Different genetic and morphological outcomes for phages targeted by single or multiple CRISPR-Cas spacers

B.N.J. Watson[1], R.A. Easingwood[2], B. Tong[1], M. Wolf[3], G.P.C. Salmond[4], R.H.J Staals[1,5], M. Bostina[1,2]* and P.C. Fineran[1,6]*

[1]Department of Microbiology and Immunology, University of Otago, PO Box 56, Dunedin 9054, New Zealand.
[2]Otago Centre for Electron Microscopy, University of Otago, PO Box 56, Dunedin 9054, New Zealand.
[3]Molecular Cryo-Electron Microscopy Unit, Okinawa Institute of Science and Technology, Graduate University, Onna, Okinawa, Japan.
[4]Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, United Kingdom.
[5]Laboratory of Microbiology, Department of Agrotechnology and Food Sciences, Wageningen University, 6708 HB Wageningen, The Netherlands.
[6]Bio-Protection Research Centre, University of Otago, Dunedin, New Zealand.

*For correspondence: mihnea.bostina@otago.ac.nz or peter.fineran@otago.ac.nz

Keywords: CRISPR-Cas, bacteriophages, tape measure protein, phage evolution, phage morphology

Summary

CRISPR-Cas systems provide bacteria and archaea with adaptive immunity against genetic invaders, such as bacteriophages. The systems integrate short sequences from the phage genome into the bacterial CRISPR array. These 'spacers' provide sequence-specific immunity but drive natural selection of evolved phage mutants that escape CRISPR-Cas defence. Spacer acquisition occurs by either naïve or primed adaptation. Naïve adaptation typically results in the incorporation of a single spacer. In contrast, priming is a positive feedback loop that often results in acquisition of multiple spacers, which occurs when a pre-existing spacer matches the invading phage. We predicted that single and multiple spacers, representative of naïve and primed adaptation respectively, would cause differing outcomes after phage infection. We investigated the response of two phages, ϕTE and ϕM1, to the *Pectobacterium atrosepticum* type I-F CRISPR-Cas system and observed that escape from single spacers typically occurred via point mutations. Alternatively, phages escaped multiple spacers through deletions, which can occur in genes encoding structural proteins. Cryo-EM analysis of the ϕTE structure revealed shortened tails in escape mutants with tape measure protein deletions. We conclude that CRISPR-Cas systems can drive phage genetic diversity, altering morphology and fitness, through selective pressures arising from naïve and primed acquisition events.

Introduction

Phages are abundant, genetically diverse and are expected to evolve faster than their bacterial hosts (1, 2). In response, bacteria employ different strategies to limit phage predation, including CRISPR-Cas systems (3, 4). Briefly, CRISPR-Cas immunity is mediated through three stages. Firstly, during adaptation, fragments of invading phage DNA are incorporated into the CRISPR array (5). During the expression phase, the Cas proteins are produced and the CRISPR array is transcribed and processed into CRISPR RNAs (crRNAs) (6). The crRNAs associate with Cas effector proteins or protein complexes to survey the cell for sequence complementary to the crRNA. Finally, in interference, phage genomic material is recognised and degraded by Cas nucleases (6, 7) (for recent reviews, see (5, 8, 9)).

The initial generation of immunity against phages occurs through naïve adaptation, which will typically result in the incorporation of a single spacer into the CRISPR array (5, 7). Consequently, CRISPR-Cas targeting has been shown to select for phages with point mutations in positions of the protospacer and protospacer adjacent motif (PAM) that are required for interference (7, 10-19). Additionally, small deletions or duplications have led to phage escape from CRISPR-Cas (11, 15, 18). In type I CRISPR-Cas systems, point mutations that abolish interference can activate primed adaptation (or priming), during which, multiple new spacers are incorporated into the CRISPR array to provide effective resistance against the invader (11, 20-22). Consequently, it is predicted to be more difficult for phages to escape from bacterial strains that have undergone primed acquisition. A recent bioinformatic study observed that priming is widespread and many CRISPR arrays contain multiple spacers targeting nearby regions within the phage genome (23). Evidence of priming was confirmed in four type I systems, I-B, I-D, I-E and I-F, as well as two type II systems, II-A and II-C, demonstrating that in many CRISPR-Cas systems, multiple spacers will target a single gene. While some studies have observed phage escape from strains with single phage targeting spacers, representing naïve adaptation, resistance acquired through primed adaptation is predicted to provide a different selection pressure for phages to overcome.

To investigate genetic changes occurring in phage populations overcoming single (naïve) or multiple (primed) spacers targeting different phage genes we used two phages, ϕTE and ϕM1, to infect *Pectobacterium atrosepticum* containing a type I-F CRISPR-Cas system (24, 25). Point mutations typically allowed phages to overcome a single spacer, whereas escape through duplications and deletions were less frequently observed. In contrast, phages only escaped multiple spacers through deletions, some of which were quite large, losing up to 61% of the targeted gene. When these deletions occurred in structural genes, the escape phages exhibited altered morphology as observed by electron microscopy. Therefore, naïve and primed adaptation provide different selection pressures that drive the evolution of phages and generate parasite genetic diversity in response to bacterial CRISPR-Cas defences.

Methods

Strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are given in table S1 and S2. *P. atrosepticum* SCRI1043 (26) was grown at 25°C and *E. coli* at 37°C in lysogeny broth (LB) at 180 rpm or on LB-agar (LBA) plates containing 1.5% (w v^{-1}) agar. When required, media were supplemented with tetracycline (Tc; 10 µg ml^{-1}) and isopropyl-ß-D-thiogalactopyranoside (IPTG) at 1 mM. Bacterial growth was measured in a Jenway 6300 spectrophotometer at 600 nm (OD_{600}).

Generation of strains with phage-targeting spacers

Phage genes were cloned into a plasmid, pPF712, containing a priming protospacer (spacer 1 from CRISPR1, with a TG PAM), a Tc resistance cassette and mCherry (27), which enabled us to exploit priming to obtain phage resistant strains that are rare under laboratory conditions (22, 28). The ϕTE genes targeted included the DNA polymerase (phiTE_10, 1,998 bp), a putative tail fibre (phiTE_215, 3,300 bp), which contains three collagen triplex helix repeats that are found in phage fibrillar proteins (29-31) and the putative tape measure protein gene (phiTE_228, 2,436 bp) (figure 1*a* - *c, f*). Genes targeted for ϕM1 included an RNA polymerase (phiM1_31, 2,445 bp) and a putative tail fibre (phiM1_44, 1,603 bp) (figure S1*a* - *c*) (32). The oligonucleotides and plasmids details are listed in table S2. *P. atrosepticum* was transformed with the priming vectors, followed by selection on Tc. Transformants were then grown in cultures without selection, with daily subculturing and plating onto LBA containing IPTG, to visualise colonies that had lost the vector (mCherry expression). Clones not expressing mCherry were patched onto LBA ± Tc to confirm loss of the plasmid and the CRISPR arrays of Tc-sensitive clones were amplified using PCR to detect spacer acquisition, using the following primer combinations: PF174 and PF175 for CRISPR1, PF176 and PF177 for CRISPR2 and PF178 and PF179 for CRISPR3. The expanded CRISPR arrays were sequenced, using PF175 for CRISPR1 and PF177 for CRISPR2, and new spacers were mapped to the relevant plasmid and phage genome(s). Strains with between one and four new phage-targeting spacers (anti-ϕ strains) were isolated, to represent strains that had acquired phage resistance through naïve and primed acquisition (figure 1*b* & *c*, figure S1*b* & *c*, table S1). Predicted functional protein domains encoded within the phage genes (figure 1) were identified using NCBI Conserved Domain search (33).

