OPEN

The ISME Journal (2017), 1–16
www.nature.com/ismej

ORIGINAL ARTICLE

Genetic hurdles limit the arms race between *Prochlorococcus* and the T7-like podoviruses infecting them

Daniel A Schwartz and Debbie Lindell
Faculty of Biology, Technion—Israel Institute of Technology, Haifa, Israel

Phages and hosts coexist in nature with a high degree of population diversity. This is often explained through coevolutionary models, such as the arms race or density-dependent fluctuating selection, which differ in assumptions regarding the emergence of phage mutants that overcome host resistance. Previously, resistance in the abundant marine cyanobacterium, *Prochlorococcus*, was found to occur frequently. However, little is known about the ability of phages to overcome this resistance. Here we report that, in some cases, T7-like cyanophage mutants emerge to infect resistant *Prochlorococcus* strains. These resistance-breaking phages retained the ability to infect the wild-type host. However, fitness of the mutant phages differed on the two hosts. Furthermore, in one case, resistance-breaking was accompanied by costs of decreased fitness on the wild-type host and decreased adsorption specificity, relative to the wild-type phage. In two other cases, fitness on the wild-type host increased. Whole-genome sequencing revealed mutations in probable tail-related genes. These were highly diverse in isolates and natural populations of T7-like cyanophages, suggesting that antagonistic coevolution enhances phage genome diversity. Intriguingly, most interactions did not yield resistance-breaking phages. Thus, resistance mutations raise genetic barriers to continuous arms race cycles and are indicative of an inherent asymmetry in coevolutionary capacity, with hosts having the advantage. Nevertheless, phages coexist with hosts, which we propose relies on combined, parallel action of a limited arms race, fluctuating selection and passive host-switching within diverse communities. Together, these processes generate a constantly changing network of interactions, enabling stable coexistence between hosts and phages in nature.

The ISME Journal advance online publication, 25 April 2017; doi:10.1038/ismej.2017.47

Introduction

Lytic bacteriophages are obligate parasites, thus their survival depends on the ability to encounter and reproduce inside a suitable host. In planktonic communities, the likelihood of finding such a host is a function of the concentration of both host and phage. However, as lytic phages kill their hosts in the course of reproduction, they reduce the density of susceptible hosts and select for variants in the host population that are resistant to infection (Bohannan and Lenski, 2000; Buckling and Rainey, 2002). Phage mutants capable of infecting resistant hosts often emerge (Lenski and Levin, 1985; Qimron *et al.*, 2006; Poullain *et al.*, 2008; Meyer *et al.*, 2010; Paterson *et al.*, 2010). Such phage–host coevolution is now realized to be a central process that generates and maintains extensive diversity among microbes (Rodriguez-Valera *et al.*, 2009; Avrani *et al.*, 2011; Cordero and Polz, 2014). However, much less attention is given to the mechanisms underlying the generation and preservation of viral diversity (Paterson *et al.*, 2010; Mizuno *et al.*, 2014).

Antagonistic host–phage coexistence in nature is often explained by reciprocal selection, whereby phages select for resistant hosts and resistant hosts select for phages capable of re-infecting resistant hosts (Avrani *et al.*, 2012; Martiny *et al.*, 2014). In experimental systems, such mutation and counter-mutation cycles that lead to an arms race result in selective sweeps in which host and phage mutants with increasing resistance and infectivity ranges replace their predecessors (Buckling and Rainey, 2002; Mizoguchi *et al.*, 2003; Marston *et al.*, 2012). However, resistance that cannot be overcome by phage mutants is sometimes observed (Lenski and Levin, 1985; Duffy *et al.*, 2006; Lennon and Martiny, 2008). In addition, ever increasing fitness costs associated with these mutations are expected to eventually limit the propagation of the arms race (Gomez and Buckling, 2011; Hall *et al.*, 2011).

In an alternative ecological model of fluctuating selection that takes fitness costs into consideration, such as that proposed in the ‘kill-the-winner’ model,

---

Correspondence: D Lindell, Faculty of Biology, Technion—Israel Institute of Technology, Emerson Building, room 5-31, Technion City, Haifa 32000, Israel.
E-mail: dlindell@tx.technion.ac.il

Received 18 August 2016; revised 31 January 2017; accepted 28 February 2017

antagonistic coexistence is explained by density-dependent fluctuations of susceptible and resistant host types (Woolhouse *et al.*, 2002; Winter *et al.*, 2010; Avrani *et al.*, 2012). These fluctuations result from alternating selective forces, whereby susceptible hosts and phages fluctuate in predator–prey-like cycles: phages infecting susceptible hosts increase in abundance, selecting for resistant hosts; consequently, phage numbers decrease as they cannot infect the resistant hosts, relieving phage predation pressure; thereupon, resistant hosts are outcompeted by susceptible hosts, as the latter are not burdened by pleiotropic costs of resistance (Gomez and Buckling, 2011; Hall *et al.*, 2011). This model does not require the emergence of resistance-breaking mutants (RBMs) but does require pleiotropic costs of resistance. Note that we refer to mutant phages capable of re-infecting resistant hosts as 'resistance-breaking mutants', following terminology used by the plant virus community (Harrison, 2002; Palloix *et al.*, 2009), rather than 'host range mutants', as these phages infect an allelic variant of the same host rather than a truly new host.

Bacteria–phage coevolution studies abound (reviewed in Buckling and Brockhurst (2012); Koskella and Brockhurst (2014) and Martiny *et al.* (2014)) with many studies focused on the host, both on the mutations that confer resistance as well as their fitness effects (Lenski, 1988a; Lennon *et al.*, 2007; Forde *et al.*, 2008; Avrani *et al.*, 2011). Clearly, the diverse mechanisms of phage resistance in bacteria (Hyman and Abedon, 2010; Labrie *et al.*, 2010) have dictated diverse mechanisms allowing phages to overcome resistance (Samson *et al.*, 2013). However, one of the most common modes of resistance is by mutations in cell surface structures that prevent adsorption of phage to the cell surface of the host (Bohannan and Lenski, 2000). In these cases, resistance-breaking phage mutants arise by changes in structural genes mediating host recognition and attachment. In tailed phages such mutations typically modify the tail and the tail fiber (Ravin *et al.*, 2002; Qimron *et al.*, 2006; Paterson *et al.*, 2010; Scanlan *et al.*, 2011; Meyer *et al.*, 2012; Perry *et al.*, 2015), although mutations in other genes have also been found (Paterson *et al.*, 2010; Marston *et al.*, 2012). Typically, RBMs retain their ability to infect the original susceptible host, and in some cases result in a reduction in phage fitness (Chao *et al.*, 1977; Duffy *et al.*, 2006; Ford *et al.*, 2014).

An example of coexistence in the environment is that of unicellular marine picocyanobacterium, *Prochlorococcus* and the phages infecting it. *Prochlorococcus*, found in warm sunlit waters of the ocean in concentrations often exceeding 10^5 cells per ml (Johnson *et al.*, 2006; Zwirglmaier *et al.*, 2008), are the world's most numerous photosynthetic organism (Partensky *et al.*, 1999; Flombaum *et al.*, 2013). Coexisting with *Prochlorococcus* are phages (Sullivan *et al.*, 2003; Dekel-Bird *et al.*, 2015) that are thought to influence their hosts' abundance, diversity, metabolism and evolution (Mann, 2003; Lindell *et al.*, 2004; Muhling *et al.*, 2005; Avrani *et al.*, 2011; Thompson *et al.*, 2011; Parsons *et al.*, 2012). Cyanophages isolated to date have double-stranded DNA genomes and belong to the three morphologically defined families of the tailed bacteriophages, the *Myoviridae*, *Podoviridae* and *Siphoviridae* (Suttle and Chan, 1993; Waterbury and Valois, 1993; Mann, 2003; Sullivan *et al.*, 2003; Wang and Chen, 2008; Sabehi *et al.*, 2012). Of relevance to this work are the cyanopodophages that resemble the enteric T7 phage.

The similarity of T7-like cyanophages to the T7 archetype is found for a number of different characteristics such as: virion structure (Liu *et al.*, 2010); genome organization, including the presence of many homologous genes, their conserved arrangement along the genome and their grouping according to functional modules (Chen and Lu, 2002; Sullivan *et al.*, 2005; Pope *et al.*, 2007; Labrie *et al.*, 2013); and their expression program (Lindell *et al.*, 2007). T7-like cyanophages are typically very host-specific, infecting only one or two closely related hosts (Sullivan *et al.*, 2003; Dekel-Bird *et al.*, 2013; Dekel-Bird *et al.*, 2015). However, it remains unclear how such narrow host range phages survive when resistance dominates.

Prochlorococcus resistance to T7-like cyanophages occurs at a high frequency and is often through the prevention of phage attachment to the cell surface (Avrani *et al.*, 2011). Many of these phage-resistant strains have fitness costs of lower growth rate or enhanced susceptibility to other phages. Most phage resistance mutations in *Prochlorococcus* are in genes of unknown function, and occur in genomic islands (Avrani *et al.*, 2011), hypervariable genomic regions with different gene content among closely related isolates (Coleman *et al.*, 2006). From the phage perspective, the abundant *Prochlorococcus* population is composed of numerous subpopulations with different phage susceptibilities, with only a fraction of these being available to any given phage type (Avrani *et al.*, 2011). A high degree of resistance would lead to low encounter rates between susceptible hosts and co-occurring phages, and was previously suggested as a mechanism facilitating cyanobacterial–cyanophage coexistence (Waterbury and Valois, 1993; Avrani *et al.*, 2011).

That evolution of phage resistance drives host diversity in genomic susceptibility regions is now well described (Rodriguez-Valera *et al.*, 2009; Avrani *et al.*, 2011; Kashtan *et al.*, 2014); however, its consequence for phage diversity and evolution remains unclear. In an experimental coevolution study in a related system, several arms race cycles were observed between a marine *Synechococcus* host and a T4-like cyanophage (Marston *et al.*, 2012). Phage mutations that likely evolved in response to host resistance occurred in genes of unknown function and their effect on phage fitness remains unknown.

