The diversity-generating benefits of a prokaryotic adaptive immune system

Stineke van Houte[1], Alice K.E. Ekroth[1], Jenny M. Broniewski[1], Hélène Chabas[1,2], Ben Ashby[3], Joseph Bondy-Denomy[4], Sylvain Gandon[2], Mike Boots[3,5], Steve Paterson[6], Angus J. Buckling[1] and Edze R. Westra[1]

[1] Environment and Sustainability Institute, Centre for Ecology and Conservation, Biosciences, Cornwall Campus, Penryn, UK.

[2] CEFE UMR 5175, CNRS-Université de Montpellier, Université Paul-Valéry Montpellier, EPHE, 1919, route de Mende 34293 Montpellier Cedex 5, France

[3] Centre for Ecology and Conservation, Biosciences, Cornwall Campus, Penryn, UK.

[4] Department of Microbiology and Immunology, University of California, San Fransisco, 94158, San Francisco, US.

[5] Department of Integrative Biology, University of California, Berkeley, US.

[6] Institute of Integrative Biology, University of Liverpool, UK.

Prokaryotic CRISPR-Cas adaptive immune systems insert spacers derived from viruses and other parasitic DNA elements into CRISPR loci to provide sequence-specific immunity[1,2]. This frequently results in high within-population spacer diversity[3-6], but it is unclear if and why this is important. Here, we show that as a result of this spacer diversity, viruses can no longer evolve to overcome CRISPR-Cas by point mutation, which results in rapid virus extinction. This effect arises from synergy between spacer diversity and the high specificity of infection, which greatly increases overall population resistance. We propose that the resulting short-lived nature of CRISPR-dependent bacteria-virus coevolution has provided strong selection for the evolution of sophisticated virus-encoded anti-CRISPR mechanisms[7].

We previously reported that *Pseudomonas aeruginosa* strain UCBPP-PA14 evolves high levels of CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats; CRISPR-associated) adaptive immunity against virus DMS3vir under laboratory conditions[6]. However, viruses can readily evolve to overcome sequence specific CRISPR immunity[8,9]. To study how CRISPR-Cas impacts virus persistence, we measured titers of virus DMS3vir over time upon infection of either wild type (WT) *P. aeruginosa* or a functional CRISPR-Cas knock-out (CRISPR KO) strain. Virus that infected the WT strain went extinct at 5 days post-infection (dpi) (Fig. 1A), whereas virus infecting the CRISPR KO strain persisted in all replicates until the experiment was terminated at 30 dpi (Fig. 1B). WT bacteria exclusively evolved CRISPR-mediated immunity, while the CRISPR KO strain evolved immunity by mutation, loss or masking of the receptor (i.e. surface mutation) (Extended Data Fig.

1). The observation that CRISPR-Cas drives virus extinct so rapidly was unexpected since viruses can escape CRISPR immunity by a single point mutation[8,9].

Virus extinction might result from the high level of spacer diversity that naturally evolves upon virus exposure in this and other CRISPR-Cas systems[3-6]. Both theory and data suggest that host genetic diversity can synergistically reduce the spread of parasites if the infection process is specific (i.e. a parasite genotype can infect a restricted number of host genotypes) and a failed infection results in parasite death[10-18]; assumptions that hold for CRISPR-Cas-virus interactions. While the protective effect of host diversity may be lost following the evolution of single viruses that escape from multiple spacers[10,17], host diversity has the additional benefit of limiting such viral adaptation. Specifically, lower virus population sizes resulting from host diversity[11,12] reduces the probability of escape mutations, and the greater the diversity the more escape mutations needed.

To examine these hypotheses, we generated bacterial populations in which we manipulated the level of spacer diversity; we used 48 individual clones with CRISPR-based immunity against virus DMS3vir to generate bacterial populations with five distinct diversity levels: monocultures or polycultures consisting of equal mixtures of either 6, 12, 24 or 48 clones. To allow for direct comparisons, each of the 48 clones was equally represented at each diversity level by adjusting the number of replicate experiments accordingly. Each population was competed against a previously described surface mutant[6] in the presence or absence of virus DMS3vir and virus levels were monitored over time.

This experiment revealed a strong inverse relationship between virus persistence and the level of spacer diversity in the bacterial population (Fig. 2). Virus titers remained high in 44 out of 48 replicates when the CRISPR population consisted of a monoculture (Fig. 2A). However, as diversity increased, virus persistence decreased (Fig. 2B-E) and virus was driven extinct rapidly and reproducibly when the CRISPR population consisted of a 48-clone mixture (Fig. 2E).

Next, we examined the fitness consequences of generating spacer diversity. In the absence of virus there was no significant effect of diversity on the relative fitness associated with CRISPR-Cas compared to a resistant surface mutant (Extended Data Fig. 2; $F_{1, 52}$=3.20, p=0.08). However, in the presence of virus CRISPR-associated fitness increased with increasing spacer diversity (Fig. 3; $F_{4,71}$=40.30 p<0.0001 and Extended Data Table 1), with mean fitness increasing 11-fold from monoculture to the highest diversity population. In monoculture, the CRISPR population was outcompeted by the surface mutant (rel. fitness < 1 ; T=-11.68, p<0.0001). However, as diversity increased, the CRISPR population consistently outcompeted the surface mutant (rel. fitness > 1; 6-clones: T=3.05, p=0.0093; 12-clones: T=3.95, p=0.0028; 24-clones: T=3.48, p=0.0088; 48-clones: T=3.06, p=0.014; all significant after sequential Bonferroni correction), showing that the generation of spacer diversity is an important fitness determinant of CRISPR-Cas (Fig. 3).

Given that all bacterial clones used in the experiment were initially resistant, we hypothesized that the benefit of spacer diversity emerges from an inability of virus to evolve escape mutants. To examine this, virus isolated from each time point (0, 16, 24, 40, 48, 64 and 72 hours post-infection) was spotted onto lawns of each of the 48 CRISPR clones. As expected, we could not detect escape virus in the ancestral virus (Fig. 4A; left column, indicated in green). However, in 43 of the 48 CRISPR monocultures, virus evolved within 2 days to overcome CRISPR immunity (Fig. 4A; indicated in red). For 5 clones no escape virus could be detected, and virus went extinct in 4 of these instances (Fig. 4A, asterisks). Three of these 5 clones carried multiple spacers targeting the virus, which limits the emergence of escape virus[16]. The emergence of escape virus decreased as diversity increased to 6, 12, 24 and 48 CRISPR alleles (Fig. 4); in the latter, no escape virus could be detected. These phenotypic data were supported by results of deep sequencing of virus genotypes isolated from 1 dpi: there was a significant inverse relationship between host diversity and the accumulation of viral mutations in the target sequences (Extended Data Fig. 3). This is because virus needs to overcome multiple spacers in the diverse host population if it is to increase in frequency (Extended Data Fig. 4). Consistent with a lack of escape virus emerging against all host genotypes, the spacer content of mixed populations of 6, 12, 24 and 48 clones did not increase between $t=0$ and $t=3$ (Wilcoxon Signed Rank $p>0.2$ for all treatments), whereas monocultures acquired novel spacers in response to emerging escape virus (Wilcoxon Signed Rank $W=333$, $DF=47$, $p<0.0001$; Extended Data Fig. 5). These data show that while escape viruses can clearly evolve against most of the clones, escape viruses do not emerge when these clones are mixed.

We hypothesized that the benefit of within-population spacer diversity is because of synergy between the different clones. However, diversity will also increase the chance that a single clone with one or more spacers that the virus is unable to overcome will be present in the population. Indeed, we observed 5 clones against which escape mutants were never detected, and presence of these clones in many of the diverse populations could explain the fitness advantage of diversity. To investigate if synergy plays an important role in the benefit of diversity beyond this “jackpot” effect, we compared the fitness of diverse populations with the fitness of the fittest constituent clone, as measured in monoculture. This analysis revealed that synergism contributed an approximately 50% growth rate advantage when in competition with surface mutants (Mean ± SEM difference in fitness between mixtures and fittest constituent monoculture = 0.47 ± 0.18; $P < 0.01$)

The short-lived nature of coevolution between CRISPR-resistant bacteria and virus escape mutants beyond a host diversity threshold may explain the evolution of sophisticated anti-CRISPR mechanisms to overcome CRISPR-Cas[7]. Indeed, a virus carrying an anti-CRISPR gene[7] was found to persist independent of CRISPR diversity levels (Extended Data Fig. 6AB) and caused similar extinction of CRISPR-resistant monocultures and 48-clone populations that competed against a surface mutant (Fisher's exact test, p=1.0 at t=1, p=0.33 at t=3 dpi; Extended Data Fig. 6C).

