolution or back mutation of host resistance to susceptibility (Weissman ISME). Here we present the first direct evidence of coevolution between CRISPR and phage under laboratory conditions, and demonstrate that it is associated with arms race dynamics (ARD). During ARD hosts increasing their resistance ranges over time through the accumulation of spacers and phages increasing their infectivity ranges over time through the accumulation of point mutations in protospacers. ARD may eventually result in phage extinction as the burden of acquiring increasing numbers of mutations in the phage genome becomes too great for the phage ([Lenski & Levin 1985](#_ENREF_15); [van Houte *et al.* 2016](#_ENREF_22)). As a consequence of the increasing costs of resistance and infectivity, ARD can in some cases give way to fluctuating selection dynamic (FSD) coevolution, which is typified by negative frequency dependent selection, resulting in selection against common bacterial and phage genotypes ([Gómez & Buckling 2011](#_ENREF_10); [Hall *et al.* 2011](#_ENREF_11)). According to our analyses, the contribution of FSD was small during the time period that was sampled here.

In nature… Banfield papers ; also discuss Nature Comm paper Sundberg salmon farm!

Why does phage coevolve with CRISPR system of *Streptococcus thermophilus* but not with that of *P. aeruginosa*? This is probably because the levels of spacer diversity that evolve in the former empirical system are much lower than those that evolve in P. aeruginosa (Westra 2015; van Houte 2016). This then raises the question why these systems evolve differently. One reason may be the difference in PAM utilization between the two systems. P. aeruginosa IF CRISPR-Cas system has a PAM of 2 nt (GG) ([Mojica *et al.* 2009](#_ENREF_17)), whereas the Type II-A system in *Streptococcus thermophilus* which has a PAM of 3-5 nt (NGGNG and NNAGAAW) ([Deveau *et al.* 2008](#_ENREF_9); [Horvath *et al.* 2008](#_ENREF_12)), which inevitably causes the latter CRISPR systems to be more restricted with regards to protospacer selection, and consequently have a reduced potential to generate diversity. Possibly an even more important effect is that of priming. The type IF system of P. aeruginosa is primed against phage DMS3vir, which relies on the presence of a partial match between a pre-existing spacer and the phage genome. This probably leads to low levels of cleavage activity against the phage, effectively generating defective phages by inactivating their infection cycle, and generating substrates for spacer selection (refs priming). By contrast, priming has not been described for *S. thermophilus* and the system therefore may rely on naturally inactivated phages, e.g. due to DNA damage, mutation, etc. The resulting low levels of CRISPR diversity allow phage to coevolve with the CRISPR system of the host, although sooner or later the phage was nevertheless driven extinct. Interestingly, previous experiments suggest that longer periods of bacteria-phage coexistence are reached when experimental treatments contained additional phages (Paez Espino 2x). Under these conditions, phage were found to escape not only by mutation, but also by recombination (Paez-Espino MBio), an observation consistent with observations from natural environments where phage recombinations were correlated with CRISPR activity (Anderson & Banfield).

While in primed systems, such as the P. aeruginosa Type IF CRISPR-Cas system, the presence of multiple phages may also contribute of bacteria-phage coexistence, another factor that is important in this context is the presence of Anti-CRISPRs in many Pseudomonas phages (Bondy-Denomy), which allow phages to amplify if their initial densities are sufficiently high (Landsberger Cell). Such Acr’s have recently also been identified in S. thermophilus phages, suggesting they are an important factor in shaping the population dynamics of bacteria with CRISPR systems and their phages. Acr’s when CRISPR can generate lots of spacer diversity.

**Figure Legends**

**Figure 1: Phage and host population dynamics over time in each replicate.**

Phage titres (plaque-forming units; pfu ml-1) are shown in black and host densities (colony-forming units; cfu ml-1) are shown in blue. Cfu ml-1 were derived by multiplying OD600values by 1x108. Values of 100 represent undetectable levels of host or phage (i.e. extinction), and not 1. **A-C)** 109-106 phage treatments, respectively, with replicate identity shown above each sub-panel.

**Figure 2: Evolution of infectivity and resistance over time.**

**A)** Infectivity is the proportion of host genotypes at all timepoints per replicate that were infected by a sympatric phage genotype. **B)** Resistance is the proportion of phage genotypes at all timepoints per replicate that were resisted by a sympatric host genotype. Means and 95% CIs are shown (*N*=8).