Here we set out to investigate whether coexistence between the highly abundant *Prochlorococus* and T7-like cyanophages is facilitated by an arms race. To this end, we tested the ability of T7-like cyanophages to overcome resistance in *Prochlorococcus* and characterized RBMs for host range, fitness costs and genetic diversity. In agreement with the arms race model, some RBMs were isolated that had an expanded host range and no fitness cost. Furthermore, the mutations identified were in tail-related genes that are highly diverse among isolates and natural populations of T7-like cyanophages. However, RBMs were not detected in most interactions investigated. This dearth of mutants indicates that ecological coexistence between these antagonists is unlikely to be solely due to continuous arms race cycles. Rather we suggest a model whereby coexistence and genetic diversity of hosts and phages is due to the combination of an arms race, cost-driven density-dependent fluctuations and passive host-switching.

Materials and methods

Cyanobacterial strains and growth conditions

Prochlorococcus MED4 (henceforth MED4-wild type (WT)) and phage-resistant strains derived from *Prochlorococcus* MED4 (MED4-R#) were the main cyanobacterial strains used in this study. The phage-resistant strains (MED4-R1, MED4-R5, MED4-R6 and MED4-R8) were previously isolated by Avrani *et al.* (2011) as colonies that grew on plates in the presence of a selecting phage (Supplementary Table S1). These strains were chosen because they have been genotypically and phenotypically characterized (Supplementary Table S1; Avrani *et al.*, 2011) and had not lost their resistance due to continued evolution (Avrani and Lindell, 2015) by the onset of this study. However, the MED4-R8 strain underwent changes in its resistance range during the period of this study (Avrani and Lindell, 2015) and was therefore used only for screening of resistance-breaking phages. Several other strains of *Prochlorococcus* and *Synechococcus* were used for phage–host range characterization and as controls for adsorption assays (see below). The doubling time of the cyanobacterial strains used in this study was between 1 and 2 days.

Prochlorococcus was grown in PRO99 medium (Moore *et al.*, 2007) with a Mediterranean seawater base. Cultures were grown at 22 °C under cool white light at an intensity of 10–20 μmol photons m^{-2} s^{-1}, with a 14:10 h light:dark cycle. *Synechococcus* was grown under the same conditions in the ASW artifical seawater medium (Lindell *et al.*, 1998). Growth in liquid cultures was monitored by chlorophyll *a* (chl *a*) fluorescence measured using a Synergy MX microplate reader (BioTek, Winooski, CA, USA), with excitation and emission wavelengths at 440 ± 20 and 680 ± 20 nm, respectively. Relative cell concentrations were determined from chl *a* fluorescence, and verified to have the same number of cells in the different strains by colony plating.

Prochlorococcus colonies and lawns were grown in pour plates of the PRO99 medium supplemented with 0.28% low-melting-point agarose (UltraPure, Invitrogen, Carlsbad, CA, USA) and 1 mM Na_2SO_3 (Lindell, 2014). To achieve high plating efficiencies, a heterotrophic 'helper' strain, *Alteromonas sp.* EZ55, was added to the medium (Morris *et al.*, 2008).

Cyanophage strains, their propagation and enumeration

Three T7-like cyanophage strains that infect *Prochlorococcus* MED4 were used in this study: P-SSP7, P-GSP1 and P-TIP38. Their ability to infect the resistant *Prochlorococcus* strains is shown in Supplementary Table S1. Phages were propagated from single plaques in liquid cultures of exponentially growing host cells. Cell debris were removed from viral lysates by centrifugation (10 000 *g* at 20 °C for 15 min) and filtration (0.22 μm Millex GV syringe filter, Millipore (Cork, Ireland); or 0.2 μm Rapidflow filter unit, Nalgene, Rochester, NY, USA). Phages were enumerated by plaque assays in which serially diluted phages were plated in pour plates after a 1–4 h adsorption step to host cells in liquid under host growth conditions.

Screening for RBM phages

We screened for RBMs of T7-like cyanophages on phage-resistant strains of *Prochlorococcus* MED4. Two approaches were employed that differ in the number of rounds of evolutionary selection allowed. The first method, plaque formation, allowed for a single round of selection. Six interactions were tested based on the resistance range of the MED4-derived resistant strains (Supplementary Table S1): the P-SSP7 phage was tested on the MED4-R1, MED4-R5, MED4-R6 and MED4-R8 strains and the P-GSP1 and P-TIP38 phages were tested on the MED4-R8 strain. Large volume lysates were produced from single plaques on the WT MED4 host. Lysates were concentrated using a centrifugal filtration at 4 °C (Amicon Ultracel 100 KDa, Millipore) and plated on lawns of phage-resistant host strains such that only resistant-breaking mutants would form plaques. Three to four assays with independent phage lysates were carried out for each host–phage pair.

The second method, host-mixing, allowed for multiple evolutionary steps. Phages were added to mixed host populations in liquid consisting of a minority of susceptible hosts and a majority of phage-resistant cells (Benmayor *et al.*, 2009). Phages were serially transferred weekly for 10 rounds: 1% of a 0.2 μm filtrate was transferred to naive, mixed host populations each round. Spontaneous mutations that arise during replication of the WT phage on the susceptible host and that enable infection of resistant cells are thus highly favored and lead to a decline in the resistant host population. Four mixed cell compositions were used with susceptible hosts added at frequencies of 0.01%, 0.1%, 1% and 10%, and the remainder made up of resistant cells. A total of 8–9 × 10^7 cells (colony-forming units) were used each round, and the first round of infection contained 5.7 × 10^6 plaque-forming units of the WT P-SSP7 phage (multiplicity of infection = 0.07). Cultures of susceptible or resistant hosts only, as well as non-infected cultures, were used as controls for phage infection and cell growth, respectively. Four replicate phage populations were evolved on each of the mixed cell compositions and on the controls for each host–phage interaction.

Isolation of RBM phages

RBM phages of P-SSP7 were isolated from assays employing both single and multiple rounds of selection. Well-isolated plaques from the single-round assays were plaque-purified a second time and then propagated in liquid on the same ‘resistant’ host. Six RBMs were isolated in this manner: RBM1–RBM6 (Supplementary Table S1). An additional six RBMs were isolated from the host-mixing assays that allowed for multiple rounds of selection (RBM7–RBM12): mutants from round 2 did not form plaques and were isolated using a liquid dilution to extinction assay, whereas mutants from rounds 4 and 6 formed plaques and were plaque-purified before propagation in liquid. WT control phages (WT1–WT3) were isolated from serial dilutions of each of the initial lysates using plaque assays on the WT MED4 host and served as paired controls for the phenotypic and genotypic characterization of the mutant phages (Supplementary Table S1).

Phenotypic characterization of isolated resistance-breaking phage mutants

Phenotypic characterization was carried out for the first three independently isolated mutants (RBM1, RBM2 and RBM3) that also underwent whole-genome sequencing (see below).

Host range. Phage–host range was tested on *Prochlorococcus* MED4-WT and three of the four resistant MED4 strains as well as on 10 other *Prochlorococcus* and *Synechococcus* strains (see Supplementary Table S1 for a list of strains used). MED4-R8 was not used due to continued evolution that led to changes in its susceptibility after RBM screenings were completed (Avrani and Lindell, 2015). Host range was tested from growth curves of liquid cultures in microtiter plates. Cyanobacterial lysis was monitored by a decline in chl *a* fluorescence relative to uninfected control cultures on the MED4-WT, MED4-R1, MED4-R5 and MED4-R6 strains. Spot assays were used to determine host range on the other cyanobacterial strains: Five microliter of phage lysates was spotted onto 2–3-day-old lawns of cyanobacteria and monitored for appearance of plaque-like clearings. Host ranges were determined using four to six replicate infections for liquid cultures and three replicates for spot assays.

Phage fitness. The production of phage progeny over multiple rounds of infection was used to determine relative phage fitness. Exponentially growing *Prochlorococcus* cultures were diluted to an equal concentration (~2 × 10^8 cells per ml) and were infected at an multiplicity of infection of 0.0001–0.001. These low phage:host ratios were used to ensure that host availability did not limit viral replication over the 2–3-day period of the experiment (Supplementary Figure S1). This long period was chosen as cell doubling and phage replication times are on the order of days and hours, respectively. The number of free phages (0.22 μm filtrate after 100-fold dilution) was measured by plaque assays on lawns of MED4-WT at the onset (N_0) and at the end of the assay (N_t) at $t = 2$ or 3 days (Supplementary Figure S1). Fitness is expressed as the number of phage population doublings per day as $(\log_2(N_t/N_0))/t$ (Bull *et al.*, 1997), where t is the assay time in days. When a mix of phages (both WT and RBM) was tested in direct competition in the same vessel, the abundance of each phage type was determined from the ability of 100 picked plaques to lyse the WT and resistant strains in replicate liquid assays. A plaque that lysed both hosts was determined to be a RBM phage, whereas those that lysed only the WT host were determined to be WT phages. Six to thirteen replicate infections were carried out for fitness measurements of each phage–host interaction.

Phage adsorption. Single-point adsorption assays were performed in which percent adsorption was determined from the number of unattached phage at 4 h relative to immediately after phage addition, as previously described (Avrani *et al.*, 2011). Briefly, exponentially growing cells were concentrated by centrifugation (12 000 *g* at 21 °C for 10 min) to ~ 2 × 10^9 colony-forming units per ml and mixed with phages at a low multiplicity of infection of 0.003–0.014. The free phage titer in the 0.22 μm filtrate (filtered after 100-fold dilution) was determined by plaque assay immediately after phage addition (0 h) and after 4 h. Five to eight replicates were carried out to measure adsorption for each phage–host interaction.