Phage storage, titration and adsorption assays

ϕTE (142, 349 bp) (34) and ϕM1 (43,827 bp) (32, 35) were stored in phage buffer (10 mM Tris-HCl pH 7.4, 10 mM $MgSO_4$ and 0.01% w v^{-1} gelatin). Phage stocks were titrated by serially-diluting phages in phage buffer, adding to 100 µl of *P. atrosepticum* culture (pre-grown in 5 ml LB overnight) in 4 ml top LBA (0.35% (ϕTE) and 0.5% (ϕM1) agar) and pouring onto LBA plates. Once set, plates were incubated at 25°C overnight, plaques were counted and the titre determined as plaque forming units (pfu) ml^{-1}. Efficiency of plating (EOP) was determined as: (pfu ml^{-1} (test strain) / pfu ml^{-1} (control strain, *P. atrosepticum*)). The plaques produced on some anti-ϕTE were tiny pin-pricks, which were difficult to accurately count. To determine EOP on these hosts, tiny plaques obtained for a proportion of the plate were counted, and taking into account the absence of plaques on lower dilutions, reproducible estimates were able to be made within an order of magnitude. Adsorption was assessed by infecting exponential phase wild type *P. atrosepticum* cultures with phages at an MOI of ~0.1. Samples taken at several time points were added to phage buffer containing chloroform to lyse cells. Phage adsorption was determined as ((pfu ml^{-1} (t=0 min) - pfu ml^{-1}

(t=0 min to 40 min) / pfu ml^{-1} (t=0 min)). The change in pfu over time was determined as (pfu ml^{-1} (t=0 min to 70 min) / pfu ml^{-1} (t=0 min).

Classification of partial resistance and genotypic escape phages

Following observation and characterisation of the plaque morphologies on anti-ϕ strains (figures 1 and 2, table S3), genotypic escape phages and those that had not heritably escaped were identified based on the following criteria. ϕTE phages that had not heritably escaped had a reduced EOP when retitrated, to the EOP value initially observed for the strain and produced plaques smaller than those on the control (*P. atrosepticum*). Additionally, plaques had defined edges, although on some anti-ϕTE strains, plaques were too small to properly observe the edges. ϕTE genotypic escape phages were typically the size of plaques produced on the control strain (*P. atrosepticum*) and the plaques all had undefined, turbid margins, which is also characteristic of plaques formed on the control. In this study, all plaques on the anti-ϕM1 strains were formed by phages that had not heritably escaped CRISPR-Cas. The representative plaques that were retitrated from anti-ϕM1 strains had a reduced EOP, to the level initially observed for the strains. Plaques were smaller than those produced on the control strain and plaque edges were not altered. To titrate both genotypic escape phages and phages that had not heritably escaped CRISPR-Cas, phages were first picked from the centre of the plaques using a toothpick and phages were suspended in phage buffer. Phages were serially diluted and 10 µl was spotted onto lawns of both the anti-ϕ strains and the WT control strain. Genotypic escape phages were plaque purified at least twice before being sequenced, with the primers used to amplify the phage genes (table S2). Each escape phage infected the anti-ϕ host strain with an EOP of 1 (table S4).

Negative stained electron microscopy

For negative stain electron microscopy, 10 µl of high-titre, highly concentrated lysate samples were loaded onto carbon-coated 300-mesh copper grids for 60 s. These grids were blotted to remove excess specimen, then stained with 10 µl of 2% phosphotungstic acid solution (pH 7) and blotted dry. A Philips CM100 Transmission electron microscope operated at 100 kV with a magnification of approximately 66,000 × was used to record micrographs. Over 30 measurements of lengths of individual tails for wild type and mutant phages were measured in 3dmod (36).

Cryo-electron microscopy

Quantifoil holey carbon grids were glow discharged (to increase hydrophilicity) and then 3 µl of purified phage sample was applied to each grid. Excess buffer from the sample was blotted away with filter paper and the grid immediately flash frozen by plunging into liquid ethane (cooled down to less than -160°C by liquid nitrogen) using a Vitrobot Mark IV plunging device. Grids were visualised using a Titan Krios operated at 300kV using the Leginon automated acquisition software (37) on a K2 Summit camera in counting mode using an energy filter with 20eV slit at a magnification corresponding to a calibrated pixel size of 1.39 Å. A number of 1,800 digital micrographs movies were recorded at 125ms/frame with 64 frames per exposure corresponding to a total dose of 47 $e^-/Å^2$.

Individual frames were aligned using Motioncor2 (38) using patch alignment, all frames being included in the final dose-weighted average. Contrast transfer function parameters were estimated using ctfind4 (39). Image processing was done in cisTEM (40). Automatic particle selection of capsid resulted in 5,590 particles. After 20 rounds of 2D classification full and empty capsid were separated in separate classes. A total of 10 rounds of 3D refinement with imposed icosahedral symmetry were performed for images of the full

capsid. Two independently refined volumes were calculated in order to estimate the resolution using the Fourier Shell Ring correlation with a 0.143 cutoff giving an estimated resolution of 7.1 Å. *UCSF Chimera* (41) was used to visualise the resultant output maps to generate the final images.

Results

Phage-targeting spacers reduce phage infectivity

To investigate the phage response to single or multiple spacers, representing typical naïve and primed adaptation events respectively, we isolated *P. atrosepticum*, strains with one or more spacers in the I-F CRISPR-Cas system targeting two different regions in ϕTE (a *Myoviridae*) (34) and ϕM1 (a *Podoviridae*) (32, 35). The ϕTE genes targeted included genes encoding a DNA polymerase, a putative tail fibre and the tape measure protein (29, 34) (figure 1*a* – *c*, *f*). Genes targeted in ϕM1 encoded an RNA polymerase and a putative tail fibre (figure S1*a* - *c*) (32). For both phages, strains with one phage-targeting spacer provided protection, from ~3 to 100-fold and additional spacers typically improved resistance, up to 10^5-fold for ϕTE (figure 1*d*, *e* & *g*) and 10^3-fold for ϕM1 (figure S1*d* & *e*). In addition, on the strains with single spacers, the plaques formed by both phages were smaller than when grown on the WT phage-sensitive control (figure 1*d*, *e* & *g*, figure S1*d* & *e*, table S3). As spacer number increased, plaque sizes were reduced further, which was more pronounced with ϕTE. However, some larger plaques were present on some anti-ϕTE strains, suggesting that these phages replicated unhindered (e.g. TE TFP 2 and 6 (figure 1*e*)). In summary, strains with spacers targeting several regions of both ϕTE and ϕM1 reduced phage infectivity and plaque size, with multiple spacers providing more resistance than a single spacer. Therefore, strains with multiple spacers representative of priming provide stronger resistance than single spacers that are typical of naïve adaptation.