**Figure 3: Spacers acquired during coexistence of S. thermophilus and phage D2972.**

**A)** Number of acquired spacers per clone at each day post-infection (dpi). Means and 95% CIs are shown (*N*=8). **B)** Mean relative frequency of clones with different numbers of acquired spacers at each dpi. No clone with >3 spacers was detected. Means and 95% CIs are shown (*N=*8). **C)** Raw pairwise differences among acquired spacers in each replicate (x-axis) at each timepoint (colours)

**Figure 4: Protospacer locations of newly acquired spacers.**

Histogram showing the location of acquired spacers in each replicate when mapped against the D2972 genome: **A)** 1 dpi, replicate 1; **B-E)** 4 dpi, replicates 3, 4, 6 and 7; **F-M)** 9 dpi, all replicates. Each dot represents a clone that had a spacer mapped to that region. Spacers in the CRISPR1 locus are shown in red, CRISPR3 in blue. Darker colours are the result of visual overlap between dots.

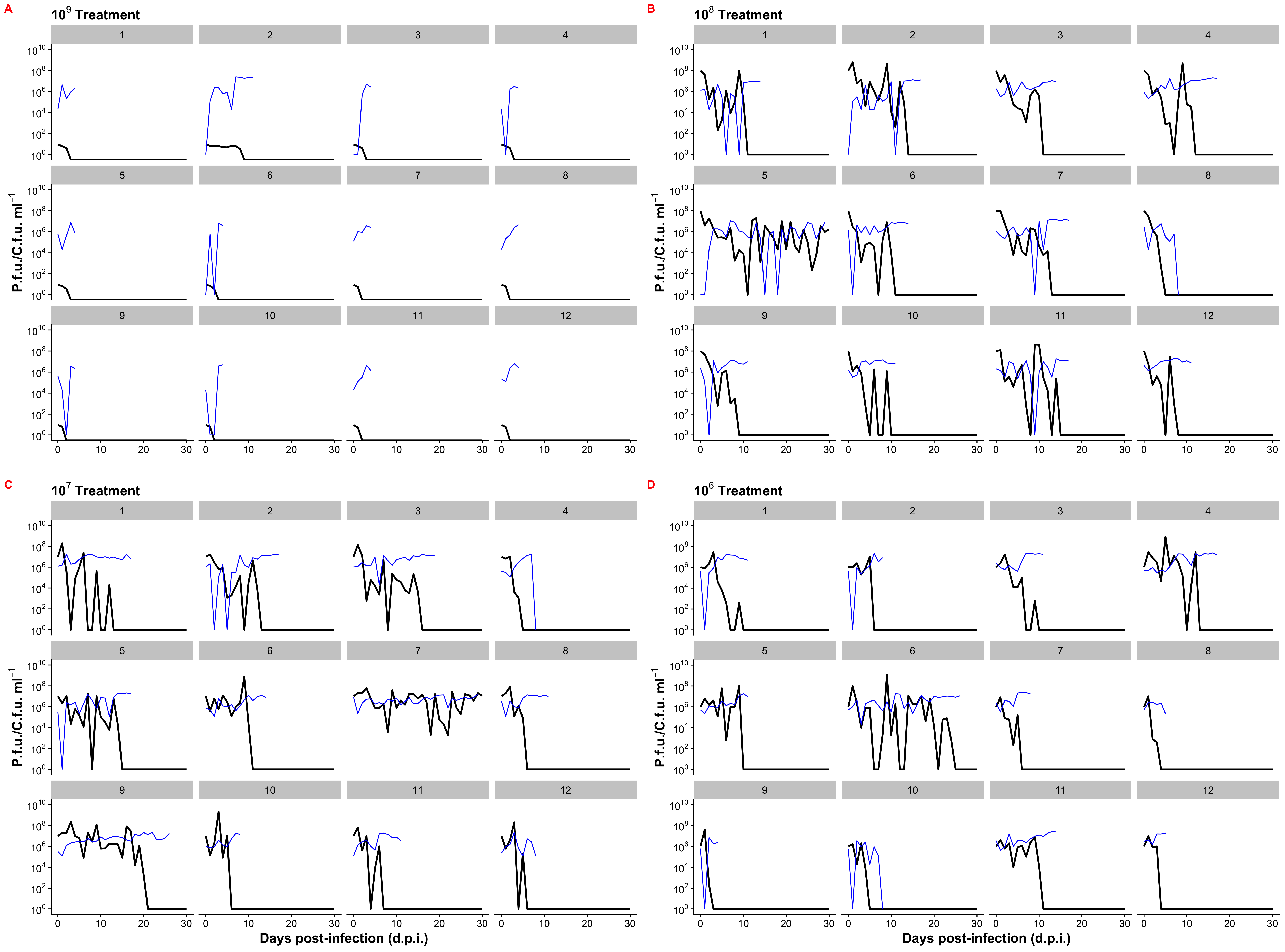
**Figure 5: The coevolutionary dynamics of CRISPR-phage interactions.**