Finally, to test that our results were not limited to the *P. aeruginosa* PA14 Type I-F CRISPR-Cas system, we performed a similar experiment with *Streptococcus thermophilus* DGCC7710 clones that evolved resistance against virus 2972 using a Type II-A CRISPR-Cas system. As shown in Extended Data Fig. 7, we found a similar effect of CRISPR resistance allele diversity on virus persistence and escape virus emergence. However, during coevolution experiments the levels of evolved diversity are lower in *S. thermophilus* (data not shown and refs. 4,5), which, consistent with theory[10,17], allows for more persistent coevolution[4,5]. Lower levels of evolved spacer diversity might be due to a more weakly primed CRISPR-Cas system[19-21].

Collectively, our data demonstrate that the propensity to generate host genetic diversity is a key fitness determinant of CRISPR-Cas adaptive immune systems because it limits the emergence of escape virus. Consistent with the idea that it is harder for a parasite to adapt to a heterogeneous host population[22], virus rapidly evolved high levels of infectivity on monocultures, but not on a diverse mix of the same host genotypes. Parasites are often invoked as the selective force driving the evolution of diversity generating mechanisms[22-26]. In most cases, individual-level selection is assumed to be the driver of these traits, because individual benefits are high, and group selective benefits would be opposed by the invasion of individuals who do not pay the fitness costs associated with these mechanisms (e.g. sex and increased mutation rates)[26-28]. In the case of CRISPR-Cas, we speculate that population-level selection may have contributed to its evolution. First, there were large benefits associated with synergy between diverse genotypes. Second, costs of CRISPR-Cas are conditional on virus exposure[6,29] and clones lacking CRISPR immunity cannot invade populations (Extended Data Figs. 8-11). Third, the highly structured nature of bacterial populations, and the resulting high relatedness, promotes between-population selection[30]. Future tests of this hypothesis are needed to reconcile the selective forces that have shaped the evolution of CRISPR-Cas systems.

Methods

Bacterial strains and viruses

P. aeruginosa UCBPP-PA14 (WT), *P. aeruginosa* UCBPP-PA14 *csy3::LacZ* (referred to as CRISPR KO, which carries a disruption of an essential *cas* gene and can therefore not evolve CRISPR immunity), the CRISPR KO-derived surface mutant and virus DMS3vir have all been described in ref. 6 and references therein. Phage DMS3vir+*acrF1*, which carries the anti-CRISPR gene *acrF1* (formerly 30-35), was made by inserting *acrF1* into the DMS3vir genome using methods described in ref. 7. *Streptococcus thermophilus* strain DGCC7710 and its virus 2972 have been described in ref. 2.

Coevolution experiments

The coevolution experiments shown in Fig. 1 were performed in glass microcosms by inoculating 6 ml M9 supplemented with 0.2% glucose with approximately 10^6 colony forming units (cfu) bacteria from fresh overnight cultures of the WT *P. aeruginosa* UCBPP-PA14 or CRISPR KO strain and adding 10^4 plaque forming units (pfu) of virus DMS3vir, followed by incubation at 37 ºC while shaking at 180 rpm (6 replicates). Cultures were transferred daily 1:100 to fresh broth. Virus titers were determined at 0, 3, 5, 11, 17, 22 and 30 days after the start of the coevolution experiment by spotting virus samples isolated by chloroform extraction on a lawn of CRISPR KO bacteria. The analysis of virus immunity was performed by cross-streak assay and PCR as described previously[6].

Generation of populations with different levels of CRISPR diversity

For the competition experiments, shown in Figs. 2-4 and Extended Data Figs. 2-6 and

8-11, we generated *P. aeruginosa* populations with varying levels of CRISPR spacer (allele) diversity. To this end, we isolated from the 6 replicates of the coevolution experiment (Fig. 1) a total of 48 individual clones that had acquired CRISPR immunity against virus DMS3vir. We have previously shown that individual clones tend to have unique spacers[6]. Using these 48 clones, populations with five different levels of CRISPR spacer (allele) diversity were generated. These populations consisted of: *1)* 1 clone (a monoculture; a clonal population carrying a single spacer); equal mixtures of *2)* 6 clones; *3)* 12 clones; *4)* 24 clones and *5)* 48 clones. In total 48 different monocultures (48 x monocultures), 8 x 6-clone populations, 4 x 12-clone populations, 2 x 24-clone populations and 1 x 48-clone population were generated (details of the composition of each population can be found below, under "number of replicate experiments").

Competition experiments

Competition experiments were done in glass microcosms in a total volume of 6 ml M9 supplemented with 0.2% glucose. Competition experiments were initiated by inoculating 1:100 from a 1:1 mixture (in M9 salts) of overnight cultures of the appropriate CRISPR population and either the surface mutant (Figs. 2-4 and Extended Data Figs. 2, 4-6, 8) or the CRISPR KO strain (Extended Data Figs. 7-11). At the start of each experiment 10^9 pfu of virus was added, unless indicated otherwise. Cultures were transferred daily 1:100 into fresh broth. At 0 and 72 hours post-infection (hpi) samples were taken and cells were serially diluted in M9 salts and plated on LB agar supplemented with 50 $\mu g \cdot ml^{-1}$ X-gal (to allow discrimination between WT-derived CRISPR clones (white) and CRISPR KO or surface mutant (blue)). The relative frequencies of the WT strain were used to calculate the relative fitness (rel. fitness =

[(fraction strain A at t=x) * (1 – (fraction strain A at t=0))] / [(fraction strain A at t=0) * (1 – (fraction strain A at t=x)]). At 0, 16, 24, 40, 48, 66 and 72 hpi, samples were taken and chloroform extractions were performed to isolate total virus, which was spotted on a lawn of CRISPR KO bacteria for quantification. All subsequent statistical analyses were carried out using JMP (v12) software.

Determination of escape virus emergence

To determine the emergence of escape virus during the competition experiments, every isolated virus sample was spotted onto 48 different bacterial lawns, corresponding to each of the different CRISPR clones. This procedure was done for each of the seven time points (see above), to enable us to track the emergence of escape virus against every individual clone over the time course of the experiment.

Deep sequencing

Isolated phage samples from t=1 dpi of the competition experiment shown in Fig. 2-4 were used to perform deep sequencing of spacer target sites on the phage genomes. To obtain sufficient material, phage were amplified by plaque assay on the CRISPR KO strain. Viruses from all replicates within a single diversity treatment were pooled. As a control, ancestral virus and escape virus from competition between *sm* and monocultures of CRISPR clones 1-3 were processed in parallel. Virus genomic DNA extraction was performed from 5 ml sample at approximately 10^{10} pfu/ml using the Norgen phage DNA isolation kit, following the manufacturer’s instructions. Barcoded Illumina Truseq Nano libraries were constructed from each DNA sample with an approximately 350bp insert size and 2x 250bp reads generated on an Illumina MiSeq platform. Reads were trimmed using Cutadapt v1.2.1 and Sickle v1.200 and then overlapping reads merged using Flash v1.2.8 to create high quality sequence at approximately 8000x coverage of DMS3vir per sample. These reads were mapped to PA14 and DMS3vir genomes using bwa mem v0.7.12 and allele frequencies of SNPs within viral target regions quantified using samtools mpileup v0.1.18. Further statistical analyses was performed in R v3.2.2. Sequence data are available from the European Nucleotide Archive under accession PRJEB12001 and analysis scripts are available from https://github.com/scottishwormboy/vanHoute.

Determining the acquisition of new spacers

To examine spacer acquisition during the competition experiments shown in Fig. 2-4, we examined by PCR for each diversity treatment the spacer content of 384 randomly isolated clones at both t=0 and t=3 (192 clones per time point). For each replicate experiment, the difference in the total number of spacers between t=0 and t=3 was divided by the number of clones that were examined to calculate the average change in the number of spacers per clone.