Phages escape strains with single spacers through point mutation

To determine how phages were propagating on the strains with single spacers (i.e. typical of naïve acquisition), phages from five independent plaques formed on representative phage resistant strains were titrated on the WT and their corresponding anti-ϕ strain from where they were isolated. For ϕM1, ~3-fold reductions in infection and small plaques were detected, similar to those observed for the original phage (figure S1*d* & *e* and figure S2*a* - *e*). Likewise, a similar result was observed for a strain with a single spacer targeting ϕTE (TE DNP 1) (figure S2*c*). Therefore, phage propagation was still hindered to the same level, indicating that these phages had not genetically escaped CRISPR-Cas interference. Apparently, interference is weak when only a single spacer targets these phage genes, which results in roughly one third of phage infections being able to outpace the CRISPR-Cas-mediated DNA degradation and form small plaques. This might be due to insufficient expression of Cas complexes loaded with the appropriate phage-targeting crRNA, relative to phage replication. Interestingly, a similar phenomenon has been observed with anti-CRISPRs, which can cause partial CRISPR-Cas immunity against phages (42, 43) and can reduce plaque size (44).

For some strains with a single spacer (e.g. TE TFP 2), rarer large plaques were visible in addition to the smaller plaques (figure 1*e* and figure S2*d* & *e*). When phages from these larger plaques were isolated, they robustly re-infected the strain with the single spacer, indicating that they bred true, as genotypic escapes. Sequencing the target regions in all of these escape phages revealed point mutations in the -1 and -2 PAM positions, as well as positions 1-5 and 7 of the protospacer (figure 2*a*, table S4), indicating the importance of these seed positions for interference in the type I-F system, hinting at a seed similar to that of the type I-E systems (10). One phage escaped targeting due to three mutations in positions 25, 26 and 32 of the protospacer (table S4, esc ϕTE19). Overall, the point mutations mostly resulted in missense mutations (65%; 15 out of 23), but some silent mutations also occurred (figure S3*a*, table S4). Phages also overcame targeting by duplication of a 15 bp sequence in the middle of the protospacer (table S4, esc ϕTE17 and ϕTE18). Four phages had escaped

targeting via deletions ranging from ~100 to 400 bp (figure S2*g*). In summary, phages mostly overcame strains with one spacer through point mutations in the PAM or seed sequence, while escape through deletions and duplications occurred less frequently.

Phages typically escape strains with multiple spacers by deletions

Next, we assessed how phages escaped strains that had acquired multiple spacers. Phages that had escaped targeting of the tail fibre gene had deletions ranging from 990 bp to 2,004 bp, comprising up to 61% of the 3,300 bp gene (figure 3*a* & *b* and figure S2*f* & *h*). Each deletion was towards the centre of the gene, which allowed the gene to remain in-frame. Therefore, the ends of the protein would remain intact, indicating that a shortened tail fibre is tolerated. Indeed, no escape phages formed on a strain that targets 44 bp from the N-terminus (figure 1*c*; TE TFP 5), further supporting the view that deletions close to the start of this gene cannot be tolerated functionally. We also characterised a few rare escape deletions that arose from single phage-targeting spacers, which revealed that similar deletions ranging from 57 to 1,719 bp were possible (figure 3*b*, table S4). However, we saw that escape through point mutation was more common for these strains (figure S3*b*). Analysis of all deletions demonstrated that they were mediated by 4 to 24 bp of sequence microhomology (figure 3*b*, table S4). In phage T4, recombination is more frequent between GC-rich sequences (49). In agreement, most microhomology sequences in ϕTE had a higher GC% than the average GC% for the tail fibre gene. In summary, large deletions enabled escape from strains with three phage targeting spacers. Therefore, multiple spacers typical of priming imposes more stringent selection that can lead to large deletions, which impact on the structural integrity and morphology of these phages.

Deletions in the tape measure protein gene allow phages to overcome CRISPR-Cas targeting

About one-third of the ϕTE genome encodes predicted structural genes (34), including the putative phage tail length determining gene, encoding the tape measure protein. Tape measure proteins dictate the length of the phage tail (50, 51) and altering their length alters tail length proportionally in members of the *Siphoviridae* (32, 50, 52, 53) and *Myoviridae* (51). We predicted that phages that escape CRISPR-Cas targeting, via deletions in the tape measure gene, would have altered morphologies due to shorter tails. Analysis of an anti-ϕTE TMP genotypic escape phage revealed a 120 bp deletion (figure 3*d*; esc ϕTE5). When, this phage was grown on a host with a different spacer targeting elsewhere in the gene, a phage variant with a further 45 bp deletion emerged (esc ϕTE40). In total, ~7% of the tape measure gene was deleted (figure 3*d*). Similar to the tail fibre deletions described earlier, the tape measure deletions were associated with microhomology sequences (table S4). No deletions were isolated on other strains with single or multiple spacers, suggesting that the other regions targeted are less tolerant, or functionally intolerant, of deletions (table S3). In summary, this confirms that CRISPR-Cas can positively select for phages containing deletions in genes encoding predicted structural proteins, including the tape measure protein, which is predicted to generate in phages with altered tail lengths.

CRISPR-Cas escape mutants have altered structural morphology

To determine if deletions in genes encoding structural proteins would lead to altered phage morphologies, we used electron microscopy (EM). Firstly, we used cryo-EM to analyse the morphology of the ϕTE phage in detail. We reconstructed the 3D structure of the capsid to a resolution of 7.1 Å (figure 4*a*, figure S4). A 2D average of tail images shows a helical, 1,460 Å long tail, composed of a long thin neck, a base-plate, and ~26 equally spaced rings with a width of ~40 Å each (figure 4*b*). To examine the effects of the deletions in the phage structural genes (i.e. the tail fibre and tape measure), a selection of escape mutants were

purified and analysed by negative stained EM. Due to difficulties in observing the small tail fibres that are positioned in different arrangements, it was not possible to visualise these structural changes. However, analysis of the escape phage with two deletions in the gene encoding the tape measure protein (figure 4*c* - *e*; esc ϕTE40) showed that the tail was shorter by 75 Å compared with the WT phage. Although the entire wild type tail length is 1,460 Å, the tape measure protein only contributes to the length of the tail ring structure, which is 1,040 Å long (figure 4*c*). In the escape phage, the tail length is reduced by 7%, resulting in the loss of 1-2 rings (figure 4*d* & *e*), correlating with the ~7% gene deletion (figure 3*d*). In summary, CRISPR-Cas interference can provide a strong selective pressure for the generation of phages carrying deleted regions of genes, leading to phage morphological changes.

Deletion escape mutations have a minor impairment in infectivity

Although deletions and point mutations allow phages to overcome CRISPR-Cas, they are predicted to have different fitness outcomes for the phages. Deletions are more likely to have a larger impact on protein function than point mutations. Since tail fibres and the tape measure protein are required for adsorption and DNA injection, phage infection might be affected in the corresponding mutants (53, 54). To assess the effect of the mutations, phage infection was measured in liquid cultures (figure 5). Each escape mutant was used to infect WT *P. atrosepticum* at a range of multiplicities of infection and compared with wild type ϕTE infection (figure 5*a* & *b*). In all cases, the escape phages infected the WT strain with an efficiency that was similar to the wild type ϕTE, with the exception of two deletion mutants (ϕTE42 and ϕTE40). These deletion mutants were less effective at infecting the WT strain at lower multiplicities of infection. When all escape phages were compared with the ability of the wild type ϕTE to infect the original phage resistant strains, the escape mutants were always more successful than the wild type phage (figure 5*c*). These experiments showed that the mutants were not markedly impaired for infection capacity, at least when grown in these single phage-host combinations under laboratory conditions. The point mutant escape phages and the tape measure deletion mutant adsorbed to the WT host similarly to wild type ϕTE, indicating no obvious major impairment in host recognition and binding (figure 5*d*). In contrast, the two mutants with deletions in the tail fibres had essentially undetectable or impaired adsorption over the same timeframe (30 min). However, we did observe an increase in phage titres in these cultures after a single round of infection, by 6- and 4-fold for ϕTE2 and ϕTE42, respectively (figure 5*e*), and since these phages still infect and lyse bacterial cultures (figure 5*b* & *c*), adsorption must still occur, albeit at a lower rate. In summary, escape phages with point mutations or deletions in phage structural genes were still functional and had a wider host range than wild type ϕTE, as they replicated extensively on the strains with spacers that targeted the wild type phage.