Genetic characterization of RBMs

Mutations in the RBMs of P-SSP7 were determined from whole-genome sequencing of three RBMs (RBM1, RBM2 and RBM3) and were compared to the sequenced genomes of three paired-control, WT phages. Phage genomic DNA was purified from concentrated lysates using the Lambda mini kit (Qiagen, Hilden, Germany) and sheared (Covaris E200, Woburn, MA, USA). DNA-sequencing libraries were constructed using TruSEQ DNA preparation kit (Illumina, San Diego, CA, USA). Paired-end sequencing (2×250 bp) was carried out on a MiSEQ sequencer (Illumina) at the Technion Genome Center (Haifa, Israel). Sequencing reads were analyzed on the GALAXY platform (Goecks *et al.*, 2010). Quality analysis was done using FastQC. Read ends were trimmed uniformly leaving only positions in which 75% of the reads or more had a PHRED quality score of >20. PCR duplicates were removed by rmdup using SAMtools (V.0.1.18) (Li *et al.*, 2009). Paired-end mapping against the P-SSP7 reference genome (GenBank accession: NC_006882) was carried out using the BWA package for Illumina (V.0.6.1; default settings except max. edit distance = 3). Variants were identified by mPileup and were inspected manually using the Integrative Genomics Viewer (V. 2.2.5; Robinson *et al.*, 2011). Mapped reads are available in the Sequence Read Archive of the National Center for Biotechnology Information under Bioproject PRJNA352672. Mutations found in the genomes of RBMs different to those in the paired controls were identified as resistance-breaking-associated mutations (Supplementary Table S2). Verification of the mutations, as well as genotyping of other phage mutants (RBM4–RBM12), was done for genes identified by whole-genome sequencing. Four independent PCRs were carried out for each gene in each phage mutant (Supplementary Table S3), pooled and Sanger sequenced.

To search for homologs of genes of interest the conserved domain and the non-redundant databases of the National Center for Biotechnology Information were searched using BLAST, PSI-BLAST (Altschul *et al.*, 1997) and HMMER (Finn *et al.*, 2011). In addition, the Pfam database was searched using HHpred (Söding, 2005) with a query of protein sequences aligned by MUSCLE (Edgar, 2004).

Diversity of resistance-breaking genes

Core genes in genomes of isolates. To compare the diversity among T7-like core genes, we calculated the average information content of each gene by applying Shannon's diversity as described previously (Avrani *et al.*, 2011). Briefly, amino-acid sequences of the homologs were globally aligned with minimum gaps (MAAFT –legacygappenalty –localpair –maxiterate 1000; Katoh *et al.*, 2002). Position-specific information content, $-p(a)\log_2(p(a))$, where $p(a)$ is the proportion of sequences containing amino acid a, was averaged across all alignment positions having at least five non-gap characters. The information content calculation was implemented using a custom R script (Supplementary File 1). Core genes were compared between eight genomes (P-SSP7, P-SSP10, P-GSP1, P-SSP9, P-HP1, P-SSP2, P-RSP5 and Syn5) spanning the phylogenetic diversity of T7-like cyanophages (Labrie *et al.*, 2013).

Environmental diversity. Metagenomic recruitment was used to identify natural populations of T7-like cyanophage isolates that are present in publically available marine viral metagenomes. A phage was defined to be present in the population when reads recruited to the reference genome at $>90\%$ BLASTn identity, had a median coverage >5 and were evenly distributed across the genome (mean/median <10). Sequences from viral metagenomes (Supplementary Table S4) were first filtered by mapping to the cyanophage reference genome using Bowtie2 (Langmead and Salzberg, 2012) set to the very-sensitive-local parameters (-D 20 -R 3 -N 0 -L 20 -i S,1,0.5). Successfully mapped reads were then recruited based on the best hit to the reference genome by BLASTn (E-value $<1 \times 10^{-5}$, >50 nt and $>90\%$ identity; Camacho *et al.*, 2009). Metagenomic islands were detected and defined largely as described in Mizuno *et al.* (2014), as stretches having $<20\%$ of the median coverage but with a minimum BLASTn identity of 90% (rather than 98%). This enabled identification of somewhat more divergent phage populations. The minimum island length was set at 300 bp (rather than 500 bp) to allow identification of islands corresponding to small genes.

Statistics

Before statistical analyses, the data were tested for normal distribution using the Shapiro–Wilk normality test ($P > 0.05$; Shahbaba, 2012). The Welch two-sample *t*-test was used when the data were normally distributed, whereas the nonparametric Mann–Whitney test was used when they were not. *P*-values were adjusted for multiple testing by the method of Benjamini and Hochberg (1995). Pearson's *r* coefficient was computed to assess the relationship between phage adsorption and fitness using the means of the measurements. Bootstrap resampling of observed metagenomic islands was used to generate, for each gene in each phage genome, a null distribution under the assumption that islands are randomly distributed in the genome. *P*-values were then estimated by comparing the observed occurrence of a gene in islands with the null distribution (Supplementary Methods). All statistical analysis were done in R (V.3.2.3 (R Development Core Team, 2015)).

Results and discussion

Phage resistance of some Prochlorococcus strains cannot be broken by phage mutants

Previously, we isolated strains of *Prochlorococcus* MED4 that were resistant to infection by host-specific T7-like cyanopodophages (Avrani *et al.*, 2011; Supplementary Figure S2). Here we investigated six phage–host interactions for the ability of the phages to overcome resistance in these cyanobacterial strains (Table 1). Specifically, populations of the P-SSP7 phage were screened for RBMs on four resistant strains (MED4-R1, MED4-R5, MED4-R6 and MED4-R8), as was the ability of the P-GSP1 and P-TIP38 phages to overcome resistance on MED4-R8. Initially an assay with a single selection step was used, whereby RBMs form plaques on lawns of resistant cells (see Materials and methods). Mutant P-SSP7 phages were found on two resistant strains, MED4-R5 and MED4-R6 (Table 1). In contrast, no RBMs were found on the other two resistant strains, MED4-R1 (with the P-SSP7 phage) and MED4-R8 (with the P-SSP7, P-GSP1 and P-TIP38 phages), despite screening over 10^{11} infective phage particles for each interaction (Table 1). Thus, RBMs were not found for four of the six interactions, suggesting that phage mutations enabling reinfection of some resistant strains are rare or even non-existent. Alternatively, multiple sequentially selected mutations may be required to overcome resistance (Scanlan *et al.*, 2011; Marston *et al.*, 2012), or RBMs may not be able to form plaques.

To address these possibilities, we employed a host-mixing experiment described by Benmayor *et al.* (2009) that facilitates selection of resistance-breakers over multiple steps of phage evolution in liquid. In this experiment, viral populations are serially transferred on mixed bacterial populations consisting of a minority of susceptible hosts and a majority of phage-resistant cells (see Materials and methods). We tested the P-SSP7 phage on two resistant hosts: one for which P-SSP7 mutants were detected in the single-step method (MED4-R6) and one for which they were not (MED4-R1). Consistent with our previous findings, resistance-breaking phage mutants were evident after three transfers in the presence of MED4-R6, as seen from lysis of the mixed susceptible and resistant host population (Figure 1). Mutants infecting MED4-R6 evolved in all host-mixing ratios tested (Supplementary Figure S3A), and were evident earlier when higher proportions of the susceptible host were present (Supplementary Figure S3B). This likely resulted from an increase in phage population size due to replication on a larger population of susceptible hosts, leading to an increased chance for resistance-breaking mutations to occur. In contrast, no RBMs capable of infecting MED4-R1 were observed even after 10 transfers of the evolving P-SSP7 population at all susceptible to resistant host ratios used (Figure 1; Supplementary Figure S3). Thus, our findings suggest that resistance-breaking phage mutants cannot emerge for some host mutations, and that such resistant hosts pose a genetic hurdle that limits the propagation of the arms race.

Inspection of our resistant strains did not yield any general insights into the types of mutations that permitted resistance-breaking versus those that posed a genetic barrier. Three of the four strains had mutations in a single cluster of non-core genes, whereas the fourth had a mutation in a core gene, all

Table 1 Frequencies of resistance-breaking mutants in populations of T7-like cyanophages forming plaques on *Prochlorococcus* MED4 phage-resistant strains

Host *Phage*	*MED4-R1* *P-SSP7*		*MED4-R5* *P-SSP7*		*MED4-R6* *P-SSP7*	
	Plated	*Freq.*	*Plated*	*Freq.*	*Plated*	*Freq.*
Lysate 1	1.6×10^{11}	0	9.9×10^{9}	1.1×10^{-8}	4.6×10^{10}	5.2×10^{-7}
Lysate 2	9.2×10^{10}	0	1.0×10^{10}	9.0×10^{-9}	4.0×10^{9}	3.0×10^{-8}
Lysate 3	9.9×10^{9}	0	1.4×10^{10}	1.8×10^{-8}	9.9×10^{9}	1.7×10^{-8}
Lysate 4			1.4×10^{10}	1.4×10^{-8}		
Total plated	$\mathbf{2.6 \times 10^{11}}$		$\mathbf{4.8 \times 10^{10}}$		$\mathbf{6.0 \times 10^{10}}$	
Mean freq.				1.3×10^{-8}		1.9×10^{-7}
± s.d.				$\pm 4.0 \times 10^{-9}$		$\pm 2.9 \times 10^{-7}$

Host	*MED4-R8*					
Phage	*P-SSP7*		*P-GSP1*		*P-TIP38*	
	Plated	*Freq.*	*Plated*	*Freq.*	*Plated*	*Freq.*
Lysate 1	1.6×10^{11}	0	7.7×10^{11}	0	7.1×10^{10}	0
Lysate 2	9.2×10^{10}	0	2.6×10^{8}	0	2.9×10^{10}	0
Lysate 3					2.2×10^{10}	0
Total plated	$\mathbf{2.5 \times 10^{11}}$		$\mathbf{7.7 \times 10^{11}}$		$\mathbf{1.2 \times 10^{11}}$	

Abbreviation: Freq., frequency.
Plated: number of infective phages plated to test for ability to infect resistant hosts. Freq.: ratio of plaque-forming units on resistant host/wild-type host. The mean of this frequency for P-SSP7 resistance-breaking mutants, when calculated from all seven experiments with non-zero frequencies, is 8.9×10^{-8} (1.9×10^{-8} s.d.). This corresponds to 1 in 6×10^{7} (4×10^{7} s.d.) phages when averaging the inverse of all frequency measurements.