**A)** Proportion of hosts infected when hosts were from the phage’s past, present or future. **B)** Proportion of hosts resistant to phage that were from the host’s past, present or future. The dotted line in A and B is shown for illustrative purposes and is not predictive in any way. **C)** Mean pairwise infectivity of phage to hosts in the timeshift assay. **D)** Mean pairwise resistance of hosts to phage to in the timeshift assay. Means and 95% CIs are shown (*N*=100048).

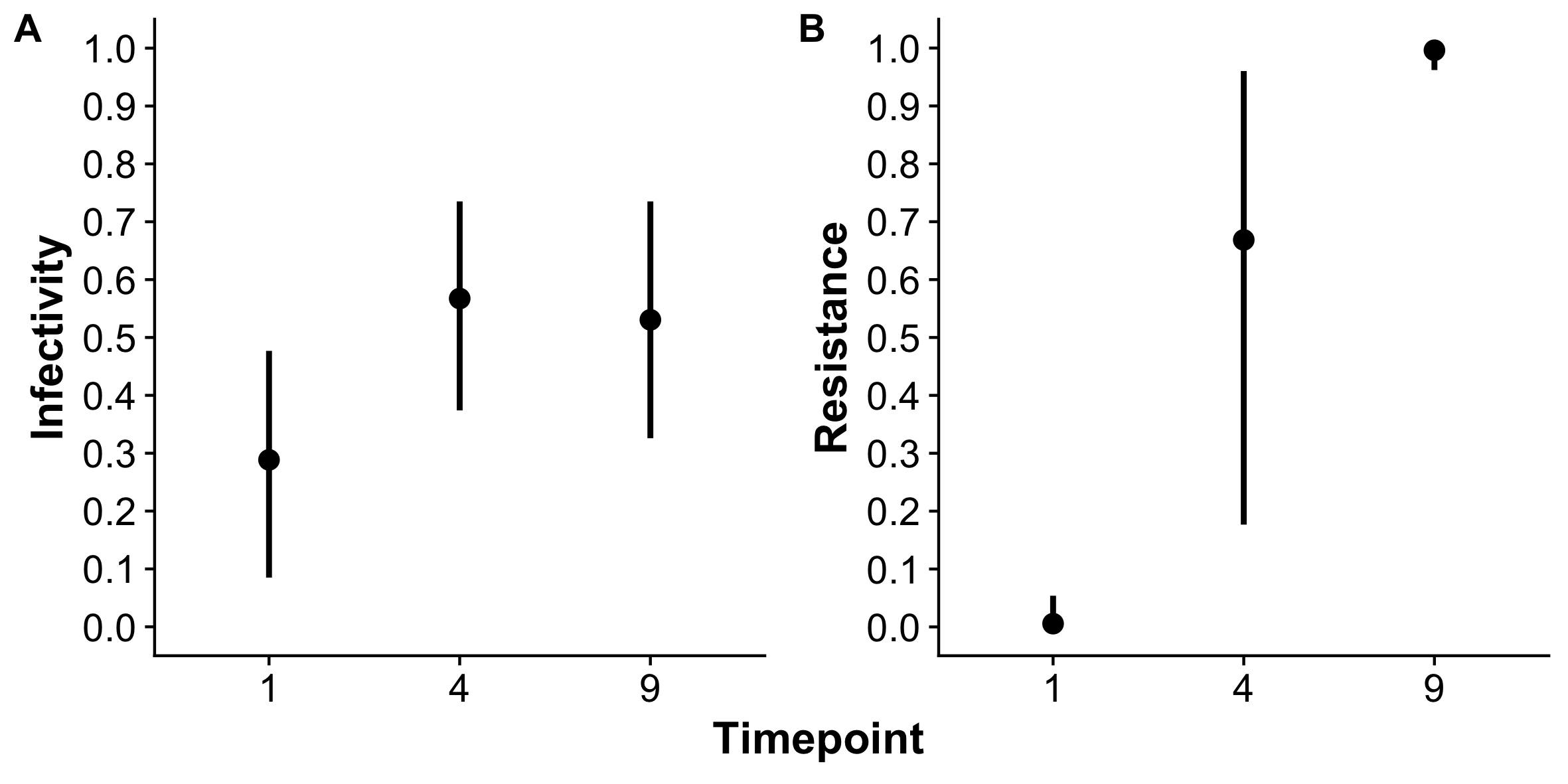
**Figure 6: The mechanistic basis of phage infectivity evolution.**