Discussion

CRISPR-Cas systems can generate immunity through naïve or primed adaptation. Naïve adaptation usually results in a single new spacer targeting the invading phage, whereas priming frequently leads to acquisition of multiple new spacers. Here, we tested the ability of the type I-F CRISPR-Cas system in *P. atrosepticum* to limit phage infection and assessed the consequences for the phage sequence, structure and function when phages escape immunity provided by one or multiple spacers. Phages overcame strains with one targeting spacer most commonly via point mutations, whereas duplications and deletions were less frequently observed. In contrast, strains with multiple phage-targeting spacers favoured the emergence of escapes that had internal gene deletions. The resulting deletions could occur in genes encoding structural proteins, demonstrating how CRISPR-Cas can influence the generation of phage diversity and morphology. Evidence of randomly occurring deletions and point mutations in phages (e.g. (45, 46, 50)), is consistent with the view that CRISPR-Cas targeting leads to the evolution of pre-existing phage variants arising through mutation and natural selection.

Our results demonstrate that the most frequent route for phages to escape strains with a single phage-targeting spacer is via point mutations, which is consistent with previous studies (7, 10-19). Analysis of the point mutants provided evidence that the PAM and positions 1-5 and 7 of the protospacer are important for efficient targeting by the type I-F system, which builds on previous work in *P. atrosepticum* and *P. aeruginosa* (12, 48). These phages with PAM and protospacer point mutations can initially escape CRISPR-Cas interference and appeared to have no detectable effect on phage infectivity. However, many of these mutants are likely to trigger primed adaptation, which will ultimately result in the acquisition of multiple new spacers targeting those phages (20, 21).

CRISPR-Cas immunity was elevated by multiple spacers, which is likely due to the decreased probability of phage escape and the higher proportion of Cas effector complexes loaded with crRNAs complementary to the invader genome. All escape phages isolated from strains with multiple spacers contained internal deletions, with some up to ~2 kb. Past studies have shown that escape from CRISPR-Cas adaptive immunity can be mediated through deletions up to 471 bp (11, 15, 18). The emergence of ϕTE escape phages with deletions in their tail fibre genes fit with observations seen in members of the *Siphoviridae,* where deletions are mediated by sequences of microhomology and the collagen-like repeats (45, 55, 56). The collagen-like repeats are recombination hotspots due to their repetitive nature and consequently, tail fibre genes are often mosaic, consisting of sequences from different phages (54, 57, 58). High levels of recombination in the cognate tail fibre genes suggests that they are under strong evolutionary pressure due to their role in recognising and binding to host receptors (57, 58), which mutate in response to phage predation (59). In our study, when compared with the tail fibre gene, much smaller deletions were detected in the tape measure protein, indicating a reduced capacity to tolerate deletions. Although deletions have been studied in the gene encoding the tape measure protein of phages λ (50, 52, 53) and T4 (51), long tails are thought to be important for efficient DNA injection (52) and most phage tape measure protein mutants are non-viable (51-53). Since the length of the tape measure protein is proportional to the length of the tail (50, 52, 53), we predicted, and demonstrated, that the mutants selected by CRISPR-Cas had shorter tails. Due to the structural and functional importance of the tail fibres and tape measure, the mutant genes encoding these proteins remained in-frame in the deletion escape mutants. Although these deletions provided a clear benefit on hosts bearing spacers targeting these regions, there were some fitness costs incurred by these deletions in terms of the ability of these phages to bind and infect the wild

type host. Taken together, our results support a model whereby the acquisition of multiple spacers via priming can lead to the selection of evolved phages, carrying deletions that can influence phage morphology and fitness.

We observed that phages were able to replicate, albeit poorly, on most of the anti-ϕ strains without heritably escaping the *P. atrosepticum* type I-F CRISPR-Cas system. The particular phage genes that were targeted influenced whether genotypic escape phages were detected. While large deletions were detected in the ϕTE tail fibre gene, no genotypic escape phages were detected on either of the ϕM1 genes targeted. In contrast, point mutations that led to genotypic escape emerged in the ϕTE tape measure and DNA polymerase genes of the phages when targeted in these regions by single spacers. This suggests that different phage genes are less tolerant to mutations. Differences in mutational tolerance will be influenced by both the essentiality of the genes and the frequency and position of sequences of recombinogenic microhomology that can mediate deletion formation.

Finally, the ability of type I CRISPR-Cas systems to rapidly gain phage resistance through priming (20-22, 28, 60), might mean that escape by point mutation is merely a short-term solution. Indeed, although overcoming single spacers through point mutations may be less detrimental to phage fitness, it is likely to result in enhanced primed CRISPR resistance against the phage. In response to priming, phage deletions appear to be a major route enabling survival when exposed to the challenge of CRISPR-Cas immunity, despite the potential for reduced function of the target gene, imposing a potential fitness cost for the phage. To conclude, single or multiple spacers, common to naïve and primed acquisition, present biologically different selection pressures on phages. The selection of deletion escape phages to overcome primed acquisition demonstrates that CRISPR-Cas can be a powerful driver of genetic and structural diversity in phages.

Acknowledgements

This work was supported by a Rutherford Discovery Fellowship from the Royal Society of New Zealand (RSNZ) (to P.C.F.), the Marsden Fund, RSNZ, the Bio-protection Research Centre (Tertiary Education Commission), a University of Otago Doctoral Scholarship (to B.N.J.W.), University of Otago Division of Health Sciences Career Development Post-doctoral Fellowship and a Veni grant (grant number 016.Veni.171.047) from the the Netherlands Organization for Scientific Research (to R.H.J.S.). GPCS was supported by the BBSRC, UK (awards BB/H002677/1 and BB/G000298/1). We thank members of the Fineran laboratory for useful discussions, Simon Jackson and Hannah Hampton for comments on the manuscript and Vivienne Young for ultracentrifuge training.

References

1. Suttle CA. 2005 Viruses in the sea. *Nature*; 437(7057):356-61.
2. Drake JW, Charlesworth B, Charlesworth D, Crow JF. 1998 Rates of spontaneous mutation. *Genetics*; 148(4):1667-86.
3. Dy RL, Richter C, Salmond GPC, Fineran PC. 2014 Remarkable mechanisms in microbes to resist viral infections. *Annual Review of Virology*; 1(1):307-31.
4. Labrie SJ, Samson JE, Moineau S. 2010 Bacteriophage resistance mechanisms. *Nature reviews Microbiology*; 8(5):317-27.
5. Jackson SA, McKenzie RE, Fagerlund RD, Kieper SN, Fineran PC, Brouns SJ. 2017 CRISPR-Cas: Adapting to change. *Science*; 356(6333):eaal5056.