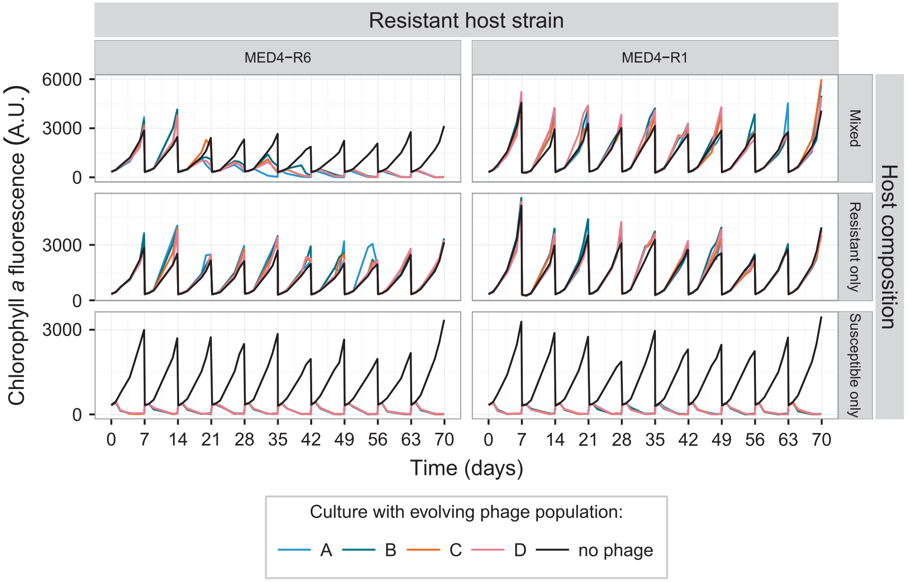

Figure 1 Experimental evolution of P-SSP7 resistance-breaking mutants. Four populations of the P-SSP7-WT phage were evolved on a host mix composed of 10% susceptible host cells (MED4-WT) and 90% phage-resistant cells: MED4-R6 (in left panels) and MED4-R1 (in right panels). P-SSP7-WT phages were added on day 0 (multiplicity of infection = 0.07). Every 7 days, for a total of 10 transfers, 1% of each of the four evolving phage populations were transferred to a naive host population of the same initial composition. Growth curves of host populations infected with descendants of the same evolving phage population are shown by the same color. Growth of these host populations without phage is shown in black ($n = 8$). Cell types infected separately served as controls for the ability of the evolving phage populations to infect the susceptible strain (MED4-WT) and for the inability to infect resistant strains (MED4-R1 and MED4-R6). a.u., arbitrary units. See Supplementary Figure S3 for similar results with different host mix ratios.

of which are located in the same genomic island (Supplementary Table S1A). Rather, the ability to overcome resistance appears related to the particular resistance gene harboring the mutation. The two resistant strains that yielded resistance-breaking phage mutants (MED4-R5 and MED4-R6) had different mutations in the PMM1247 gene, a gene of unknown function in *Prochlorococcus* MED4. However, a second mutation common to both MED4-R5 and MED4-R1 (in the PMM1249 gene) appears unrelated to the resistance-breaking phenomenon as the MED4-R1 strain did not yield any RBMs nor could it be infected by RBMs isolated on other strains.

Phenotype of RBMs

Six RBMs derived from P-SSP7 were isolated from plaques formed on lawns of the MED4-R5 and MED4-R6 strains during the single selection assay (Supplementary Table S1B). We began phenotypic characterization of three independently isolated phage mutants, RBM1, RBM2 and RBM3, by assessing their host range. All three RBMs maintained their ability to infect the WT MED4 and expanded their host range to infect both MED4-R5 and MED4-R6 (Supplementary Figure S4). However, they could not infect MED4-R1, or other *Prochlorococcus* or *Synechococcus* strains tested (Supplementary Table S1B). Thus, these mutants had a limited expansion in host range among some resistant MED4 types, but not beyond.

An expansion of host range, even if limited, is clearly beneficial to resistance-breaking phage mutants, providing them with an increased pool of susceptible hosts. However, the relative fitness of a resistance-breaking phage may not be the same on the original (MED4-WT) and acquired (MED4-R6) hosts. To evaluate their relative fitness, we measured phage fitness on each host as the phage progeny yield over multiple cycles of infection. Indeed, we found that all three of the RBMs had different fitness levels on their two hosts (Figure 2). Fitness of RBM1 and RBM2 was considerably lower on the acquired MED4-R6 host than on the original host by 54% and 66% for the two phages, respectively ($n = 7–13$, $P = 0.001–0.0002$). In contrast, the fitness of RBM3 was high on both its hosts, being only 8% greater on the acquired host ($n = 7–10$, $P = 0.01$). It is important to point out, however, that the lower fitness on the new host, as found for RBM1 and RBM2, should not be considered a cost, as this is a marked improvement over the WT phage that cannot infect this host at all.

Next, we assessed how the fitness of the RBMs compared to that of the ancestral WT phage (P-SSP7-WT) on their common MED4-WT host. Surprisingly, the fitness of RBM1 and RBM2 was significantly greater (30%) on the MED4-WT host relative to the ancestral WT phage ($n = 6–10$; $P = 0.001$). However, the fitness of RBM3 was significantly lower (16%) than that of the WT phage on their common host ($n = 7–10$; $P = 0.001$; Figure 2). Thus, the ability of RBM3 to efficiently infect the new host came with a cost on the original host, whereas less efficient infection of the acquired host by RBM1 and RBM2 did not lead to a fitness cost on the original host.

We expected that the gain in ability to infect resistant hosts would be manifested through adsorption to the resistant strains, as resistance of MED4-R5 (Avrani, 2011) and MED4-R6 (Figure 3a) impaired adsorption of P-SSP7-WT to the resistant strains. Consistent with this expectation, adsorption of RBM3 was significantly greater than that of P-SSP7-WT to MED4-R6 ($P = 0.005$). However, adsorption of RBM1 and RBM2 to this host remained low with no significant difference detectable in comparison to the WT phage (Figures 3a, $P = 0.57$ for RBM1 and $P = 0.24$ for RBM2). As adsorption is an essential first step for infection, some level of adsorption must have been regained and we assume that the increase was too small to be detected using the adsorption assay. Indeed, average phage fitness was linearly correlated with average phage adsorption (Figure 3b), suggesting that an important determinant of phage fitness that evolved in these phages was their capacity to adsorb to the resistant hosts.

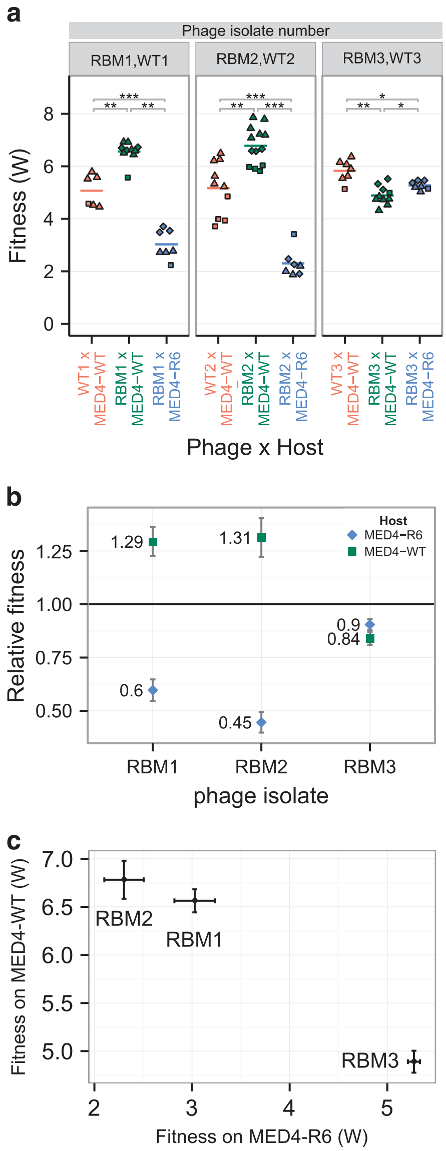

Figure 2 Fitness of resistance-breaking mutants (RBMs) and wild-type (WT) P-SSP7 phages. (**a**) Fitness is shown as the number of phage population doublings per day (W). Fitness was measured for all phages at 2 (triangle and diamond shapes) or 3 days (squares) after infection (Supplementary Figure S1) when infecting the MED4-WT host as well as for RBMs when infecting MED4-R6. Fitness on MED4-WT includes results from infections of RBM and WT phages separately and in direct competition (Supplementary Figure S9). Bars indicate means of all experiments ($n = 6$–13). Experiments carried out at the same time have the same shaped data symbols. Significant differences in fitness are indicated ($*P < 0.05$; $**P < 0.01$; $***P < 0.001$). The WT phage used is the paired control for the specific RBM. (**b**) Relative fitness of RBMs is shown as the ratio between mean fitness of RBMs on the tested host and the WT phage on the MED4-WT host (set as 1). Error bars show the propagated s.e.m. for the ratios of the fitness measurements in **a**. (**c**) Relationship between fitness on the original WT host (MED4-WT) and on the acquired resistant host (MED4-R6) for the three RBMs. Each datum point represents the mean ± s.e.m. of the measurements shown in **a**.

Interestingly, adsorption of RBM3 to the non-host *Prochlorococcus* strain, MIT9515, also increased significantly relative to the P-SSP7-WT phage (Figures 3a, $P = 0.02$). In contrast, adsorption of RBM1 and RBM2 to this non-host did not differ significantly from P-SSP7-WT (Figures 3a, $P = 0.86$ for RBM1 and $P = 0.24$ for RBM2). This suggests that the benefit of efficient adsorption of RBM3 to the MED4-R6 host comes with a cost of decreased adsorption specificity. This is likely to be quite detrimental in natural communities in which susceptible hosts are only a minute fraction of the community (Waterbury and Valois, 1993).

In summary, we found that RBM3, which efficiently infects the resistant host, suffers decreased fitness on its original host as well as lower adsorption specificity, whereas RBM1 and RBM2 are both less efficient in infecting the new host and have no measurable fitness cost associated with their mutation. Although our sample size of RBMs is small, these findings suggest a negative relationship between infectivity on the new host relative to fitness on the WT host (Figure 2c).