Number of replicate experiments

To ensure equal representation of each of the 48 clones across the different treatments, the number of replicate experiments for a given diversity treatment was adjusted accordingly, with a total number of replicates of at least 6 for sufficient statistical power. Hence, competition experiments with the 1-clone (monoculture) populations were performed in 48 independent replicates, each corresponding to a unique monoculture of a CRISPR clone (clones 1-48; each clone is equally represented). Competition experiments with the 6-clone populations were performed in eight independent replicates, each corresponding to a unique polyculture population

(population 1: equal mixture of clones 1-6; population 2: clones 7-12; population 3: clones 13-18; population 4: clones 19-24; population 5: clones 25-30; population 6: clones 31-36; population 7: clones 37-42; population 8: clones 43-48). Competition experiments with the 12-clone populations were also performed in eight replicates, corresponding to 4 unique polyculture populations (replicate 1 and 2: clones 1-12; replicate 3 and 4: clones 13-24; replicate 5 and 6: clones 25-36; replicate 7 and 8: clones 37-48). Competition experiments with the 24-clone populations were performed in six replicates, corresponding to 2 unique polyculture populations (replicate 1-3: clones 1-24; replicate 4-6: clones 25-48). Competition experiments with the 48-clone populations were performed in six replicates, each corresponding to the same polyculture population (replicate 1-6: clones 1-48).

Escape phage degradation and fitness

In the experiment shown in Extended Data Fig. 3, approximately 10^8 pfus of either ancestral virus or escape virus, which was isolated from the competitions between monocultures 1-6 and the surface mutant, was used to infect a monoculture of the corresponding CRISPR clone or the 48-clone polyculture. Phage samples were taken at 0, 9, 20 and 28 hpi by chloroform extraction and titrated on a lawn of the CRISPR KO strain. Fitness of each of the escape phages was determined by a competition experiment between ancestral and escape virus; a 50:50 ratio of escape and ancestral phage (10^9 pfus total) was used to infect either a monoculture of the corresponding CRISPR clone or the 48-clone polyculture. Virus samples were taken at t=0 and t=20 hpi by chloroform extraction and used in a plaque assay on CRISPR KO. Next, individual plaques (48 plaques per replicate) were isolated and amplified on the CRISPR KO strain. To determine the ratio of escape and ancestral virus, virus from each individual plaque was spotted on a lawn of *1*) CRISPR KO (both ancestral and escape virus form plaques) and *2*) the corresponding CRISPR immune clone (only escape virus can form a plaque).

Effect of spacer diversity in *Streptococcus thermophilus*

Streptococcus thermophilus DGCC7710 was grown in M17 medium supplemented with 0.5% α-lactose (LM17) at 42°C. Virus 2972 was used throughout the experiments. Virus infections were carried out using 10^6 pfus of phage 2972 and 10mM $CaCl_2$ to facilitate the infection process. To obtain virus-resistant *S. thermophilus* clones, a sample of virus lysate at 24 hpi was plated on LM17 agar plates. Individual colonies were picked and PCR-screened for the acquisition of novel spacers in each of the 4 CRISPR loci, as described in ref. 2. A total of 44 individual clones were selected to generate 44 monocultures and a single polyculture comprised of a mix of 44 clones. These cultures were infected with 10^7 pfu of virus, and samples were taken after the indicated periods of time to isolate virus. We determined virus titers by spotting viral dilutions on lawns of ancestral bacteria, and the emergence of escape virus by spotting virus on lawns corresponding to each of the 44 CRISPR resistant clones.

References

1. van der Oost, J., Westra, E. R., Jackson, R. N., Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nature Rev. Microbiol.* **12**, 479-92 (2014).
2. Barrangou, R., *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-12 (2007).
3. Andersson, A.F., Banfield, J. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320**, 1047-50 (2008).
4. Paez-Espino, D., *et al*. Strong bias in bacterial CRISPR elements that confer immunity to phage. *Nature Comm.* **4**, 413 (2013).
5. Paez-Espino, D. *et al*. CRISPR immunity drives rapid phage genome evolution in *Streptococcus thermophilus*. *MBio* **6**, e00262-15 (2015).
6. Westra *et al*. Parasite exposure drives selective evolution of constitutive versus inducible defense. *Curr. Biol.* **25**, 1043-9 (2015).
7. Bondy-Denomy, J., Pawluk, A., Maxwell, K. L., Davidson, A. R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429-32 (2013).
8. Deveau, H., *et al*. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* **190**, 1390-400 (2008).
9. Semenova, E. *et al.* Interference by clustered regularly interspaced short palindromic repeats (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci USA.* **108**, 10098-103 (2011).
10. Childs, L. M., England, W. E., Young, M. J., Weitz, J. S., Whitaker, R. J. CRISPR-induced distributed immunity in microbial populations. *PLoS One* **9**, e101710 (2014).

11. Lively, C. M. The effect of host genetic diversity on disease spread. *Am Nat* **175**, 149-52 (2010).

12. King, K. C., Lively, C. M. Does genetic diversity limit disease spread in natural populations? *Heredity* **109**, 199-203 (2012).

13. Van Baalen M, Beekman M. The costs and benefits of genetic heterogeneity in resistance against parasites in social insects. *Am Nat* **167**, 568–577 (2006).

14. Altermatt F, Ebert D. Genetic diversity of Daphnia magna populations enhances resistance to parasites. *Ecol Lett* **11**, 918–928 (2008).

15. Schmid-Hempel P, Crozier RH. Polygyny versus polyandry versus parasites. *Phil Trans R Soc Lond B Biol Sci* **354**, 507–515 (1999).

16. Levin, B. R., Moineau, S., Bushman, M., Barrangou, R. The population and evolutionary dynamics of phage and bacteria with CRISPR immunity. *PLoS Genet.* **9**, e1003312 (2013).

17. Iranzo, J., Lobkovsky, A. E., Wolf, Y. I., Koonin, E. V. Evolutionary dynamics of the prokaryotic adaptive immune system CRISPR-Cas in an explicit ecological context. *J. Bacteriol.* **195**, 3834-44 (2013).

18. Keesing, F., *et al.* Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* **468**, 647-52 (2010).

19. Datsenko, K. A., *et al.* Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nature Comm.* **3**, 945 (2012).

20. Swarts, D. C., Mosterd, C. van Passel, M. W., Brouns, S. J. CRISPR interference directs strand specific spacer acquisition. *PLoS One*. **7**, e35888 (2012).

21. Fineran, P. C., et al. Degenerate target sites mediate rapid primed CRISPR adaptation. *Proc. Natl Acad Sci USA*. **111**, 1629-38 (2014).

22. Hamilton, W. D., Axelrod, R., Tanese, R. Sexual reproduction as an adaptation to resist parasites (a review). *Proc Natl Acad Sci USA* **87**, 3566–3573 (1990).
23. Pal, C., Maciá, M. D., Oliver, A., Schachar, I., Buckling, A. Coevolution with viruses drives the evolution of bacterial mutation rates. *Nature* **450**, 1079-81 (2007).
24. Morran, L. T., Schmidt, O. G., Gelarden, I. A., Parrish R. C. 2nd, Lively, C. M. Running with the red queen: host-parasite coevolution selects for biparental sex. *Science* **333**, 216-8 (2011).
25. Howard, R. S., Lively, C. M. Parasitism, mutation accumulation and the maintenance of sex. *Nature* **367**, 554-7 (1994).
26. Ashby, B., King, K. C. Diversity and the maintenance of sex by parasites. *J. Evol. Biol.* **28**, 511-20 (2015).
27. Lively, C. An epidemiological model of host-parasite coevolution and sex. *J. Evol. Biol.* **23**, 1490-7 (2010).
28. Peters, A.D., Lively, C. The Red Queen and fluctuating epistasis: a population genetic analysis of antagonistic coevolution. *Am. Nat.* **154**, 393-405 (1999).
29. Vale, P.F., et al. Costs of CRISPR-Cas-mediated resistance in *Streptococcus thermophilus*. *Proc. Biol. Sci.* **282**, 20151270 (2015).
30. Gardner, A. & Grafen, A. Capturing the superorganism: a formal theory of group adaptation. *J. Evol. Biol.* **22**, 659-71 (2009).