6. Brouns SJ*, et al.* 2008 Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*; 321(5891):960-4.
7. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. 2007 CRISPR provides acquired resistance against viruses in prokaryotes. *Science*; 315(5819):1709-12.
8. Hille F, Charpentier E. 2016 CRISPR-Cas: biology, mechanisms and relevance. *Philos Trans R Soc Lond B Biol Sci*; 371(1707).
9. Barrangou R, Horvath P. 2017 A decade of discovery: CRISPR functions and applications. *Nat Microbiol*; 2:17092.
10. Semenova EV, Jore MM, Westra ER, Oost vdJ, Brouns SJJ. 2011 Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci U S A*; 108(25):10098-8424.
11. Strotskaya A, Savitskaya E, Metlitskaya A, Morozova N, Datsenko KA, Semenova E, Severinov K. 2017 The action of *Escherichia coli* CRISPR-Cas system on lytic bacteriophages with different lifestyles and development strategies. *Nucleic acids research*; 45(4):1946-57.
12. Cady KC, Bondy-Denomy J, Heussler GE, Davidson AR, O'Toole GA. 2012 The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J Bacteriol*; 194(21):5728-38.
13. Westra ER*, et al.* 2015 Parasite exposure drives selective evolution of constitutive versus inducible defense. *Current biology : CB*; 25(8):1043-9.
14. Tao P, Wu X, Rao V. 2018 Unexpected evolutionary benefit to phages imparted by bacterial CRISPR-Cas9. *Science Advances*; 4(2):eaar4134.
15. Deveau H, Barrangou R, Garneau JE, Labonte J, Fremaux C, Boyaval P, Romero DA, Horvath P, Moineau S. 2008 Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol*; 190(4):1390-400.
16. Levin BR, Moineau S, Bushman M, Barrangou R. 2013 The population and evolutionary dynamics of phage and bacteria with CRISPR-mediated immunity. *PLoS Genet*; 9(3):e1003312.
17. Sun CL, Barrangou R, Thomas BC, Horvath P, Fremaux C, Banfield JF. 2013 Phage mutations in response to CRISPR diversification in a bacterial population. *Environ Microbiol*; 15(2):463-70.
18. Martel B, Moineau S. 2014 CRISPR-Cas: an efficient tool for genome engineering of virulent bacteriophages. *Nucleic acids research*; 42(14):9504-13.
19. Paez-Espino D, Sharon I, Morovic W, Stahl B, Thomas BC, Barrangou R, Banfield JF. 2015 CRISPR immunity drives rapid phage genome evolution in *Streptococcus thermophilus*. *mBio*; 6(2):e00262-15.
20. Datsenko KA, Pougach K, Tikhonov A, Wanner BL, Severinov K, Semenova E. 2012 Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nature communications*; 3:945.
21. Swarts DC, Mosterd C, van Passel MW, Brouns SJ. 2012 CRISPR interference directs strand specific spacer acquisition. *PloS one*; 7(4):e35888.
22. Staals RH, Jackson SA, Biswas A, Brouns SJ, Brown CM, Fineran PC. 2016 Interference-driven spacer acquisition is dominant over naive and primed adaptation in a native CRISPR-Cas system. *Nature communications*; 7:12853.
23. Nicholson TJ, Jackson, S. A., Croft, B. I., Staals, R. H. J., Fineran, P. C., Brown, C. M. 2018 Bioinformatic evidence of widespread priming in Type I and II CRISPR-Cas systems. *submitted*.

24. Przybilski R, Richter C, Gristwood T, Clulow JS, Vercoe RB, Fineran PC. 2011 Csy4 is responsible for CRISPR RNA processing in *Pectobacterium atrosepticum*. *RNA biology*; 8(3):517-28.
25. Richter C, Fineran PC. 2013 The subtype I-F CRISPR-Cas system influences pathogenicity island retention in *Pectobacterium atrosepticum* via crRNA generation and Csy complex formation. *Biochem Soc Trans*; 41(6):1468-74.
26. Bell KS*, et al.* 2004 Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors. *Proc Natl Acad Sci U S A*; 101(30):11105-10.
27. Pawluk A, Staals RHJ, Taylor C, Watson BNJ, Saha S, Fineran PC, Maxwell KL, Davidson AR. 2016 Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nature Microbiology*; 1(8):16085.
28. Richter C, Dy RL, McKenzie RE, Watson BN, Taylor C, Chang JT, McNeil MB, Staals RH, Fineran PC. 2014 Priming in the Type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic acids research*; 42(13):8516-26.
29. Smith MCM, Burns N, Sayers JR, Sorrell JA, Casjens SR, Hendrix RW, Engel J. 1998 Bacteriophage Collagen. *Science*; 279(5358):1831.
30. Ghosh N, McKillop TJ, Jowitt TA, Howard M, Davies H, Holmes DF, Roberts IS, Bella J. 2012 Collagen-like proteins in pathogenic *E. coli* strains. *PloS one*; 7(6):e37872.
31. Mondal SI, Islam MR, Sawaguchi A, Asadulghani M, Ooka T, Gotoh Y, Kasahara Y, Ogura Y, Hayashi T. 2016 Genes essential for the morphogenesis of the Shiga toxin 2-transducing phage from *Escherichia coli* O157:H7. *Sci Rep*; 6:39036.
32. Blower TR*, et al.* 2017 Evolution of *Pectobacterium* bacteriophage phiM1 to escape two bifunctional type III toxin-antitoxin and abortive infection systems through mutations in a single viral gene. *Applied and environmental microbiology*; 83(8):e03229-16.
33. Marchler-Bauer A*, et al.* 2005 CDD: a Conserved Domain Database for protein classification. *Nucleic acids research*; 33(Database issue):D192-6.
34. Blower TR, Evans TJ, Przybilski R, Fineran PC, Salmond GPC. 2012 Viral evasion of a bacterial suicide system by RNA–based molecular mimicry enables infectious altruism. *PLoS Genet*; 8(10):e1003023.
35. Toth IK, Mulholland V, Cooper V, Bentley S, Shih YL, Perombelon MCM, Salmond GPC. 1997 Generalized transduction in the potato blackleg pathogen *Erwinia carotovora* subsp. *atroseptica* by bacteriophage phiM1. *Microbiology*; 143(7):2433.
36. Kremer JR, Mastronarde DN, McIntosh JR. 1996 Computer visualization of three-dimensional image data using IMOD. *J Struct Biol*; 116(1):71-6.
37. Carragher B, Kisseberth N, Kriegman D, Milligan RA, Potter CS, Pulokas J, Reilein A. 2000 Leginon: an automated system for acquisition of images from vitreous ice specimens. *Journal of Structural Biology*; 132(1):33-45.
38. Zheng SQ, Palovcak E, Armache J-P, Verba KA, Cheng Y, Agard DA. 2017 MotionCor2 - anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods*; 14(4):331-2.
39. Mindell JA, Grigorieff N. 2003 Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol*; 142(3):334-47.
40. Grant T, Rohou A, Grigorieff N. 2018 cisTEM, user-friendly software for single-particle image processing. *eLife*; 7:e35383.
41. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004 UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry*; 25(13):1605-12.