Genotype of RBM phages

To determine the mutations responsible for the resistance-breaking phenotypes, we sequenced the whole genome of the RBM1, RBM2 and RBM3 phages. Whole-genome sequencing revealed that they all had amino-acid changing mutations in two regions (Table 2; Figure 4a): all three mutants harbored the exact same mutation in gene *23* (numbering is according to the P-SSP7 genome). This gene encodes a short protein (93 aa) with a proline-rich N terminus and is positioned upstream of the portal protein gene in the P-SSP7 genome. In addition, each phage mutant had one to two mutations in genes *37* and *38*, which together with gene *39* form a cassette of three paralogous genes in P-SSP7 (Supplementary Figure S5) located directly downstream of the tail fiber gene, gene *36* (Sullivan

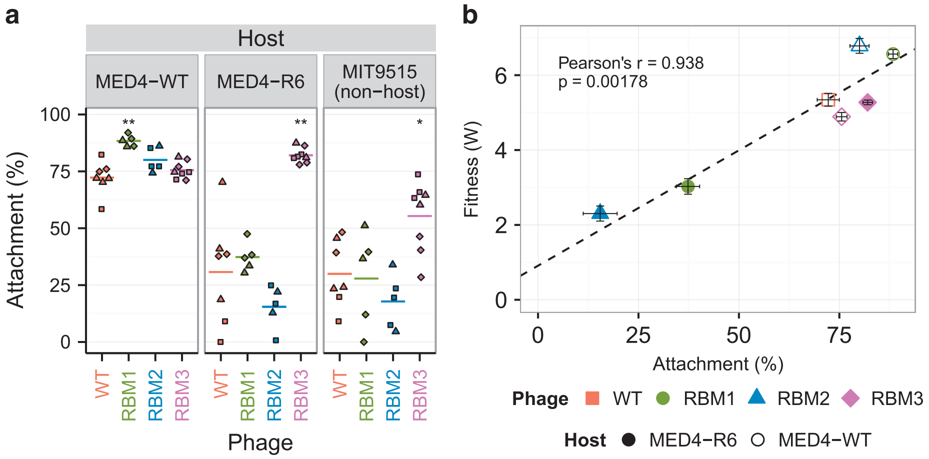

Figure 3 Attachment of resistance-breaking mutants (RBMs) and wild-type (WT) P-SSP7 phages to *Prochlorococcus*. (**a**) Percent attachment of the three RBMs and their WT ancestral phages to MED4-WT and MED4-R6 4 h after phage addition. *Prochlorococcus* sp. MIT9515 served as a control for attachment to a non-host strain. Significant differences in attachment between each of the RBMs and the WT phage are shown for each of the three hosts ($*P<0.05$; $**P<0.01$). Attachment of P-SSP7-WT to MED4-R6 is lower than to WT host ($P=0.027$). Attachment to resistant and non-host strains was similar for all phages ($P=0.95$) except RBM3. Bars indicate means of all experiments ($n=5$–8). Experiments performed at the same time have the same shaped data symbols. (**b**) Positive correlation between phage attachment and fitness (calculated for the means). Each data point represents the mean ± s.e.m. of the measurements shown in Figures 2a and 3a. Results from all three paired-control WT phages are combined.

et al., 2005; Liu *et al*., 2010). The RBM1 mutant also had a non-synonymous substitution in the exonuclease gene (gene *19*) in a non-conserved position (Table 2).

Genotyping of nine additional isolates of RBMs by sequencing PCR amplicons of genes *23*, *37* and *38* revealed all but one to have mutations in at least one of these genes (Table 2). Such evolutionary convergence of resistance-breaking mutations suggests that the protein products of genes *23*, 37 and *38* are important for determining host range in P-SSP7. Furthermore, our data suggest that a mutation in gene *23* is typically the first to enable resistance-breaking and that a subsequent mutation in genes *37* or *38* increases their virulence (Supplementary Results and Discussion). However, as 1 of the 12 mutants did not have a mutation in either of these genes, mutations in other regions can also confer the resistance-breaking phenotype.

Putative structural function of mutant genes

As genes *23*, *37* and *38* appear to be involved in host recognition and attachment of the phage to the cell surface of the host, one would expect them to form part of the tail structure. Although P-SSP7 gp23, gp37 and gp38 proteins are found in virion particles (Lindell *et al*., 2007), their exact location in the virion is unknown (Liu *et al*., 2010). Nonetheless, two lines of evidence suggest that these genes are likely to be tail-related: their distant homology to tail-related genes in other phages and the conservation of genomic position relative to homologous tail-related genes in the genome of T7.

We identified homologs of the P-SSP7 gene *23* in nearly all T7-like cyanophage genomes based on proline richness of the N terminus (Supplementary Figure S6; see Supplementary Results and Discussion). Furthermore, the position of this gene is conserved, being upstream of the portal protein gene (Supplementary Table S5). In coliphage T7, gene *7.3* is in this same genomic position and it shares remote homology with the C terminus of the cyanophage proline-rich proteins (Supplementary Results and Discussion). In T7, gp7.3 is involved in host recognition and attachment (Dunn and Studier, 1983; Kemp *et al*., 2005). Although its exact function and location in the T7 virion remains unknown (Kemp *et al*., 2005; Cuervo *et al*., 2013), it is presently thought to act as a scaffold for tail assembly (Kemp *et al*., 2005).

Unlike gene *23*, the P-SSP7 genes *37* and *38* have no homologs in other T7-like cyanophages (Supplementary Results and Discussion). Nevertheless, they do have a low level of identity to a tail fiber-like protein of a T4-like cyanophage (Supplementary Results and Discussion). The genomic position of these genes downstream of the tail fiber gene further alludes to a potential role in tail fiber structure.

Although direct evidence is lacking for the structural role of the protein products of genes *23*, *37* and *38*, current information suggests that these are previously unrecognized host-recognition proteins that function in tail assembly and in the tail fiber structure. Furthermore, these findings suggest that similar to other T7-like phages (Studier, 1975; Garcia *et al*., 2003; Qimron *et al*., 2006; Paterson *et al*., 2010; Scanlan *et al*., 2011), over-coming resistance in these T7-like cyanophages is related to tail fiber genes and genes involved in tail assembly.

Hyperdiversity of resistance-breaking genes

Coevolution of phages with their host populations can drive the rapid divergence of phage genes

Table 2 Characteristics of P-SSP7 resistance-breaking mutants (RBMs)

Phage	*Source*	*Isolation host*	*Isolation method*	*Genotyping*	*Gene 23*[a]	*Gene 37*[a]	*Gene 38*[a]	*T7-like exonuclease*
RBM1	Lysate	R6	Plaque	WGS	P42L	No mut	R127S; Y163H	No mut
RBM2	Lysate	R6	Plaque	WGS	P42L	No mut	M126I; R127H	E183K
RBM3	Lysate	R5	Plaque	WGS	P42L	SGTW[130-3]R[b]	No mut	No mut
RBM4	Lysate	R5	Plaque	PCR	No mut	No mut	SS [165-6] FRF[c]	ND
RBM5	Lysate	R5	Plaque	PCR	A43E	No mut	No mut	ND
RBM6	Lysate	R5	Plaque	PCR	N41T	No mut	No mut	ND
RBM7	Mix. pop. B (T2)[d]	R6	Liquid	PCR	P42L[e]	No mut	No mut	ND
RBM8	Mix. pop. B (T4)[d]	R6	Plaque	PCR	P42L[e]	No mut	No mut	ND
RBM9	Mix. pop. B (T6)[d]	R6	Plaque	PCR	P42L[e]	No mut	Y163H	ND
RBM10	Mix. pop. C (T2)[d]	R6	Liquid	PCR	P42L	No mut	No mut	ND
RBM11	Mix. pop. C (T4)[d]	R6	Plaque	PCR	No mut	No mut	No mut	ND
RBM12	Mix. pop. C (T6)[d]	R6	Plaque	PCR	P42L	No mut	N94H	ND

Abbreviations: WGS, whole-genome sequencing; no mut, no mutation found; ND, no data.
[a] Gene numbers are as in the P-SSP7 reference genome (NC_006882) and mutations are relative to wild-type paired controls.
[b] Deletion.
[c] Substitution and insertion.
[d] Isolated from host-mixing experiments (mix. pop.) from population B or C at transfer 2, 4 or 6 (T#).
[e] Carries an additional, synonymous, mutation in N41.

involved in recognition and attachment to the host relative to the rest of the phage genome (Paterson *et al.*, 2010; Scanlan *et al.*, 2011; Koskella and Brockhurst, 2014). To test for signals of hyperdiversity in resistant-breaking mutant genes, we analyzed their diversity across the genomes of T7-like cyanophages. The occurrence of genes *37* and *38* in the P-SSP7 genome alone indicates that these genes are part of the flexible genome and that their hyperdiversity is at the level of gene content. In contrast, we identified homologs of gene *23* in all T7-like cyanophage genomes but one. Thus, we consider this to be a core gene (Supplementary Results and Discussion). A comparison of protein sequence diversity of gene *23* to that of other T7-like cyanophage core genes showed that it is among the most diverse core genes (Figure 5). The two genes that are more diverse, the tail fiber protein and one of the three internal core proteins that form the channel that transverses the host cell envelope (Hu *et al.,* 2013), are also likely to be involved in host range in T7 and other T7-like phages (Qimron *et al.*, 2006; Paterson *et al.*, 2010; Scanlan *et al.*, 2011). Thus, the most diverse core genes in T7-like cyanophages function in host recognition, and their hyperdiversity is likely the result of rapid coevolution with the host.

Next, we examined the diversity of these resistance-breaking genes in natural populations in the oceans using recruitment analysis. We focused on 12 viral metagenomes: 3 from the Pacific Ocean and the others from the Indian Ocean (Supplementary Figure S7), in which P-SSP7-like phages were present. The presence of P-SSP7 in these metagenomes is evident from the high coverage of metagenomic sequences along most of the P-SSP7 genome (see Materials and methods). However, some genomic regions have very low coverage. Such under-recruiting regions, known as metagenomic islands, have been previously observed for bacteria and phages (Rodriguez-Valera *et al.*, 2009; Mizuno *et al.*, 2014) and can be explained by a high level of sequence divergence or even absence of these genes from natural populations.

The most under-represented genomic region of P-SSP7 in the metagenomes is that of the tail fiber gene and the four genes downstream of it, including resistance-breaking genes *37* and *38* (Figure 4). The occurrence of these genes in metagenomic islands (that is, their divergence or absence in viromes) is significantly higher than would be expected had the islands been randomly distributed over the genome ($P<0.012$). A similar pattern of under-recruitment downstream of the tail fiber gene was found for seven of the eight other T7-like cyanophages that are well represented in viral metagenomes (Supplementary Figure S8). These findings are in agreement with previous comparative genomic analyses showing variable gene content downstream of the tail fiber gene in T7-like cyanophages (Island III in Labrie *et al.* (2013)) and in viral metagenomes mapping to T7-like pelagiphages (Mizuno *et al.*, 2014).