Acknowledgements

We thank Dan Morley and Suzanne Kay for experimental contributions and Andy Gardner for comments on the manuscript. SVH received funding from the People

Programme (Marie Curie Actions) of the European Union's Horizon 2020. ERW received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n^{0} [327606]. ERW, AB and MB also acknowledge NERC, BBSRC, the Royal Society, the Leverhulme Trust, the Welcome Trust and the AXA research fund for funding. JBD was supported by the University of California San Francisco Program for Breakthrough in Biomedical Research, the Sandler Foundation, and an NIH Director's Early Independence Award (DP5-OD021344).

Author contributions

EW, AB and SvH designed experiments. SvH, EW AE, JB and HC performed the experiments. SvH, EW and AB analyzed the data and wrote the manuscript. BA and MB contributed to discussions and provided feedback throughout the project. SP performed and analyzed deep sequencing of virus genomes. SG and HC helped to set up the experiments with *Streptococcus thermophilus*. JBD supplied phage with anti-CRISPR gene.

Author information

Reprints and permissions information is available at www.nature.com/reprints. The authors declare no conflict of interest. Correspondence and requests for materials should be addressed to: vanhoute.stineke@gmail.com, A.J.Buckling@exeter.ac.uk or westra.edze@gmail.com.

Figure legends

Figure 1

Evolution of CRISPR-mediated immunity leads to rapid extinction of virus. Titer (pfu/ml) of virus DMS3vir over time upon infection of **A)** WT *P. aeruginosa* and **B)** *P. aeruginosa* strain *csy3::LacZ* (CRISPR KO strain). Each line indicates an individual replicate experiment (n=6). The limit of detection is 200 pfu/ml.

Figure 2

Virus persistence inversely correlates with the level of spacer diversity. Virus titers (pfu/ml) over time upon infection of a bacterial population consisting of an equal mixture of a surface mutant and **A)** a monoculture with CRISPR-mediated immunity (n=48), or polycultures with CRISPR-mediated immunity consisting of **B)** 6 clones (n=8), **C)** 12 clones (n=8), **D)** 24 clones (n=6), **E)** 48 clones (n=6). The number of replicates is chosen such that all clones are equally represented in each treatment. Each line indicates an individual replicate experiment. The limit of detection is 200 pfu/ml.

Figure 3

Relative fitness of bacterial populations with CRISPR-mediated immunity positively correlates with increasing spacer diversity. Relative fitness of bacterial populations with CRISPR-mediated immunity, with spacer diversity as indicated, at 3 days post-infection when competing with a surface mutant. Error bars indicate 95% confidence intervals.

Figure 4

Emergence of virus that overcomes host CRISPR immunity (escape virus) during the experiment shown in Figures 2 and 3. Each column in a table represents a time point where virus was isolated (0, 16, 24, 40, 48, 64 and 72 hours post-infection, as indicated below the table (in days post-infection)). Green: no escape virus. Red: escape virus. Panels A-E correspond to each of the experiments shown in Figure 2 A-E. Bold numbers indicate replicate experiments. Numbers between parentheses indicate the identity of the clones that are present in the CRISPR population. Asterisks indicate that virus went extinct during the experiment.

Figure 1

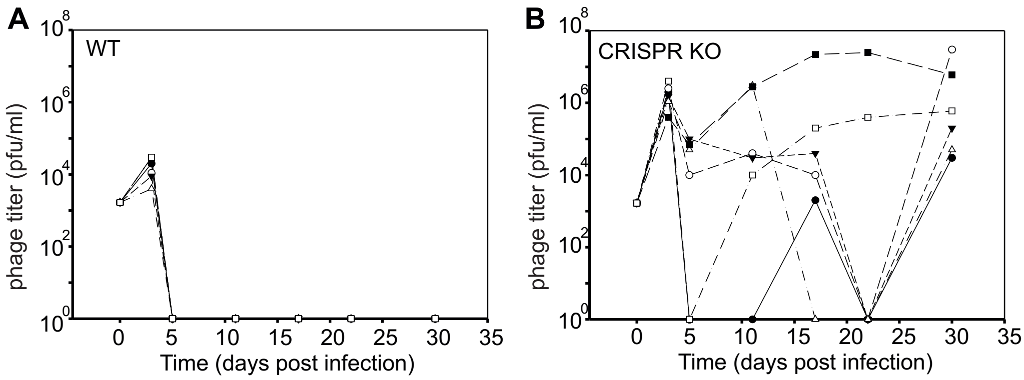
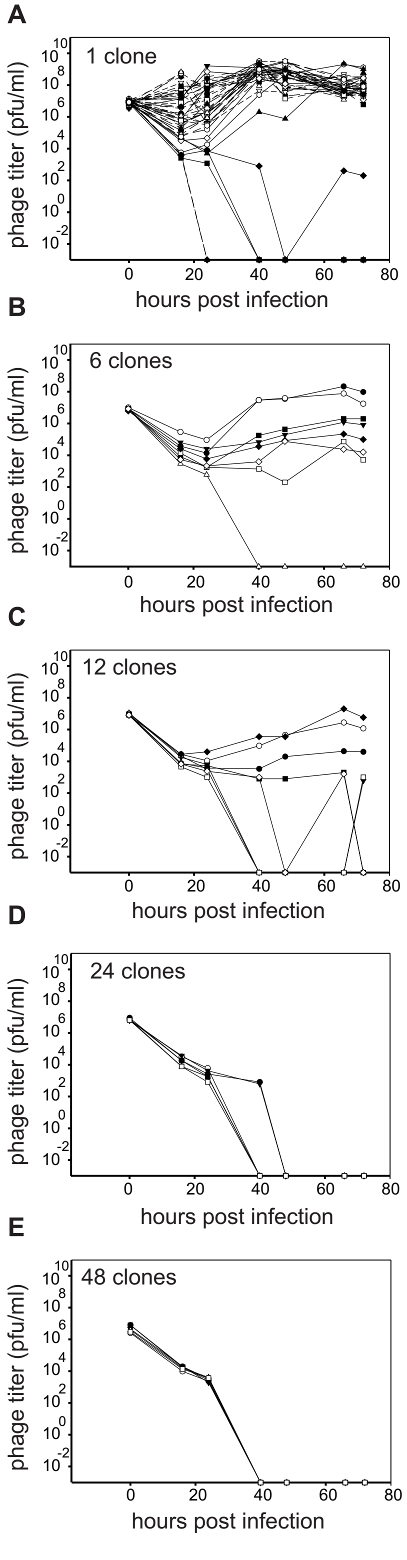

A
WT
phage titer (pfu/ml)
10^8
10^6
10^4
10^2
10^0
0 5 10 15 20 25 30 35
Time (days post infection)
B
CRISPR KO
phage titer (pfu/ml)
10^8
10^6
10^4
10^2
10^0
0 5 10 15 20 25 30 35
Time (days post infection)

Figure 2



A
1 clone
B
6 clones
C
12 clones
D
24 clones
E
48 clones
phage titer (pfu/ml)
hours post infection
0
20
40
60
80
10^{10}
10^{8}
10^{6}
10^{4}
10^{2}
10^{0}
10^{-2}