**Figures**

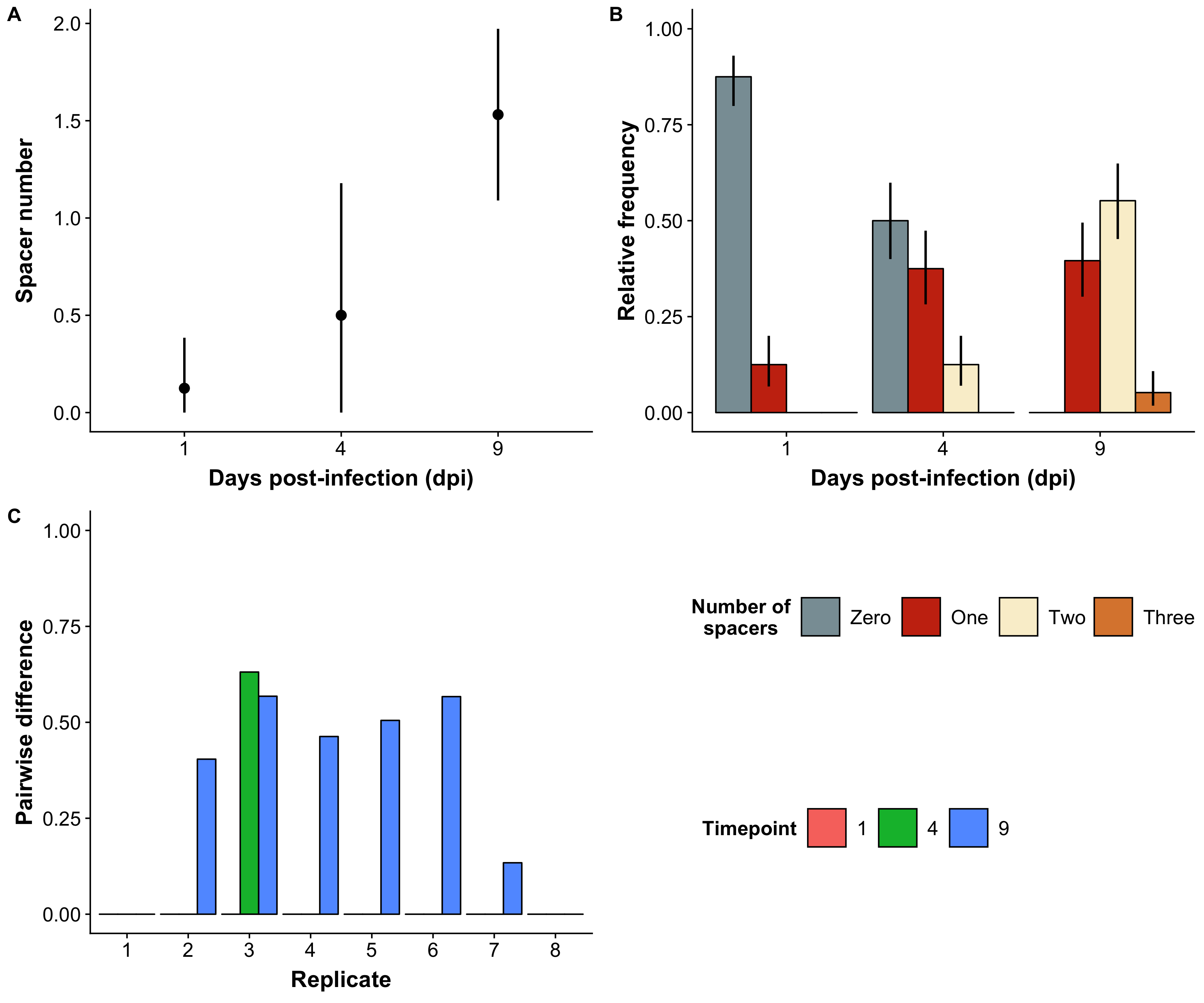
**Figure 1**

****

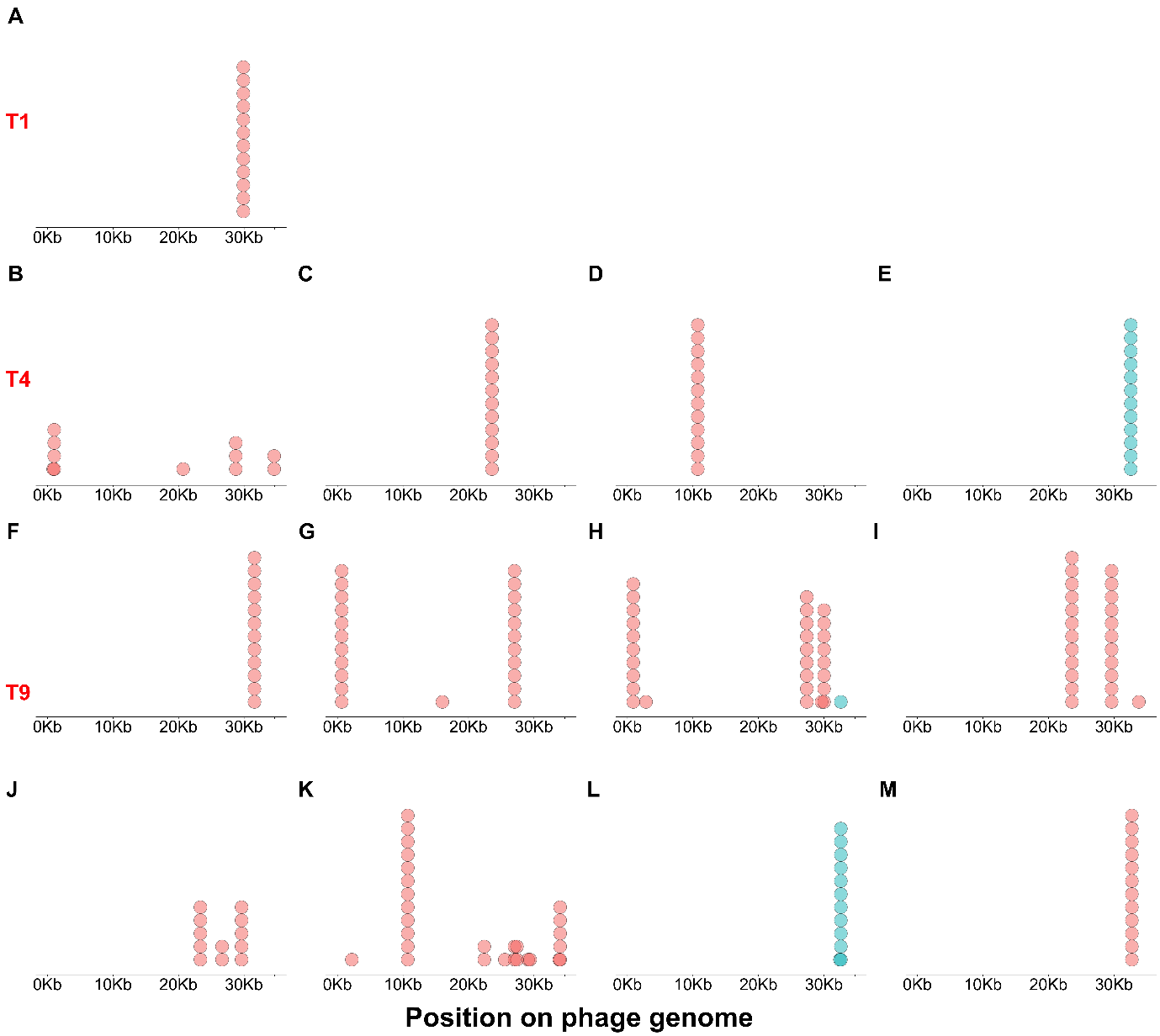
**Figure 2**



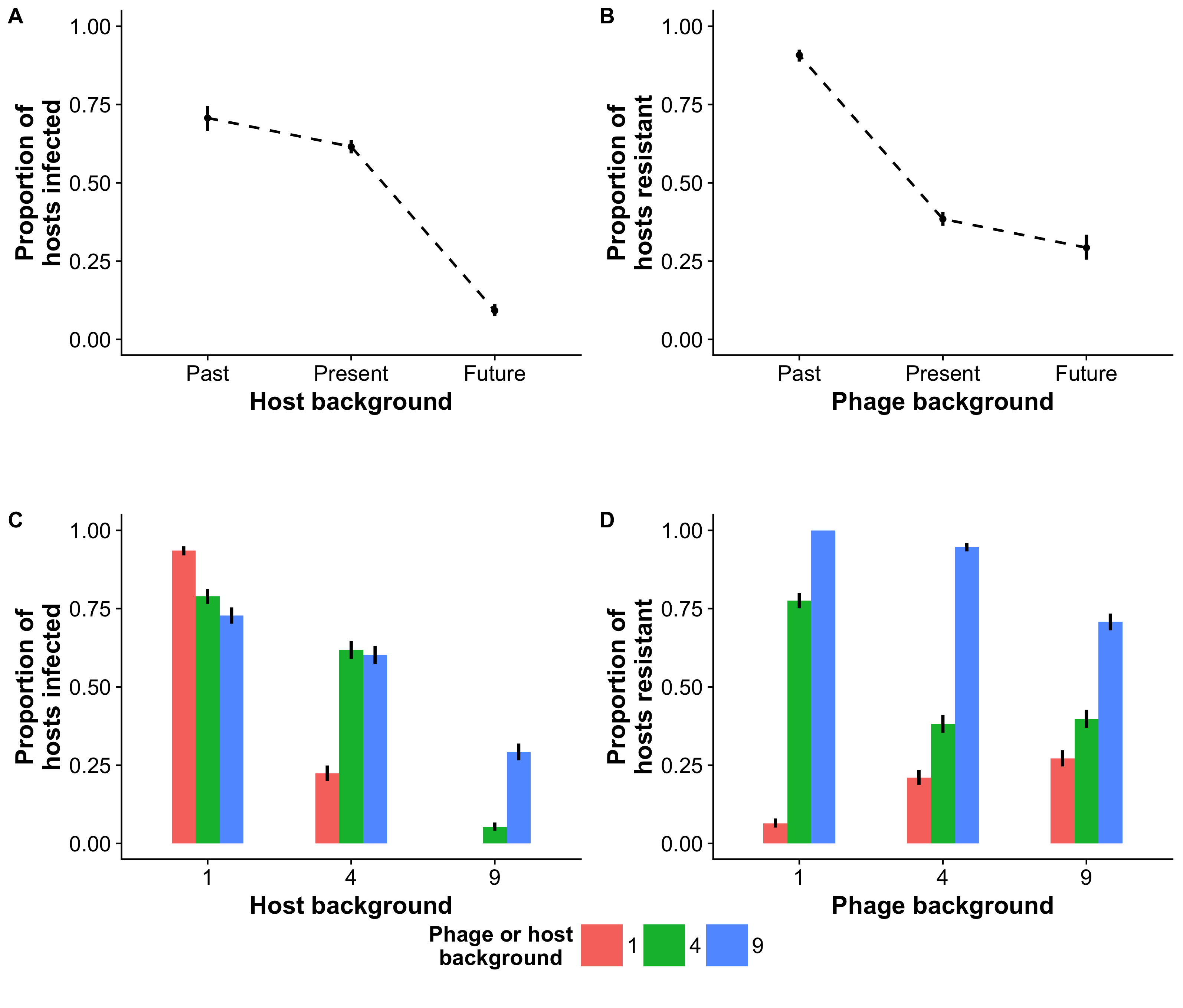
**Figure 3**

****

**Figure 4**

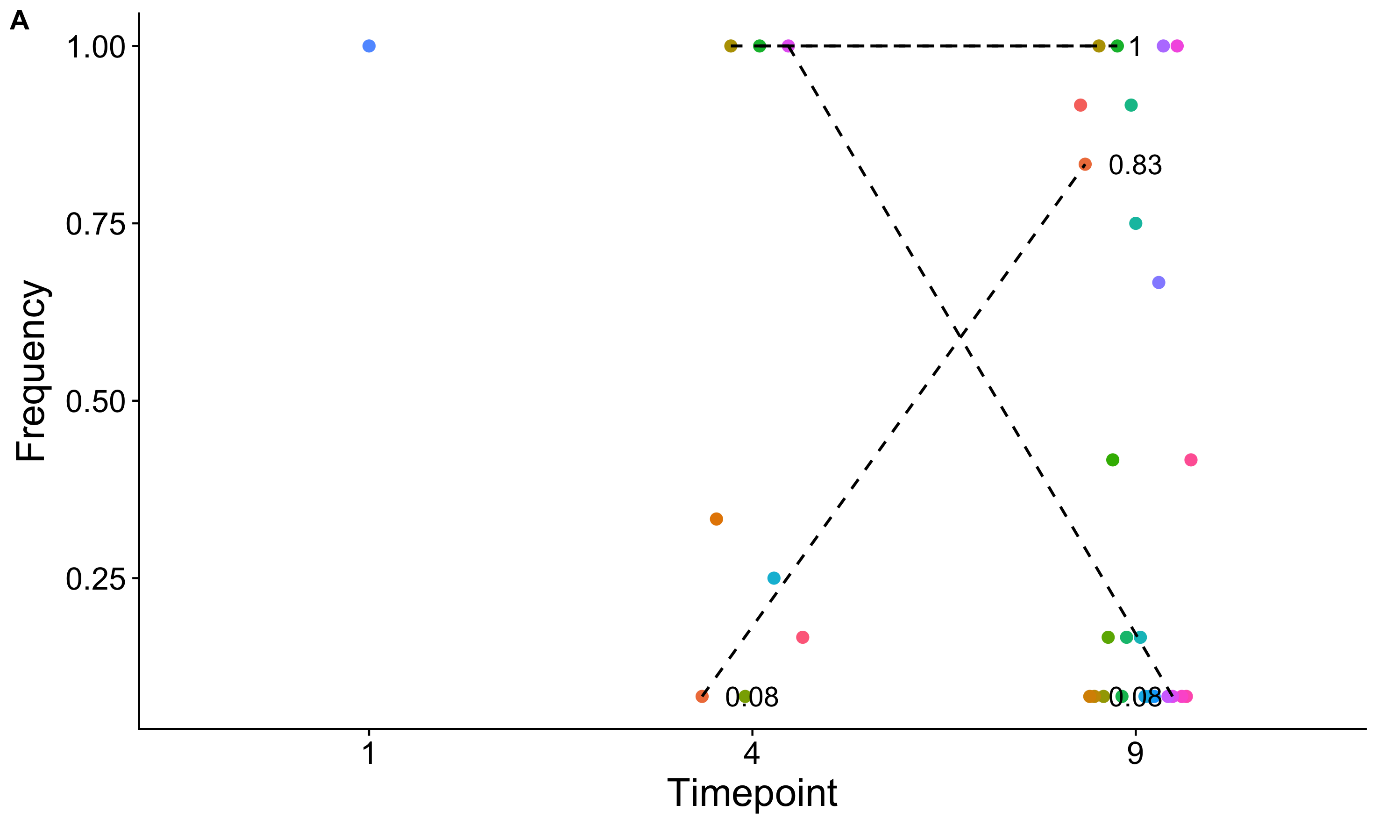
****

**Figure 5**

****

**Figure 6**

**Figure 7**

****

**Fig 7: Dynamics of spacers and point mutations gained and lost.