In P-SSP7, gene *23* was also significantly under-represented in viral metagenomes, having a low recruitment in 8 of 19 of the metagenomic data sets (Figure 5). Of the reads that did recruit to gene *23*, half (49% of 600) had polymorphisms in at least one of the three amino-acid positions for which we experimentally found resistance-breaking mutations (in positions 41, 42 and 43). Only two code for an amino acid identical to that in our mutants (N41T) with the vast majority found in position 43. These data indicate that the diversity of gene *23* is both at the level of sequence diversification and gene content in natural populations of P-SSP7-like phages. It should be noted that unlike for P-SSP7 the proline-rich homologs of gene *23* were not under-represented in any of the other T7-like cyanophages (Supplementary Figure S8).

In summary, our results directly link evolution of host recognition with diversity in T7-like cyanophages and suggest that the coevolutionary process drives this diversity. Indeed, resistance-breaking loci

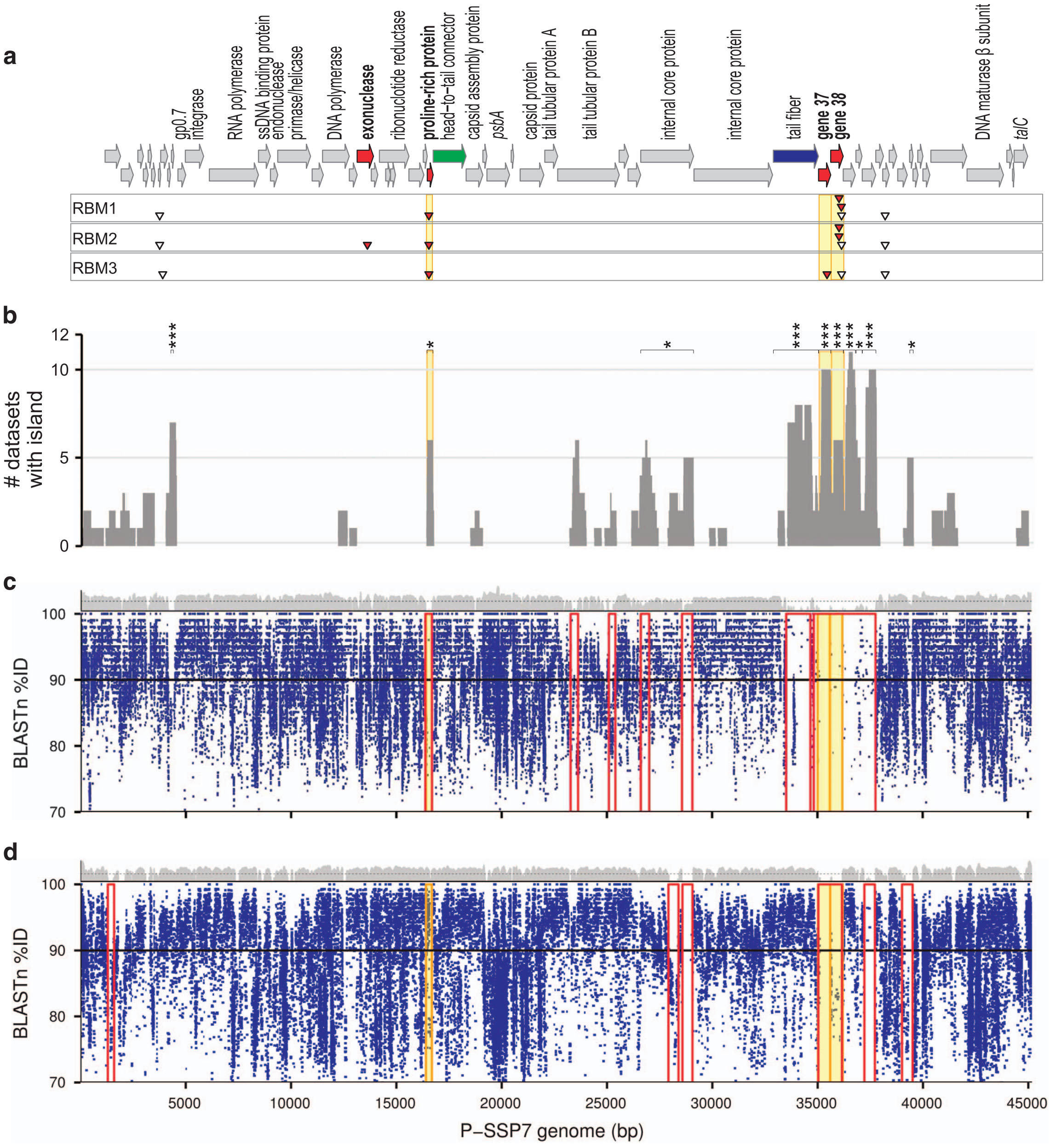

Figure 4 Resistance-breaking mutations are in genes occurring in metagenomic islands in wild populations of P-SSP7. (**a**) Convergence of resistance-breaking mutations in three independent phage mutants. The genomic map of P-SSP7 is shown with the positions of resistance-breaking genes marked in red. The portal protein and tail fiber genes are marked in green and blue, respectively. Triangles mark the position of mutations in the genomes of each of the three completely sequenced P-SSP7 resistance-breaking mutants (RBMs) compared with the reference genome (NC_006882). Mutations unique to the RBM are in red, whereas mutations shared with wild-type paired-control phages are in white. (see Table S2 and Supplementary Results and Discussion for details). (**b**) The locations of metagenomic islands that under-recruited to the genome of P-SSP7 are shown for 12 viral metagenomes for which P-SSP7 was present (detailed in Supplementary File 3). Significant enrichment of genes in observed metagenomic islands compared with randomly positioned islands of equivalent size in 10 000 bootstrap replicates are indicated (*$P<0.05$; ***$P<0.001$). (**c**, **d**) Examples of recruitment plots of viral metagenomes from the Indian (**c**, TARA station 38, 5 m) and Pacific (**d**, ALOHA station, 120 m) oceans. These are representative of data sets with (**c**) and without (**d**) metagenomic islands for gene *23*, encoding the proline-rich protein. Recruited sequence fragments (in blue) are mapped along the P-SSP7 genome (*x* axis) at their identity level to that genome (*y* axis). At the top of each recruitment plot, the base-specific read coverage of the metagenomic fragments with over 90% BLASTn identity is shown in log scale. The dashed line in this plot marks the coverage threshold used for designating regions as metagenomic islands. The positions of resistance-breaking genes are indicated by yellow boxes and shading in all panels. Metagenomic islands are indicated by red boxes in **c** and **d**.

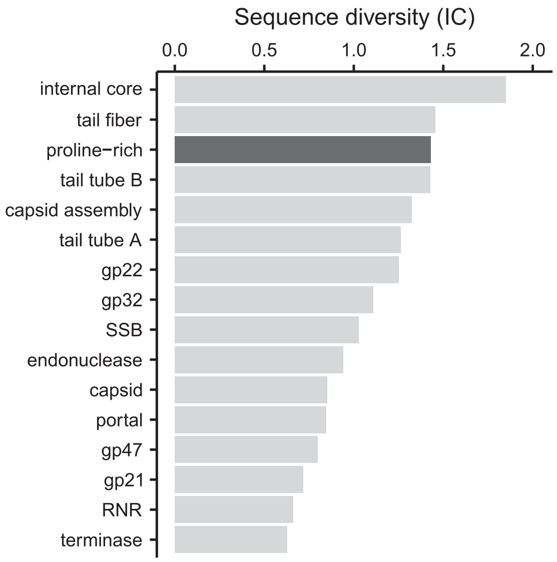

Figure 5 Diversity of the proline-rich proteins (black) relative to proteins of other core genes (gray) in T7-like cyanophages. Diversity across eight genomes was calculated as information content (IC; Avrani *et al.*, 2011), ranging from 0, which indicates identity of all sequences, to 4.3, indicating no identity across the entire alignment. For genes of unknown function, the P-SSP7 protein numbers are indicated. DNA and RNA polymerases were not included in the analysis due to the occurrence of frameshifts in gene sequence for many genomes that hinders the inference of the corresponding amino-acid sequence; however, these are among the most conserved core genes. SSB, single-strand binding; RNR, ribonucleotide reductase.

identified experimentally are highly diverse in wild populations of T7-like cyanophages in both gene sequence and gene content. The diversity of these resistance-breaking loci is reminiscent of the diversity of susceptibility regions observed in *Prochlorococcus* (Avrani *et al.*, 2011). Thus, antagonistic coexistence has led to rapid evolution and population-level diversity in genomic islands harboring genes involved in phage recognition of hosts in both host and phage genomes.

A combination of models is needed to explain coexistence in nature

Our combined findings suggest that an integrative model incorporating multiple coevolutionary processes is required to adequately explain antagonistic coexistence between *Prochlorococcus* and its T7-like cyanophages. On one hand, our ability to isolate resistant hosts (Avrani *et al.*, 2011) and RBMs with an expanded host range for at least some resistant *Prochlorococcus* hosts (this study) demonstrates that arms race cycles can occur (Figure 6). Furthermore, hyperdiversity of resistance and resistance-breaking loci in hosts and phages suggests that mutations and counter-mutations with reciprocal selection cycles occur in natural populations and are not purged by selective sweeps or genetic barriers to coevolution. On the other hand, our finding of many cases for which RBMs did not evolve suggests that phage counter-mutations alone cannot explain phage survival. Indeed, if resistant hosts sweep through the population, host-specific phages would remain host-less and become extinct. Such cases would constitute a dead end to the cyclic nature of the arms race.