Figure 3

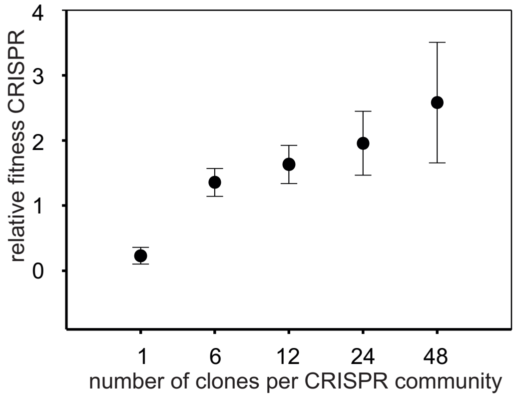
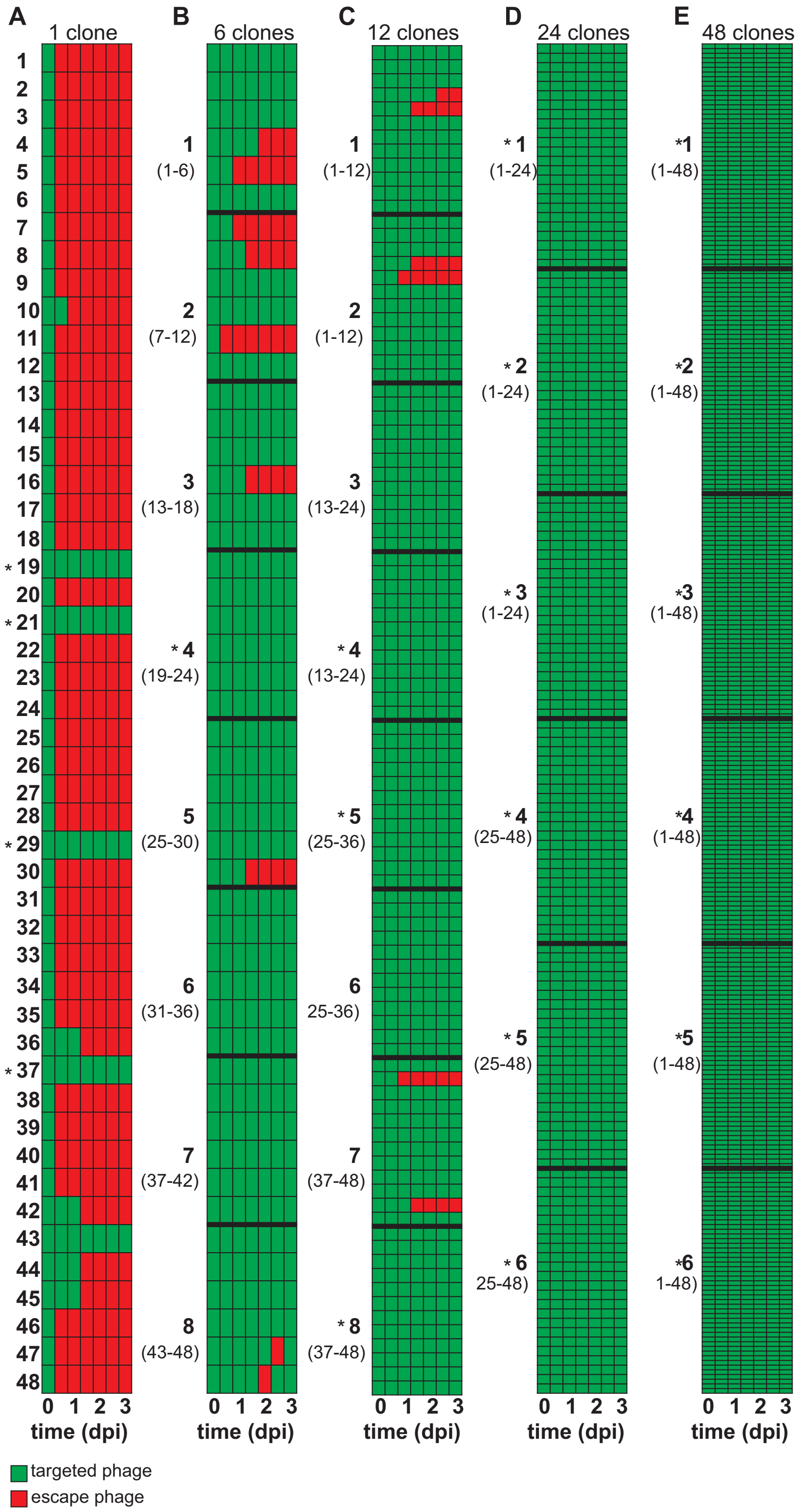

relative fitness CRISPR
0
1
2
3
4
1
6
12
24
48
number of clones per CRISPR community

Figure 4


A
1 clone
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
* 19
20
* 21
22
23
24
25
26
27
28
* 29
30
31
32
33
34
35
36
* 37
38
39
40
41
42
43
44
45
46
47
48
B
6 clones
1 (1-6)
2 (7-12)
3 (13-18)
* 4 (19-24)
5 (25-30)
6 (31-36)
7 (37-42)
8 (43-48)
C
12 clones
1 (1-12)
2 (1-12)
3 (13-24)
* 4 (13-24)
* 5 (25-36)
6 25-36)
7 (37-48)
* 8 (37-48)
D
24 clones
* 1 (1-24)
* 2 (1-24)
* 3 (1-24)
* 4 (25-48)
* 5 (25-48)
* 6 25-48)
E
48 clones
*1 (1-48)
*2 (1-48)
*3 (1-48)
*4 (1-48)
*5 (1-48)
*6 1-48)
0 1 2 3
time (dpi)
targeted phage
escape phage

Extended Data

The diversity-generating benefits of a prokaryotic adaptive immune system

Stineke van Houte[1], Alice K.E. Ekroth[1], Jenny M. Broniewski[1], Hélène Chabas[1,2], Ben Ashby[3], Joe Bondy-Denomy[4], Sylvain Gandon[2], Mike Boots[3,5], Steve Paterson[6], Angus J. Buckling[1] and Edze R. Westra[1]

[1] Environment and Sustainability Institute, Centre for Ecology and Conservation, Biosciences, Cornwall Campus, Penryn, UK.

[2] CEFE UMR 5175, CNRS-Université de Montpellier, Université Paul-Valéry Montpellier, EPHE, 1919, route de Mende 34293 Montpellier Cedex 5, France

[3] Centre for Ecology and Conservation, Biosciences, Cornwall Campus, Penryn, UK.

[4] Department of Microbiology and Immunology, University of California, San Fransisco, 94158, San Francisco, US.

[5] Department of Integrative Biology, University of California, Berkeley, US.

[6] Institute of Integrative Biology, University of Liverpool, UK.

Extended Data Table 1: Tukey HSD all pairwise comparisons of the data in Figure 3. 1 = monoculture, 6 = 6-clone polyculture, 12 = 12-clone polyculture, 24 = 24-clone polyculture, 48 = 48-clone polyculture

<table>
<tr><th colspan="2">Comparison</th><th>Difference</th><th>Std Error</th><th>t Ratio</th><th>Prob>|t|</th><th>Lower 95%</th><th>Upper 95%</th></tr>
<tr><td>1</td><td>6</td><td>-1.12680</td><td>0.2141986</td><td>-5.26</td><td><.0001*</td><td>-1.72637</td><td>-0.52724</td></tr>
<tr><td>1</td><td>12</td><td>-1.40303</td><td>0.2141986</td><td>-6.55</td><td><.0001*</td><td>-2.00259</td><td>-0.80346</td></tr>
<tr><td>1</td><td>24</td><td>-1.72790</td><td>0.2428783</td><td>-7.11</td><td><.0001*</td><td>-2.40775</td><td>-1.04806</td></tr>
<tr><td>1</td><td>48</td><td>-2.35252</td><td>0.2428783</td><td>-9.69</td><td><.0001*</td><td>-3.03236</td><td>-1.67267</td></tr>
<tr><td>6</td><td>12</td><td>-0.27622</td><td>0.2804518</td><td>-0.98</td><td>0.8612</td><td>-1.06124</td><td>0.50879</td></tr>
<tr><td>6</td><td>24</td><td>-0.60110</td><td>0.3029225</td><td>-1.98</td><td>0.2842</td><td>-1.44901</td><td>0.24682</td></tr>
<tr><td>6</td><td>48</td><td>-1.22571</td><td>0.3029225</td><td>-4.05</td><td>0.0012*</td><td>-2.07363</td><td>-0.37780</td></tr>
<tr><td>12</td><td>24</td><td>-0.32488</td><td>0.3029225</td><td>-1.07</td><td>0.8200</td><td>-1.17279</td><td>0.52304</td></tr>
<tr><td>12</td><td>48</td><td>-0.94949</td><td>0.3029225</td><td>-3.13</td><td>0.0205*</td><td>-1.79741</td><td>-0.10158</td></tr>
<tr><td>24</td><td>48</td><td>-0.62462</td><td>0.3238378</td><td>-1.93</td><td>0.3119</td><td>-1.53108</td><td>0.28184</td></tr>
</table>

Extended data Figure 1

Infection with virus DMS3vir leads to rapid evolution of CRISPR-mediated immunity in WT bacteria, while CRISPR KO bacteria primarily evolve virus immunity by surface mutation. Percentage bacteria at 5 days post-infection that have evolved immunity by CRISPR-Cas (white bar), surface mutation (black bar) or that have not evolved immunity (sensitive; grey bars). Error bars indicate 95% confidence intervals (CI).