42. Landsberger M, Gandon S, Meaden S, Rollie C, Chevallereau A, Chabas H, Buckling A, Westra ER, van Houte S. 2018 Anti-CRISPR Phages Cooperate to Overcome CRISPR-Cas Immunity. *Cell*; 174(4):908-16.
43. Borges AL, Zhang JY, Rollins MF, Osuna BA, Wiedenheft B, Bondy-Denomy J. 2018 Bacteriophage Cooperation Suppresses CRISPR-Cas3 and Cas9 Immunity. *Cell*; 174(4):917-25
44. Hynes AP, Rousseau GM, Lemay ML, Horvath P, Romero DA, Fremaux C, Moineau S. 2017 An anti-CRISPR from a virulent streptococcal phage inhibits *Streptococcus pyogenes* Cas9. *Nat Microbiol*; 2(10):1374-80.
45. Desiere F, Lucchini S, Brüssow H. 1998 Evolution of *Streptococcus thermophilus* bacteriophage genomes by modular exchanges followed by point mutations and small deletions and insertions. *Virology*; 241(2):345-56.
46. Bull JJ, Badgett MR, Rokyta D, Molineux IJ. 2003 Experimental evolution yields hundreds of mutations in a functional viral genome. *Journal of molecular evolution*; 57(3):241-8.
47. Wichman HA, Millstein J, Bull JJ. 2005 Adaptive molecular evolution for 13,000 phage generations: A possible arms race. *Genetics*; 170(1):19-31.
48. Vercoe RB*, et al.* 2013 Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLoS Genet*; 9(4):e1003454.
49. Singer BS, Westlye J. 1988 Deletion formation in bacteriophage T4. *Journal of Molecular Biology*; 202(2):233-43.
50. Katsura I, Hendrix RW. 1984 Length determination in bacteriophage lambda tails. *Cell*; 39(3 Pt 2):691-8.
51. Abuladze NK, Gingery M, Tsai J, Eiserling FA. 1994 Tail Length Determination in Bacteriophage T4. *Virology*; 199(2):301-10.
52. Katsura I. 1987 Determination of bacteriophage lambda tail length by a protein ruler. *Nature*; 327(6117):73-5.
53. Mahony J, Alqarni M, Stockdale S, Spinelli S, Feyereisen M, Cambillau C, Sinderen DV. 2016 Functional and structural dissection of the tape measure protein of lactococcal phage TP901-1. *Sci Rep*; 6:36667.
54. Haggard-Ljungquist E, Halling C, Calendar R. 1992 DNA sequences of the tail fiber genes of bacteriophage P2: evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages. *J Bacteriol*; 174(5):1462-77.
55. Duplessis M, Moineau S. 2001 Identification of a genetic determinant responsible for host specificity in *Streptococcus thermophilus* bacteriophages. *Mol Microbiol*; 41(2):325-36.
56. Boyce JD, Davidson BE, Hillier AJ. 1995 Spontaneous deletion mutants of the *Lactococcus lactis* temperate bacteriophage BK5-T and localization of the BK5-T attP site. *Applied and environmental microbiology*; 61(11):4105-9.
57. Tetart F, Desplats C, Krisch HM. 1998 Genome plasticity in the distal tail fiber locus of the T-even bacteriophage: recombination between conserved motifs swaps adhesin specificity. *J Mol Biol*; 282(3):543-56.
58. Hendrix RW. 2002 Bacteriophages: evolution of the majority. *Theoretical Population Biology*; 61(4):471-80.
59. Clement JM, Lepouce E, Marchal C, Hofnung M. 1983 Genetic study of a membrane protein: DNA sequence alterations due to 17 *lamB* point mutations affecting adsorption of phage lambda. *The EMBO journal*; 2(1):77-80.
60. Fineran PC, Gerritzen MJH, Suárez-Diez M, Künne T, Boekhorst J, van Hijum SAFT, Staals RHJ, Brouns SJJ. 2014 Degenerate target sites mediate rapid primed CRISPR adaptation. *Proc Natl Acad Sci U S A*; 111(16):E1629.

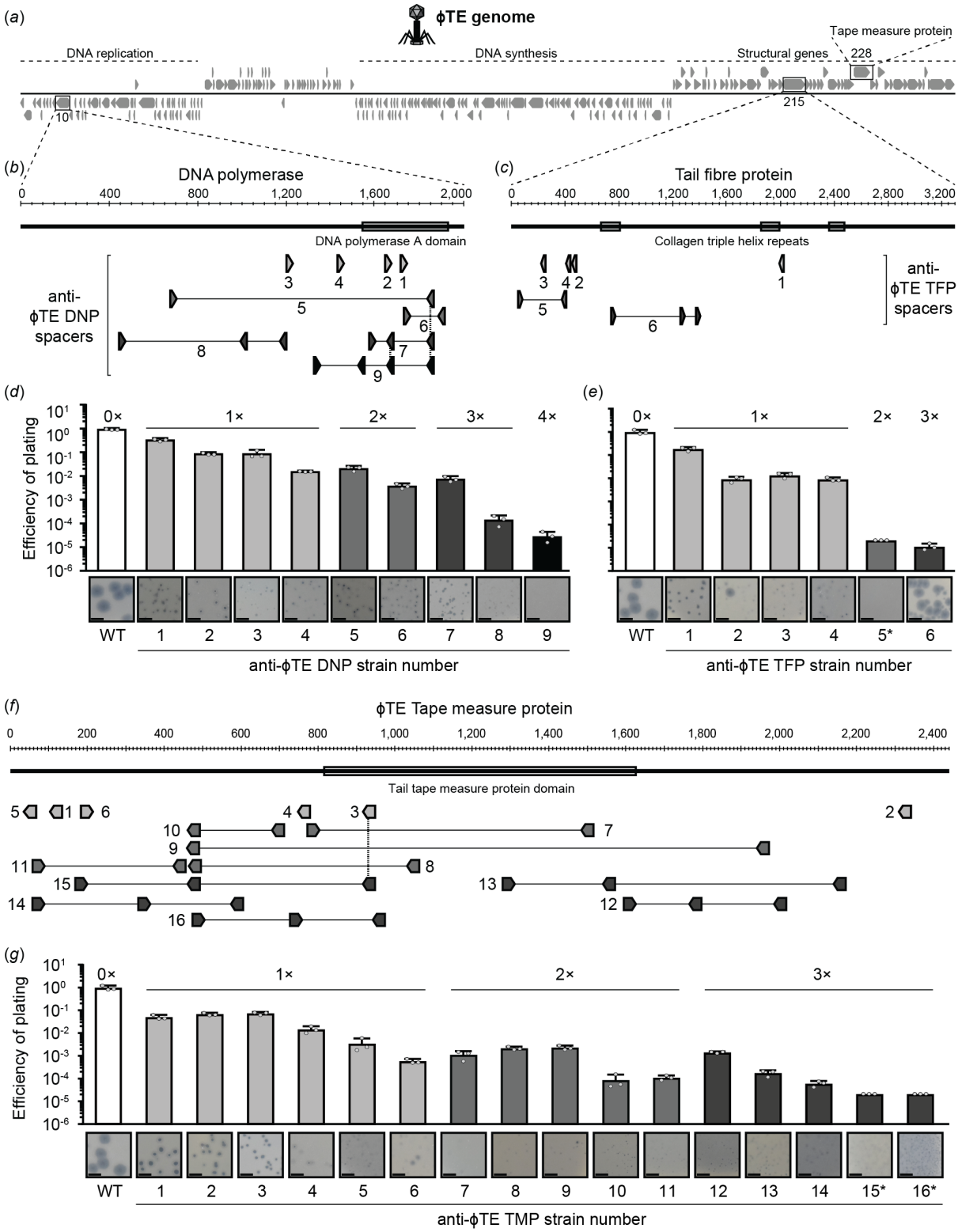
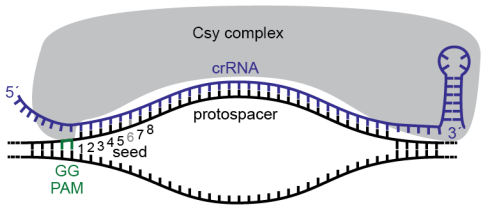