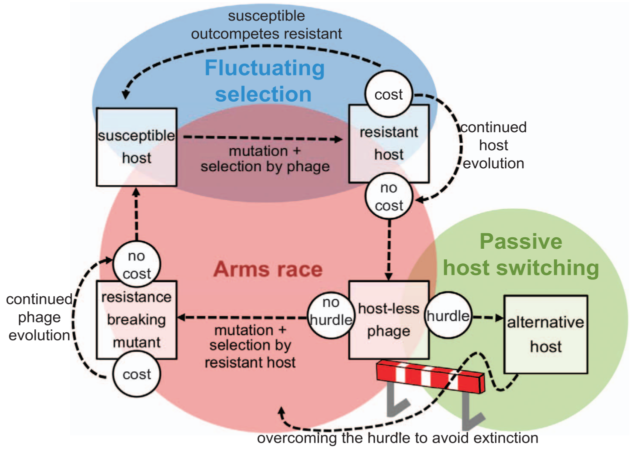

Figure 6 Model of multiple coevolutionary processes proposed to act at the same time on different interacting partners to enable community-level coexistence of *Prochlorococcus* and T7-like cyanophages. Each evolutionary process is shaded by a different color. Hosts and phages in different phenotypic states with respect to the interaction with each other are shown in white squares. Different manifestations of these states are shown in white circles. Descriptions of transitions between the different states are indicated by arrows. Phages that overcome the hurdle of a genetic barrier by passive host-switching avoid extinction and return to the evolutionary playing field. The hurdle imposed by a genetic barrier (marked by a white circle) is the same hurdle that is overcome by passive host-switching (marked by the red and white bar).

Alternatively, density-dependent fluctuations between subpopulations of hosts that differ in their genetic susceptibility to phages could explain coexistence, even in the absence of resistance-breaking phage mutants (Avrani *et al.*, 2011; Avrani *et al.*, 2012). These fluctuations require, and are driven by, costs of resistance such as reduced growth rate and greater susceptibility to other phage types (Avrani *et al.*, 2011; Figure 6). However, apparent cost-less phage resistance mutations (Duffy *et al.*, 2006; Lennon *et al.*, 2007; Avrani *et al.*, 2011; Ford *et al.*, 2014) and costs that are ameliorated with continued evolution (Lenski, 1988b; Kashiwagi and Yomo, 2011; Avrani and Lindell, 2015) are known to occur. Thus, these situations would still leave host-specific phages host-less, leading to their extinction.

Thus, a mechanism that enables phage persistence in the face of a reduction in the pool of susceptible hosts must exist. Previously, we suggested that passive host-switching can balance such host loss without a mutation occurring in the phage to overcome host resistance (Avrani *et al.*, 2012). In this scenario, a phage acquires a new host as a result of a mutation in a non-host cell that turns it into a host. This can be the result of continued evolution that returns a resistant host to susceptibility (Meyer *et al.*, 2010; Avrani and Lindell, 2015) or due to a mutation that grants a cell resistance to one phage type while making it susceptible to another, as in enhanced infection (Lennon *et al.*, 2007; Avrani *et al.*, 2011; Marston *et al.*, 2012; Castillo *et al.*, 2014; Figure 6). These processes can allow the recoupling of a phage with the original host it could previously infect or enable a phage to infect a different host entirely. Thus, in the context of the community, passive host-switching would prevent phage extinction despite resistance mutations in hosts that cannot be overcome by counter-mutations in phages.

In light of the above, we propose a three-pronged model, whereby a limited evolutionary arms race, density-dependent fluctuating selection and passive host-switching all work together to enable host–phage coexistence in the wild (Figure 6). This model paints a picture of complex host–phage coexistence that shifts between fluctuating selection and arms race dynamics as a function of fitness costs in the host and their elimination by continued evolution. And, when genetic barriers arise to the emergence of RBMs in the phage, passive host-switching supplies alternative hosts from within the diverse host community. This combined model is compatible with the high degree of microbial and viral diversity seen in wild communities of both *Prochlorococcus* and their phages in genomic regions previously hypothesized to be (Rodriguez-Valera *et al.*, 2009; Labrie *et al.*, 2013; Kashtan *et al.*, 2014), and now empirically shown to be (Avrani *et al.*, 2011; Avrani and Lindell, 2015, this study), involved in phage recognition of hosts. This population-level diversity leads to a mix of host subpopulations with a variety of susceptibilities to co-occurring phages and a mix of phage subpopulations with varying abilities to infect hosts. The composition and structure of such populations would constantly change due to the coevolutionary process, leading to a complex and dynamic network of host–phage interactions that, perhaps counter-intuitively, result in a stable community in nature. Indeed, such diversity-induced stable coexistence likely extends to many host–phage systems (Rodriguez-Valera *et al.*, 2009; Flores *et al.*, 2011; Cordero and Polz, 2014; Mizuno *et al.*, 2014).

Our findings show a clear asymmetry in the evolutionary capacity between *Prochlorococcus* and T7-like cyanophages. We found that resistant hosts are commonly isolated (Avrani *et al.*, 2011), yet phage mutants that overcome this resistance were often not found (this study). Even when phage RBMs were observed, they occurred at a frequency that is ~100–1000-fold lower than the frequencies reported for phage resistance in the hosts (1 in 3×10^4 (2×10^4 s.d.) resistant hosts (Avrani *et al.*, 2011) versus 1 in 6×10^7 (4×10^7 s.d.) resistance-breaking phages (Table 1)). Evolutionary asymmetry between hosts and phages is not unique to our system. This asymmetry, in which hosts outplay phages, has been observed in different model systems from a variety of environments including the gut, soil, phyllosphere and the oceans (Lenski, 1984; Bohannan and Lenski, 2000; Gomez and Buckling, 2011; Koskella and Parr, 2015; Scanlan *et al.*, 2015, this study), although the form in which this asymmetry is manifested may vary for different model systems. The advantage for the host may be due to the relative ease of acquiring a loss-of-function mutation in a host cell surface molecule, whereas obtaining a specific gain-of-function mutation by the phage is likely to be more limited. This would be most severe when a host mutation leads to complete loss of the phage receptor, as was previously suggested by Bohannan and Lenski (2000). Irrespective of the reasons for the asymmetry, these findings indicate that the coevolutionary mechanisms supporting coexistence are not necessarily a mirror image between hosts and their phages. Rather, multilayered coevolutionary processes, of arms race, fluctuating selection and passive host-switching, are likely to be in play at any one time among different subpopulations of each of the antagonists and these may well differ between microbial hosts and their viruses.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank Alon Philosof and Fabian Glaser for advice on recruitment and homology analysis, respectively, Sarit Avrani and members of the Lindell and Beja labs for discussions, Michael Carlson and Sarit Avrani for comments on the manuscript, and three reviewers for their comments that helped improve the manuscript. Genome sequencing was carried out at the Technion Genome Center in the Lorry I. Lokey Interdisciplinary Center for Life Sciences and Engineering at the Technion - Israel Institute of Technology. This research was supported by grants from the European Research Council (Starting Grant 203406) to D.L., from the Simons Foundation (SCOPE Award ID 329108, D.L.), and by the Lorry I. Lokey Interdisciplinary Center for Life Sciences and Engineering and the Russell Berrie Nanotechnology Institute at the Technion.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Avrani S, Wurtzel O, Sharon I, Sorek R, Lindell D. (2011). Genomic island variability facilitates *Prochlorococcus*-virus coexistence. *Nature* **474**: 604–608.

Avrani S, Schwartz DA, Lindell D. (2012). Virus-host swinging party in the oceans: incorporating biological complexity into paradigms of antagonistic coexistence. *Mob Genet Elements* **2**: 88–95.

Avrani S, Lindell D. (2015). Convergent evolution toward an improved growth rate and a reduced resistance range in *Prochlorococcus* strains resistant to phage. *Proc Natl Acad Sci USA* **112**: E2191–E2200.

Benjamini Y, Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Methodol* **57**: 289–300.

Benmayor R, Hodgson DJ, Perron GG, Buckling A. (2009). Host mixing and disease emergence. *Curr Biol* **19**: 764–767.

Bohannan BJM, Lenski RE. (2000). Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. *Ecol Lett* **3**: 362–377.

Buckling A, Rainey PB. (2002). Antagonistic coevolution between a bacterium and a bacteriophage. *Proc Biol Sci* **269**: 931–936.

Buckling A, Brockhurst M. (2012). Bacteria-virus coevolution. *Adv Exp Med Biol* **751**: 347–370.

Bull JJ, Badgett MR, Wichman HA, Huelsenbeck JP, Hillis DM, Gulati A *et al.* (1997). Exceptional convergent evolution in a virus. *Genetics* **147**: 1497–1507.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K *et al.* (2009). BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.

Castillo D, Christiansen RH, Espejo R, Middelboe M. (2014). Diversity and geographical distribution of *Flavobacterium psychrophilum* isolates and their phages: patterns of susceptibility to phage infection and phage host range. *Microb Ecol* **67**: 748–757.

Chao L, Levin BR, Stewart FM. (1977). A complex community in a simple habitat: An experimental study with bacteria and phage. *Ecology* **58**: 369–378.

Chen F, Lu J. (2002). Genomic sequence and evolution of marine cyanophage P60: a new insight on lytic and lysogenic phages. *Appl Environ Microbiol* **68**: 2589–2594.

Coleman ML, Sullivan MB, Martiny AC, Steglich C, Barry K, Delong EF *et al.* (2006). Genomic islands and the ecology and evolution of *Prochlorococcus*. *Science* **311**: 1768–1770.

Cordero OX, Polz MF. (2014). Explaining microbial genomic diversity in light of evolutionary ecology. *Nat Rev Microbiol* **12**: 263–273.

Cuervo A, Pulido-Cid M, Chagoyen M, Arranz R, Gonzalez-Garcia VA, Garcia-Doval C *et al.* (2013). Structural characterization of the bacteriophage T7 tail machinery. *J Biol Chem* **288**: 26290–26299.

Dekel-Bird NP, Avrani S, Sabehi G, Pekarsky I, Marston MF, Kirzner S *et al.* (2013). Diversity and evolutionary relationships of T7-like podoviruses infecting marine cyanobacteria. *Environ Microbiol* **15**: 1476–1491.

Dekel-Bird NP, Sabehi G, Mosevitzky B, Lindell D. (2015). Host-dependent differences in abundance, composition and host range of cyanophages from the Red Sea. *Environ Microbiol* **17**: 1286–1299.

Duffy S, Turner PE, Burch CL. (2006). Pleiotropic costs of niche expansion in the RNA bacteriophage Φ6. *Genetics* **172**: 751–757.

Dunn JJ, Studier FW. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J Mol Biol* **166**: 477–535.

Edgar RC. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.

Finn RD, Clements J, Eddy SR. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* **39**: W29–W37.

Flombaum P, Gallegos JL, Gordillo RA, Rincon J, Zabala LL, Jiao N *et al.* (2013). Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc Natl Acad Sci USA* **110**: 9824–9829.