Extended data Figure 2

No benefit of increasing spacer diversity in the absence of virus. Relative fitness of CRISPR immune monocultures (single spacer; low diversity) and polycultures (48 spacers; high diversity) at 3 days post-infection when competing with a surface mutant (*sm*) in the absence of virus. Error bars indicate 95% CI.

Extended data Figure 3

Deep sequencing analysis of the frequency of mutations in the target sequence (seed sequence and the adjoining PAM) of virus isolated at t=1 from the experiment shown in Fig. 4. **A)** Frequency of point mutation in the single target sequence of a viral population isolated from the monocultures of clones 1-3. **B)** Average frequency of point mutation across all target sites in the ancestral virus genome and in the genomes of virus from pooled samples of all replicates from a single diversity treatment. Error bars indicate 95% (CI).

Extended data Figure 4

Escape virus titers decline upon infection of diverse CRISPR populations despite increased fitness over ancestral virus. Escape virus was isolated from monocultures of clones 1-6 competing with the surface mutant at 24 hpi (Fig. 3 and Extended Data Fig. 2). **A)** Virus titers (pfu/ml) over time upon infection with approximately 10^7 pfu individual escape virus or ancestral virus of a bacterial population consisting of a monoculture (dotted line) or 48-clone polyculture (solid line). **B)** Relative fitness of escape virus and ancestral virus during infection of a CRISPR resistant monoculture or polycultures consisting of 48 clones. All experiments were performed in 6 replicates. Error bars indicate 95% CI. The limit of detection is 200 pfu/ml.

Extended data Figure 5

Diverse populations do not acquire additional spacers during the experiments shown in Figures 2-4. For each diversity treatment we examined the spacer content of 192 randomly isolated clones at both t=0 and t=3 (384 clones in total per diversity treatment). The change in the total number of spacers between t=0 and t=3 was calculated independently for each replicate experiment and divided by the number of clones that were examined. The graph indicates the average across the replicates of the change in spacer content per clone and error bars indicate 95% CI.

Extended data Figure 6

Persistence of phage that encodes an anti-CRISPR gene is independent of spacer diversity. **A)** Virus titers (pfu/ml) over time upon infection of a bacterial population consisting of an equal mixture of a surface mutant and **A)** a monoculture with CRISPR-mediated immunity (n=48) or **B)** a 48-clone polyculture with CRISPR-mediated immunity (n=6). Each clone is equally represented in each treatment. Each line indicates an individual replicate experiment. The limit of detection is 200 pfu/ml. **C)** The number of replicate experiments in which the CRISPR immune population went extinct (no detectable white colonies) at 1 and 3 dpi.

Extended data Figure 7

Virus persistence inversely correlates with the level of CRISPR spacer diversity in CRISPR immune populations of *Streptococcus thermophilus*. Virus titers (pfu/ml) over time upon infection of a bacterial population consisting of **A)** a monoculture with CRISPR-mediated immunity (n=44) or **B)** 44-clone polycultures with CRISPR-mediated immunity (n=28). Each clone is equally represented in each treatment. Each line indicates an individual replicate experiment. The limit of detection is 200 pfu/ml. **C)** OD600 of monocultures and polycultures at 1 and 2 days post infection. Error bars indicate 95% confidence intervals. **D)** Emergence of virus mutants that overcome CRISPR-mediated immunity after 0, 16, and 24 hours post-infection. Green indicates no escape virus. Red indicates emergence of escape virus. All polyculture experiments showed no escape virus.

Extended data Figure 8

Sensitive bacteria are unable to invade bacterial populations with CRISPR-mediated immunity in the presence of virus. Relative fitness of CRISPR populations with indicated spacer diversity at 3 days post-infection when competing with the sensitive CRISPR KO strain. Relative fitness of CRISPR populations decreases with increasing spacer diversity due to the rapid virus extinction, which benefits sensitive bacteria, but is higher than 1 in all cases. Error bars indicate 95% CI.

Extended data Figure 9

Virus persistence inversely correlates with the level of CRISPR spacer diversity during competition between CRISPR immune populations and the sensitive CRISPR KO strain. Virus titers (pfu/ml) over time upon infection of a bacterial population consisting of an equal mixture of a CRISPR KO clone and **A)** a monoculture with CRISPR-mediated immunity (n=48), or polycultures with CRISPR-mediated immunity consisting of **B)** 6 clones (n=8), **C)** 12 clones (n=8), **D)** 24 clones (n=6), **E)** 48 clones (n=6). The number of replicates is chosen such that all clones are equally represented in each treatment. Each line indicates an individual replicate experiment. The limit of detection is 200 pfu/ml.

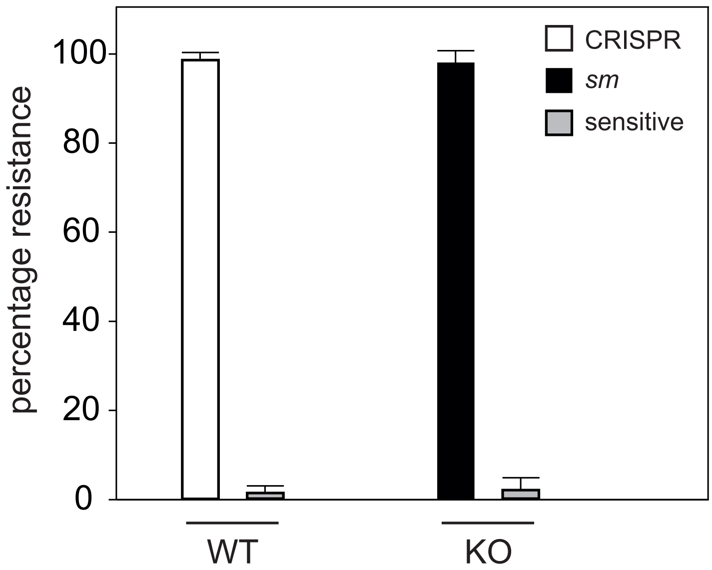
Extended data Figure 10

Emergence of virus mutants that overcome CRISPR-mediated immunity during the experiment shown in Extended Data Figure 9. Each column in a table represents a time point (0, 16, 24, 40, 48, 64 and 72 hours post-infection, as indicated below the table (in days post-infection)) where virus was isolated. Green indicates no escape virus. Red indicates emergence of escape virus. Panels A-E correspond to each of the experiments shown in Extended Data Figure 9 A-E. Bold numbers indicate replicate experiments. Numbers between parentheses indicate the identity of clones that are present in a population with CRISPR-mediated immunity. Asterisks indicate replicate experiments where virus went extinct during the experiment.

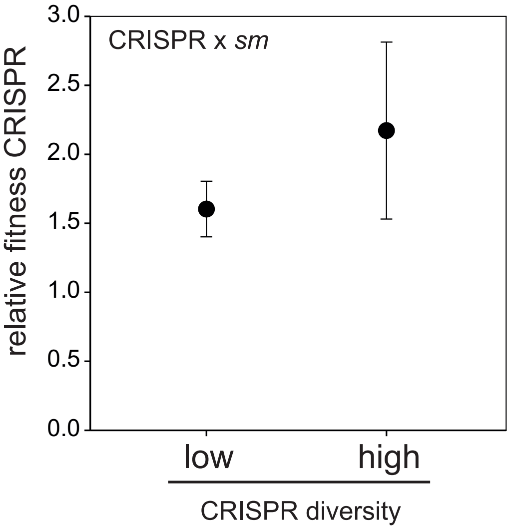
Extended data Figure 11

Sensitive bacteria are unable to invade bacterial populations with CRISPR-mediated immunity in the absence of virus, independent of the level of spacer diversity. Relative fitness of monoculture (single spacer; low diversity) and polyculture (48 spacers; high diversity) at 3 days post-infection when competing with the CRISPR KO strain (sensitive) in the absence of virus. Error bars indicate 95% CI.