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | Phage | | |
| **1** | **4** | **9** |
| Host | **1** | Present | Past | Past |
| **4** | Future | Present | Past |
| **9** | Future | Future | Present |

**Table 1:** Pairwise challenges between phage and hosts in the time shift assay. Numbers indicate the time points (days post-infection) analysed. Past, present or future refer to if hosts were contemporaneous or not with respect to the phage.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Host origin (day)** | **Phage origin (day)** | **Mean Infectivity** | **Infectivity 95% CI** | **Mean Resistance** | **Resistance 95% CI** |
| 1 | 1 | 0.93 | 0.92---0.95 | 0.06 | 0.05---0.08 |
| 4 | 4 | 0.22 | 0.20---0.25 | 0.78 | 0.75---0.79 |
| 9 | 9 | 0 | 0.00—0.00 | 1.00 | 1.00---1.00 |
| 1 | 1 | 0.79 | 0.76---0.81 | 0.21 | 0.19---0.24 |
| 4 | 4 | 0.62 | 0.59---0.65 | 0.38 | 0.35---0.41 |
| 9 | 9 | 0.05 | 0.04---0.07 | 0.95 | 0.93---0.96 |
| 1 | 1 | 0.73 | 0.70---0.75 | 0.27 | 0.25---0.29 |
| 4 | 4 | 0.6 | 0.57---0.63 | 0.39 | 0.37---0.43 |
| 9 | 9 | 0.29 | 0.27---0.32 | 0.71 | 0.68---0.73 |