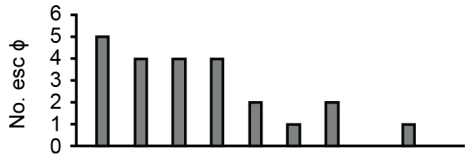
Figure 1. Phage-targeting spacers reduce phage infectivity. (*a*) Schematic of the ϕTE genome showing the locations of (*b*) the DNA polymerase gene (DNP, phiTE_10), (*c*) a tail fibre gene (TFP, phiTE_215) and (*f*) the tape measure protein gene (TMP, phiTE_228) with the targeting spacers. Efficiency of plating (EOP) and plaque morphologies of the (*d*) anti-ϕTE DNP, (*e*) anti-ϕTE TFP and (*g*) anti-ϕTE TMP strains, with white bars: WT (0 spacers), light grey: 1×, mid grey: 2×, dark grey: 3× and black: 4×anti-ϕ spacers. Data shown is the mean +SD (*n*=3). The limit of detection for the assays was 2×10^{-11}. The plaque image scale bars represent 5 mm. In (*b*) and (*f*), the dashed vertical lines represent spacers that are

shared between strains. In (*e*) and (*g*), the EOP values for TE TFP 5 and TE TMP15 and 16 (indicated by an asterisk) were determined based on estimated plaque counts (see Materials and Methods). The plaques formed on these, and some other strains were too small to be captured by camera. Full details of EOP and plaque size measurements are in table S3 and equivalent data for ϕM1 is provided in figure S1.

(*a*)



(*b*)



(*c*)

Strain	esc ϕ	PAM		protospacer							
		-2	-1	1	2	3	4	5	6	7	8
TFP 4	TE25	**C**/G	G	T	T	G	A	C	C	T	T
TMP 4	TE39	**A**/G	G	A	A	C	C	T	G	C	T
TFP 3	TE20	**A**/G	G	T	C	A	C	T	C	G	T
TMP 4	TE22	**T**/G	G	A	A	C	C	T	G	C	T
TFP 2	TE29	**T**/G	G	A	C	T	G	T	G	T	C
TMP 2	TE7	G	**C**/G	C	G	T	C	T	T	T	C
TFP 2	TE31	G	**C**/G	A	C	T	G	T	G	T	C
TMP 1	TE8	G	**A**/G	C	A	G	T	C	C	A	A
TMP 5	TE41	G	**A**/G	C	C	T	T	G	G	T	G
TMP 3	TE10	G	G	**C**/T	T	G	A	C	G	A	C
TMP 6	TE16	G	G	**C**/T	T	T	G	C	G	A	A
TFP 3	TE34	G	G	**C**/T	C	A	C	T	C	G	T
TFP 3	TE33	G	G	**G**/T	C	A	C	T	C	G	T
TMP 1	TE6	G	G	C	**C**/A	G	T	C	C	A	A
TMP 2	TE13	G	G	C	**C**/G	T	C	T	T	T	C
TMP 3	TE21	G	G	T	**C**/T	G	A	C	G	A	C
TMP 1	TE9	G	G	C	**G**/A	G	T	C	C	A	A
TFP 2	TE3	G	G	A	C	**C**/T	G	T	G	T	C
TFP 4	TE27	G	G	T	T	**A**/G	A	C	C	T	T
TFP 4	TE26	G	G	T	T	G	**G**/A	C	C	T	T
TFP 2	TE12	G	G	A	C	T	G	**C**/T	G	T	C
TMP 6	TE11	G	G	T	T	T	G	**A**/C	G	A	A
TFP 2	TE4	G	G	A	C	T	G	T	G	**C**/T	C

Figure 2. Phage escape from strains with one spacer through point mutations. (*a*) Schematic of protospacer binding by the Csy complex of the type I-F CRISPR-Cas system, highlighting the position of the PAM sequence (green, top strand) and the first eight bases of the protospacer sequence. (*b*) The frequency of point mutation at each seed and PAM position and (*c*) protospacer seed and PAM sequences of point mutant escape phages, showing the point mutation changes (red) and the original nucleotide (/X). (*a*) & (*c*) position 6 of the protospacer is faded because this position does not influence target binding (48). Effects of the point mutations on the amino acid sequences of the phage proteins are provided in figure S3*a* and table S4.

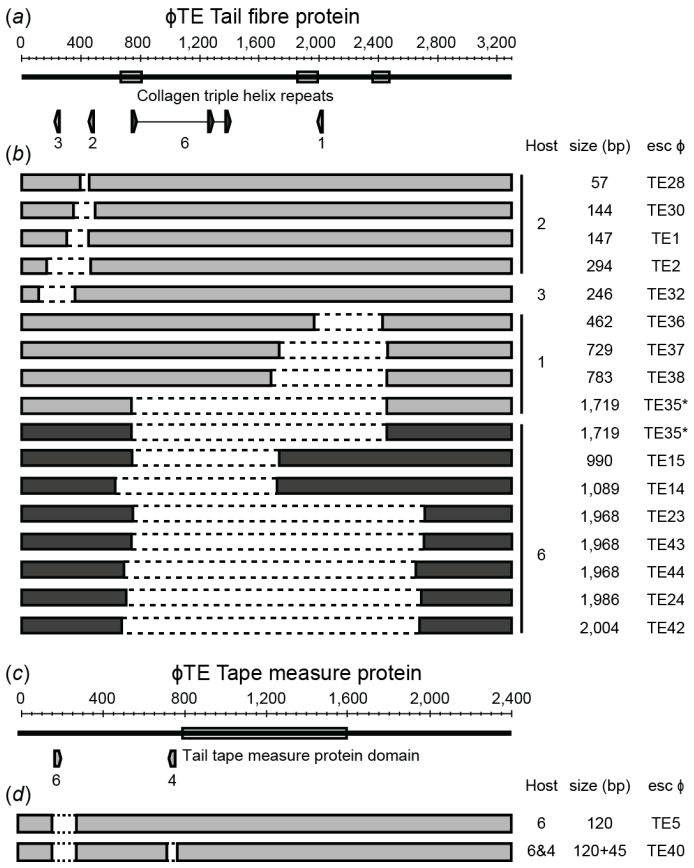


Figure 3. Phages escape strains with multiple spacers by deletions. (*a*) Schematic of the ϕTE tail fibre and (*c*) tape measure protein genes showing the regions targeted by different anti-ϕTE strains. (*b*) and (*d*) Diagrams of the deletion escape phages, showing the deleted regions (dashed bars), from 1×anti-ϕTE strains (light grey bars) and the 3×anti-ϕTE TFP strain (dark grey bars). The host strain (TE TFP/ TMP#), deletion size and escape phage number are shown for each phage mutant. Phage ϕTE35 (*) was isolated on two strains, TE TFP1 and 6. For further details see table S4.