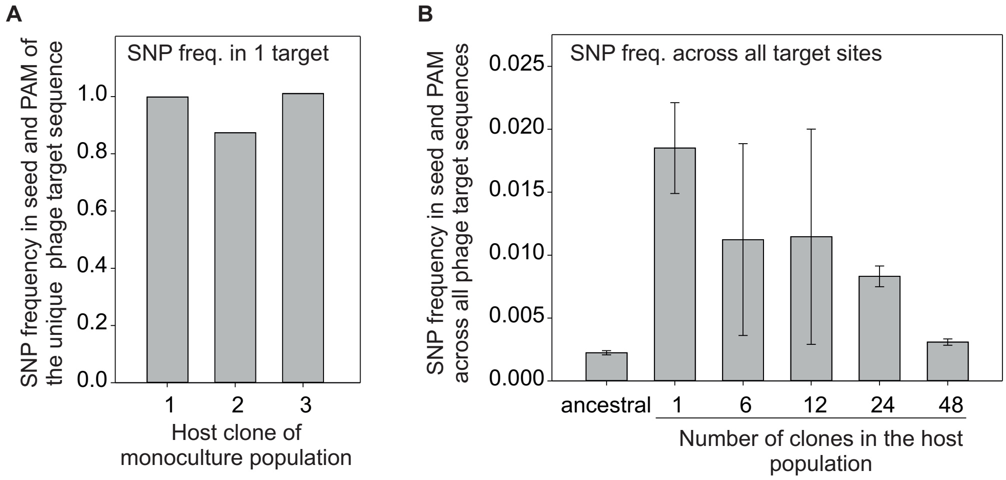
Extended data Figure 1



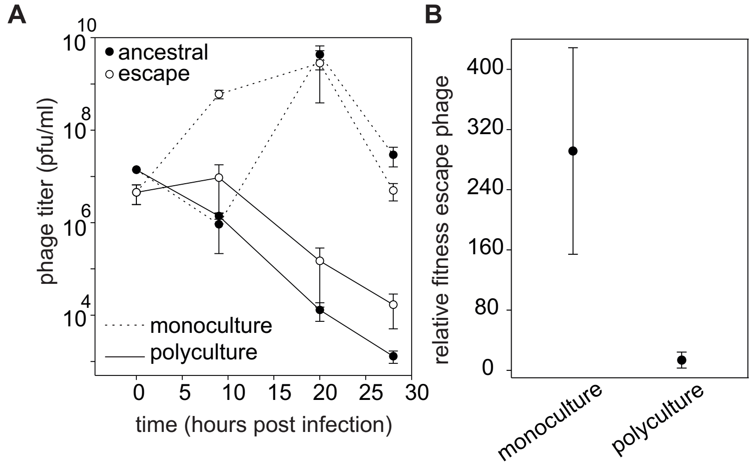
Extended Data Figure 2



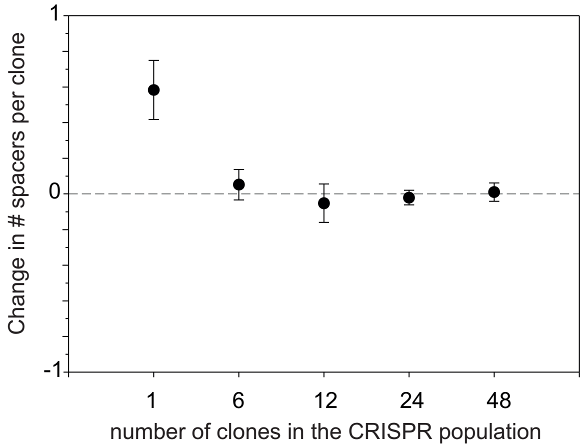
Extended data Figure 3



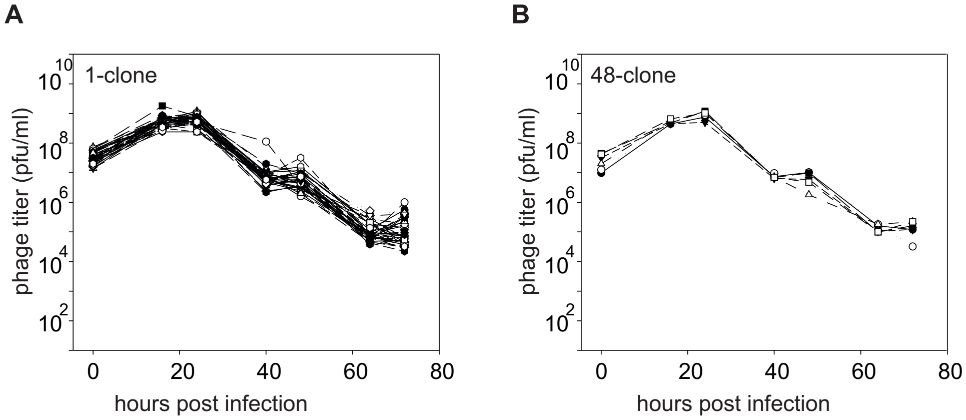
Extended Data Figure 4



Extended Data Figure 5



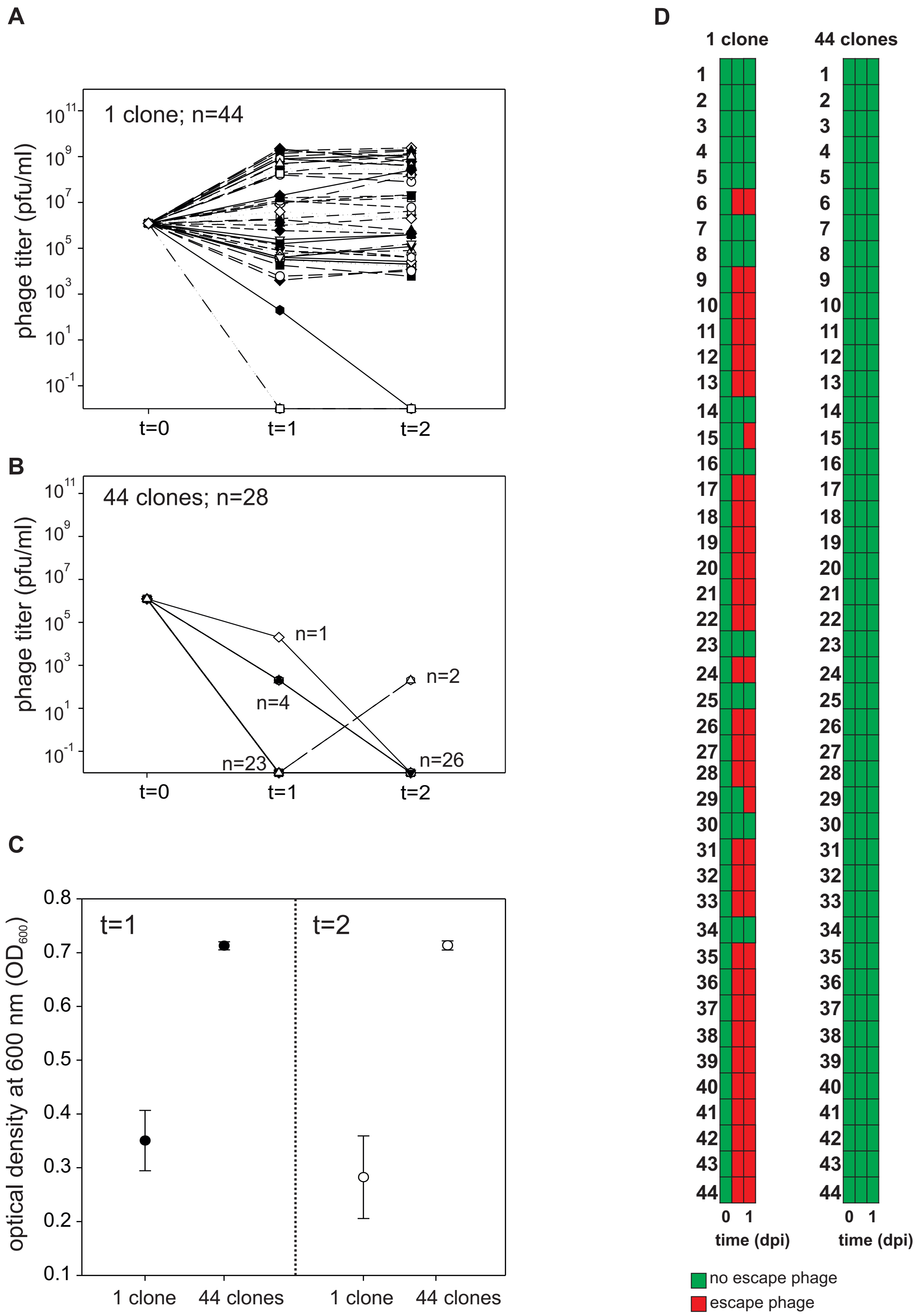
Extended data Figure 6