**Table 2**: Mean proportion and 95% confidence interval (CI) of hosts infected and phage resisted in pairwise challenges in the time-shift assay, broken down by the day from which the host or phage originated. Values are rounded to two decimal places.

**REFERENCES** (Change to Vancouver style for final copy)

Achigar R, Magadán AH, Tremblay DM, Pianzzola MJ, Moineau S (2017) Phage-host interactions in Streptococcus thermophilus: Genome analysis of phages isolated in Uruguay and ectopic spacer acquisition in CRISPR array. *Scientific reports* **7**, 43438.

Agrawal A, Lively CM (2002) Infection genetics: gene-for-gene versus matching-alleles models and all points in between. *Evolutionary Ecology Research* **4**, 91-107.

Akaike H (1973) Information theory and an extension of the maximum likelihood principle,[w:] proceedings of the 2nd international symposium on information, bn petrow, f. *Czaki, Akademiai Kiado, Budapest*.

Bates D, Mächler M, Bolker B, Walker S (2014) Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Bhaya D, Davison M, Barrangou R (2011) CRISPR-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Regulation. *Annual Review of Genetics* **45**, 273-297.

Burnham KP, Anderson DR (2003) *Model selection and multimodel inference: a practical information-theoretic approach* Springer Science & Business Media.

Burnham KP, Anderson DR (2004) Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research* **33**, 261-304.

Childs LM, England WE, Young MJ, Weitz JS, Whitaker RJ (2014) CRISPR-induced distributed immunity in microbial populations. *PLoS One* **9**, e101710.

Deveau H, Barrangou R, Garneau JE*, et al.* (2008) Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *Journal of bacteriology* **190**, 1390-1400.

Gómez P, Buckling A (2011) Bacteria-phage antagonistic coevolution in soil. *Science* **332**, 106-109.

Hall AR, Scanlan PD, Morgan AD, Buckling A (2011) Host–parasite coevolutionary arms races give way to fluctuating selection. *Ecology Letters* **14**, 635-642.

Horvath P, Romero DA, Coûté-Monvoisin A-C*, et al.* (2008) Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. *Journal of bacteriology* **190**, 1401-1412.

Iranzo J, Lobkovsky AE, Wolf YI, Koonin EV (2013) Evolutionary dynamics of archaeal and bacterial adaptive immunity systems, CRISPR-Cas, in an explicit ecological context. *Journal of bacteriology*, JB. 00412-00413.

Kearse M, Moir R, Wilson A*, et al.* (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647-1649.

Lenski RE, Levin BR (1985) Constraints on the coevolution of bacteria and virulent phage: a model, some experiments, and predictions for natural communities. *The American Naturalist* **125**, 585-602.

Levin BR, Moineau S, Bushman M, Barrangou R (2013) The population and evolutionary dynamics of phage and bacteria with CRISPR–mediated immunity. *PLoS genetics* **9**, e1003312.

Mojica FJ, Díez-Villaseñor C, García-Martínez J, Almendros C (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**, 733-740.

Paez-Espino D, Sharon I, Morovic W*, et al.* (2015) CRISPR immunity drives rapid phage genome evolution in Streptococcus thermophilus. *MBio* **6**, e00262-00215.

Semenova E, Jore MM, Datsenko KA*, et al.* (2011) Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proceedings of the National Academy of Sciences* **108**, 10098-10103.

Team RC (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013.

Therneau TM, Lumley T (2015) Package ‘survival’. *R Top Doc* **128**.

van Houte S, Ekroth AK, Broniewski JM*, et al.* (2016) The diversity-generating benefits of a prokaryotic adaptive immune system. *Nature* **532**, 385.

Weissman JL, Holmes R, Barrangou R*, et al.* (2018) Immune loss as a driver of coexistence during host-phage coevolution. *The ISME journal* **12**, 585.

Wickham H (2009) ggplot2: Elegant Graphics for Data Analysis Springer-Verlag New York. Version.