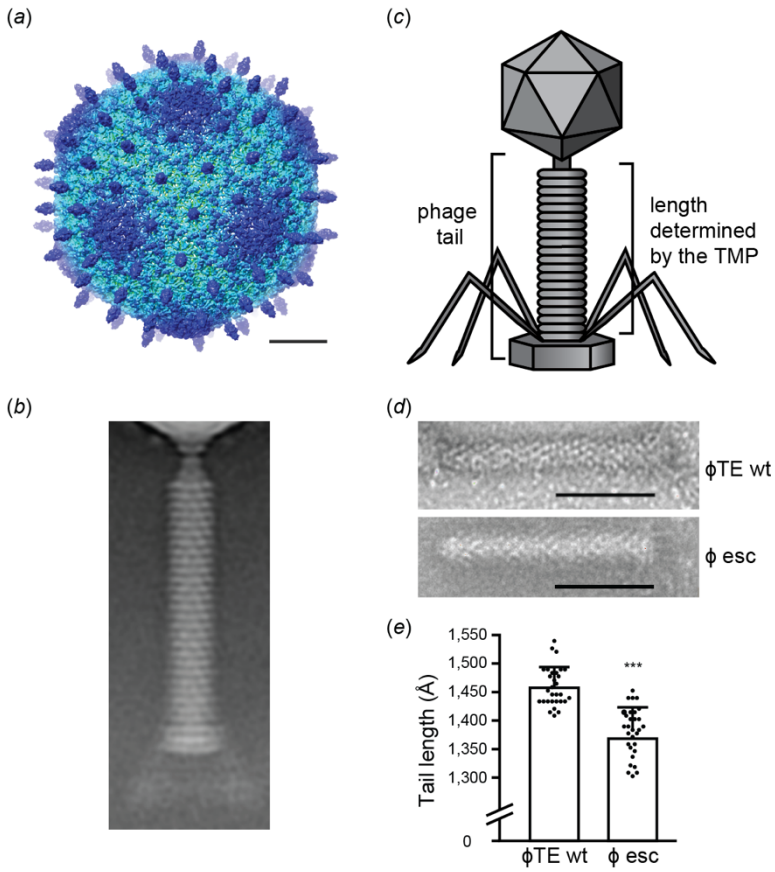


Figure 4. CRISPR-Cas escape mutants have altered structural morphology. (*a*) Radially coloured surface representation of the wild type ϕTE capsid; hexamers of the major capsid protein are visible with a protruding decoration protein at their centre. The scale bar is 200 Å. (*b*) A two-dimensional average of single images of the wild type ϕTE tail. (*c*) Schematic of ϕTE showing the phage tail, and the region of the tail of which the length is determined by the tape measure protein. (*d*) Negative stain images of the wild type ϕTE tail (top) and the tape measure gene escape phage (ϕTE40) tail (bottom) with (*e*) the analysis of the tail lengths (scale bar 600 Å). Data shown is the mean +SD (*n*=30). Statistical significance was calculated using an unpaired t test (***, $P \leq 0.0001$). For more data see figure S4.

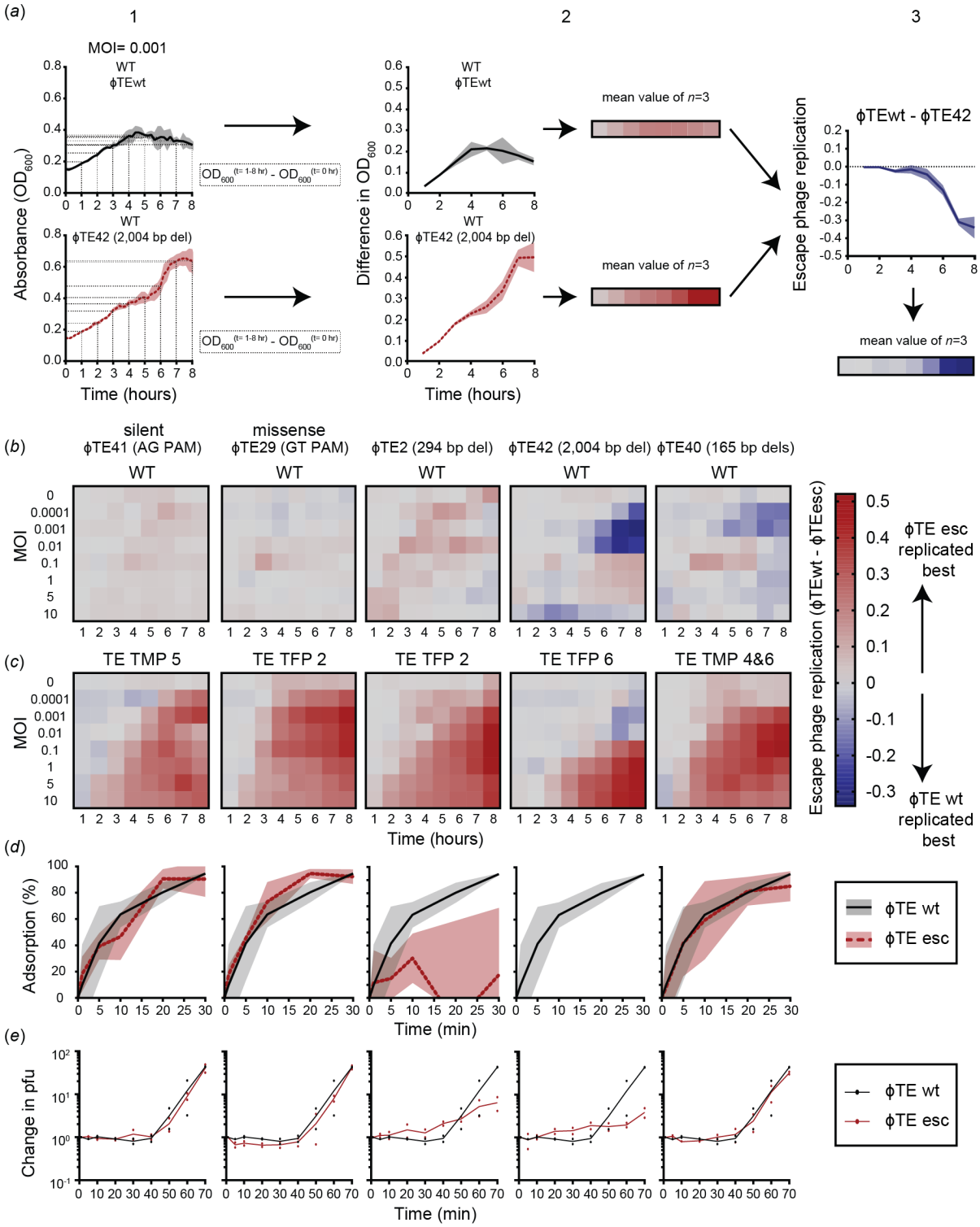


Figure 5. Deletion escape mutations cause a minor impairment in phage infectivity. (*a*) Example of how the relative phage escape growth was calculated. 1) the growth curves for WT infected with wild type ϕTE and ϕTE42 at an MOI of 0.001. 2) the difference in OD_{600} values for the first eight hours of growth, compared to the initial value, was calculated. 3) the values for ϕTE42 were subtracted from the wild type ϕTE values to determine the relative escape phage growth. (*b*) & (*c*) WT and anti-ϕ strains were grown with wild type ϕTE and escape phages at different multiplicities of infection (MOIs) and the difference in absorbance (OD_{600}) was calculated. Escape phage replication was determined as (the difference in OD for wild type ϕTE infected cultures) - (the difference in OD for ϕ esc infected cultures), for both (*b*) WT and (*c*) the anti-ϕ hosts. (*d*) Phage adsorption was determined for wild type ϕTE

(black lines) and each escape phage (red dashed lines) at 0, 1, 5, 10, 20 and 30 min. (*e*) The fold change in plaque forming units (pfu) was also determined for each phage over time. Data shown is the mean ±SD (represented by the shading) (*n*=3). For (*e*) (*n*=2).