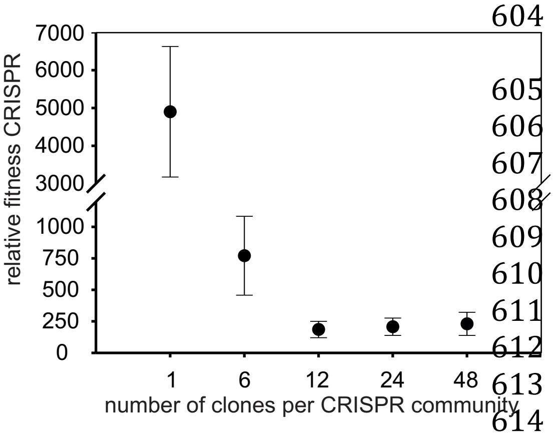
C

	monoculture		48-clone polyculture	
	extinct	non-extinct	extinct	non-extinct
t=1	37	11	5	1
t=3	37	11	6	0

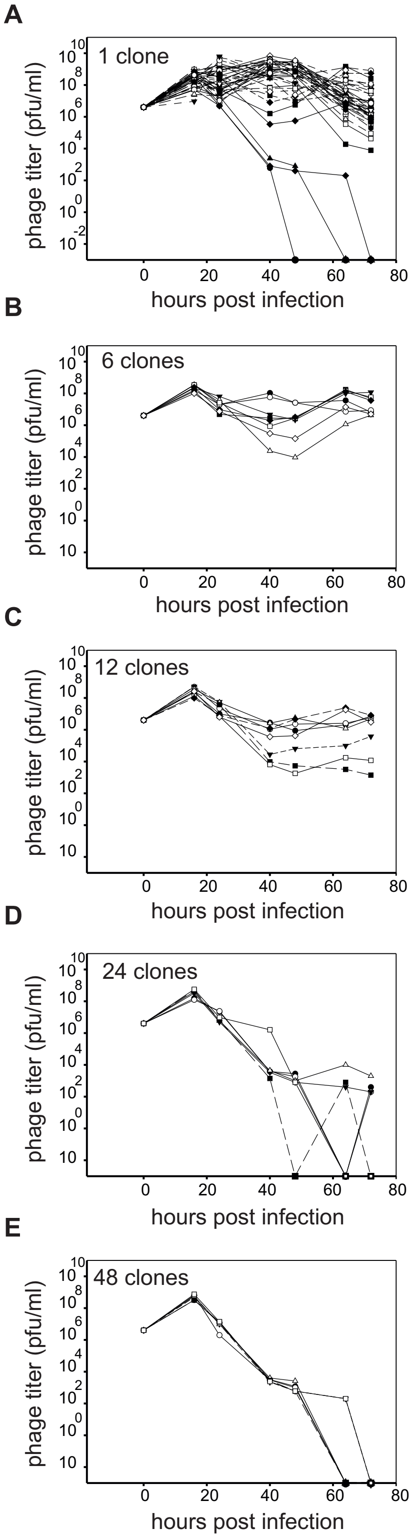
Extended data Figure 7



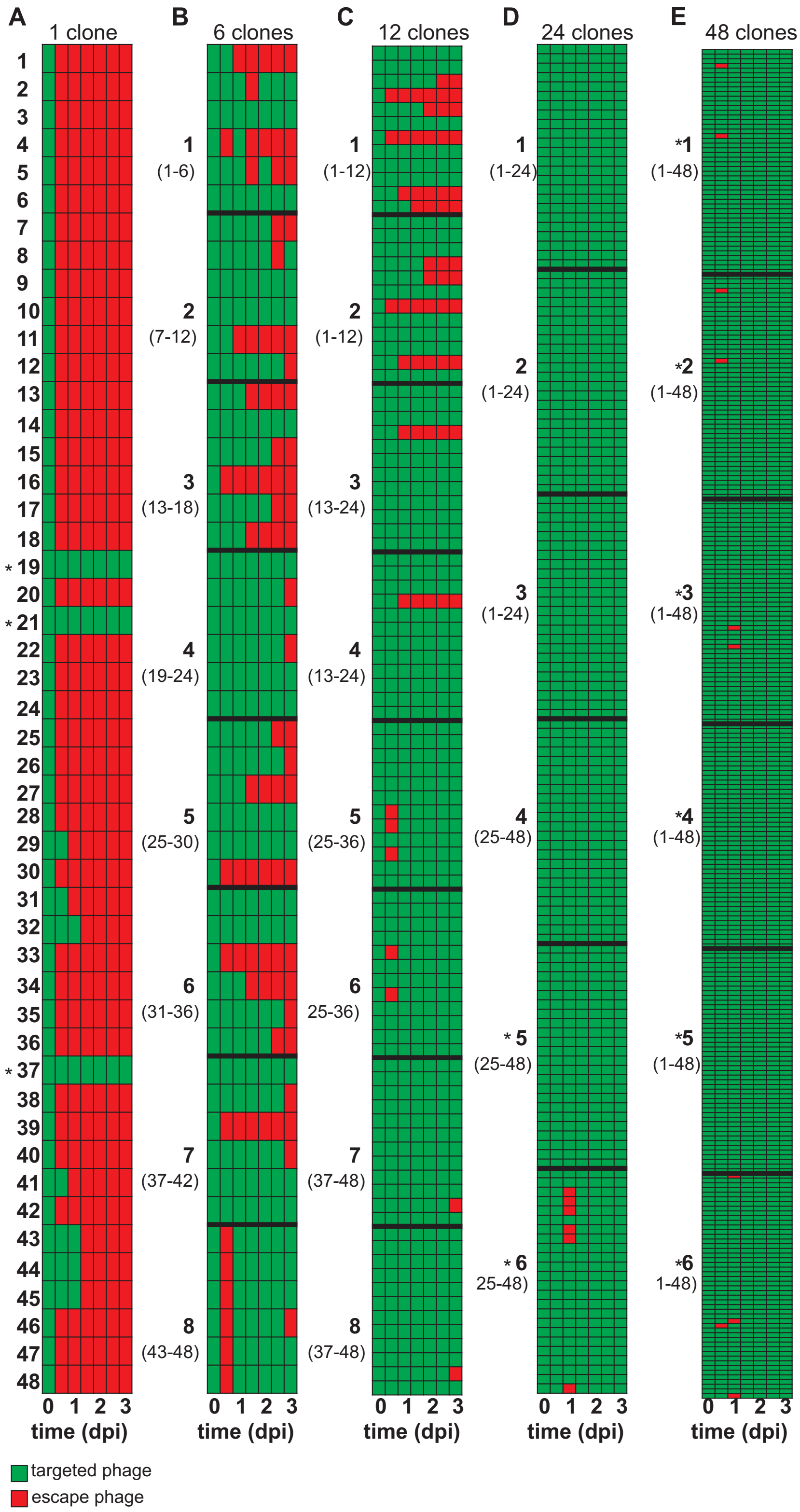
Extended data Figure 8



Extended data Figure 9



Extended data Figure 10



Extended data Figure 11