Flores CO, Meyer JR, Valverde S, Farr L, Weitz JS. (2011). Statistical structure of host-phage interactions. *Proc Natl Acad Sci USA* **108**: E288–E297.

Ford BE, Sun B, Carpino J, Chapler ES, Ching J, Choi Y *et al.* (2014). Frequency and fitness consequences of bacteriophage Φ6 host range mutations. *PLoS One* **9**: e113078.

Forde SE, Thompson JN, Holt RD, Bohannan BJ. (2008). Coevolution drives temporal changes in fitness and diversity across environments in a bacteria-bacteriophage interaction. *Evolution* **62**: 1830–1839.

Garcia E, Elliott JM, Ramanculov E, Chain PS, Chu MC, Molineux IJ. (2003). The genome sequence of *Yersinia pestis* bacteriophage ϕA1122 reveals an intimate history with the coliphage T3 and T7 genomes. *J Bacteriol* **185**: 5248–5262.

Goecks J, Nekrutenko A, Taylor J, Team TG. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* **11**: R86.

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

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

Harrison B. (2002). Virus variation in relation to resistance-breaking in plants. *Euphytica* **124**: 181–192.

Hu B, Margolin W, Molineux IJ, Liu J. (2013). The bacteriophage T7 virion undergoes extensive structural remodeling during infection. *Science* **339**: 576–579.

Hyman P, Abedon ST. (2010). Bacteriophage host range and bacterial resistance. *Adv Appl Microbiol* **70**: 217–248.

Johnson ZI, Zinser ER, Coe A, McNulty NP, Woodward EM, Chisholm SW. (2006). Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* **311**: 1737–1740.

Kashiwagi A, Yomo T. (2011). Ongoing phenotypic and genomic changes in experimental coevolution of RNA bacteriophage Qβ and *Escherichia coli*. *PLoS Genet* **7**: e1002188.

Kashtan N, Roggensack SE, Rodrigue S, Thompson JW, Biller SJ, Coe A *et al.* (2014). Single-cell genomics reveals hundreds of coexisting subpopulations in wild *Prochlorococcus*. *Science* **344**: 416–420.

Katoh K, Misawa K, Kuma K, Miyata T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**: 3059–3066.

Kemp P, Garcia LR, Molineux IJ. (2005). Changes in bacteriophage T7 virion structure at the initiation of infection. *Virology* **340**: 307–317.

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

Koskella B, Parr N. (2015). The evolution of bacterial resistance against bacteriophages in the horse chestnut phyllosphere is general across both space and time. *Philos Trans R Soc Lond B Biol Sci* **370**: 20140297.

Labrie SJ, Samson JE, Moineau S. (2010). Bacteriophage resistance mechanisms. *Nat Rev Microbiol* **8**: 317–327.

Labrie SJ, Frois-Moniz K, Osburne MS, Kelly L, Roggensack SE, Sullivan MB *et al.* (2013). Genomes of marine cyanopodoviruses reveal multiple origins of diversity. *Environ Microbiol* **15**: 1356–1376.

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

Lennon JT, Khatana SAM, Marston MF, Martiny JBH. (2007). Is there a cost of virus resistance in marine cyanobacteria? *ISME J* **1**: 300–312.

Lennon JT, Martiny JBH. (2008). Rapid evolution buffers ecosystem impacts of viruses in a microbial food web. *Ecol Lett* **11**: 1178–1188.

Lenski RE. (1984). Coevolution of bacteria and phage: are there endless cycles of bacterial defenses and phage counterdefenses? *J Theor Biol* **108**: 319–325.

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

Lenski RE. (1988a). Experimental studies of pleiotropy and epistasis in *Escherichia coli.* I. Variation in competitive fitness among mutants resistant to virus T4. *Evolution*, 425–432.

Lenski RE. (1988b). Experimental studies of pleiotropy and epistasis in *Escherichia coli.* II. compensation for maldaptive effects associated with resistance to virus T4. *Evolution* **42**: 433–440.

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

Lindell D, Padan E, Post AF. (1998). Regulation of ntcA expression and nitrite uptake in the marine *Synechococcus* sp. strain WH 7803. *J Bacteriol* **180**: 1878–1886.

Lindell D, Sullivan MB, Johnson ZI, Tolonen AC, Rohwer F, Chisholm SW. (2004). Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. *Proc Natl Acad Sci USA* **101**: 11013–11018.

Lindell D, Jaffe JD, Coleman ML, Futschik ME, Axmann IM, Rector T *et al.* (2007). Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. *Nature* **449**: 83–86.

Lindell D The genus *Prochlorococcus*, phylum Cyanobacteria. In Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (eds). *The Prokaryotes: Other Major Lineages of Bacteria and the Archaea*. Springer: Berlin, Heidelberg, Germany, 2014, pp 829–845.

Liu X, Zhang Q, Murata K, Baker ML, Sullivan MB, Fu C *et al.* (2010). Structural changes in a marine podovirus associated with release of its genome into *Prochlorococcus*. *Nat Struct Mol Biol* **17**: 830–836.

Mann NH. (2003). Phages of the marine cyanobacterial picophytoplankton. *FEMS Microbiol Rev* **27**: 17–34.

Marston MF, Pierciey FJ Jr, Shepard A, Gearin G, Qi J, Yandava C *et al.* (2012). Rapid diversification of coevolving marine *Synechococcus* and a virus. *Proc Natl Acad Sci USA* **109**: 4544–4549.

Martiny JBH, Riemann L, Marston MF, Middelboe M. (2014). Antagonistic coevolution of marine planktonic viruses and their hosts. *Ann Rev Mar Sci* **6**: 393–414.

Meyer JR, Agrawal AA, Quick RT, Dobias DT, Schneider D, Lenski RE. (2010). Parallel changes in host resistance to viral infection during 45 000 generations of relaxed selection. *Evolution* **64**: 3024–3034.

Meyer JR, Dobias DT, Weitz JS, Barrick JE, Quick RT, Lenski RE. (2012). Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* **335**: 428–432.

Mizoguchi K, Morita M, Fischer CR, Yoichi M, Tanji Y, Unno H. (2003). Coevolution of bacteriophage PP01 and *Escherichia coli* O157:H7 in continuous culture. *Appl Environ Microbiol* **69**: 170–176.

Mizuno CM, Ghai R, Rodriguez-Valera F. (2014). Evidence for metaviromic islands in marine phages. *Front Microbiol* **5**: 27.

Moore LR, Coe A, Zinser ER, Saito MA, Sullivan MB, Lindell D *et al.* (2007). Culturing the marine cyanobacterium *Prochlorococcus*. *Limnol Oceanogr Methods* **5**: 353–362.

Morris JJ, Kirkegaard R, Szul MJ, Johnson ZI, Zinser ER. (2008). Facilitation of robust growth of *Prochlorococcus* colonies and dilute liquid cultures by 'helper' heterotrophic bacteria. *Appl Environ Microbiol* **74**: 4530–4534.

Muhling M, Fuller NJ, Millard A, Somerfield PJ, Marie D, Wilson WH *et al.* (2005). Genetic diversity of marine *Synechococcus* and co-occurring cyanophage communities: evidence for viral control of phytoplankton. *Environ Microbiol* **7**: 499–508.

Palloix A, Ayme V, Moury B. (2009). Durability of plant major resistance genes to pathogens depends on the genetic background, experimental evidence and consequences for breeding strategies. *New Phytol* **183**: 190–199.

Parsons RJ, Breitbart M, Lomas MW, Carlson CA. (2012). Ocean time-series reveals recurring seasonal patterns of virioplankton dynamics in the northwestern Sargasso Sea. *ISME J* **6**: 273–284.

Partensky F, Hess WR, Vaulot D. (1999). *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol Mol Biol Rev* **63**: 106–127.

Paterson S, Vogwill T, Buckling A, Benmayor R, Spiers AJ, Thomson NR *et al.* (2010). Antagonistic coevolution accelerates molecular evolution. *Nature* **464**: 275–278.

Perry EB, Barrick JE, Bohannan BJM. (2015). The molecular and genetic basis of repeatable coevolution between *Escherichia coli* and bacteriophage T3 in a laboratory microcosm. *PLoS ONE* **10**: e0130639.

Pope WH, Weigele PR, Chang J, Pedulla ML, Ford ME, Houtz JM *et al.* (2007). Genome sequence, structural proteins, and capsid organization of the cyanophage Syn5: a "horned" bacteriophage of marine *Synechococcus*. *J Mol Biol* **368**: 966–981.

Poullain V, Gandon S, Brockhurst MA, Buckling A, Hochberg ME. (2008). The evolution of specificity in evolving and coevolving antagonistic interactions between a bacteria and its phage. *Evolution* **62**: 1–11.

Qimron U, Marintcheva B, Tabor S, Richardson CC. (2006). Genomewide screens for *Escherichia coli* genes affecting growth of T7 bacteriophage. *Proc Natl Acad Sci USA* **103**: 19039–19044.

R Development Core Team. (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria. Available at: http://www.R-project.org

Ravin V, Raisanen L, Alatossava T. (2002). A conserved C-terminal region in Gp71 of the small isometric-head phage LL-H and ORF474 of the prolate-head phage JCL1032 is implicated in specificity of adsorption of phage to its host, *Lactobacillus delbrueckii*. *J Bacteriol* **184**: 2455–2459.

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G *et al.* (2011). Integrative genomics viewer. *Nat Biotechnol* **29**: 24–26.

Rodriguez-Valera F, Martin-Cuadrado AB, Rodriguez-Brito B, Pasic L, Thingstad TF, Rohwer F *et al.* (2009). Explaining microbial population genomics through phage predation. *Nat Rev Microbiol* **7**: 828–836.

Sabehi G, Shaulov L, Silver DH, Yanai I, Harel A, Lindell D. (2012). A novel lineage of myoviruses infecting cyanobacteria is widespread in the oceans. *Proc Natl Acad Sci USA* **109**: 2037–2042.

Samson JE, Magadan AH, Sabri M, Moineau S. (2013). Revenge of the phages: defeating bacterial defences. *Nat Rev Microbiol* **11**: 675–687.

Scanlan PD, Hall AR, Lopez-Pascua LD, Buckling A. (2011). Genetic basis of infectivity evolution in a bacteriophage. *Mol Ecol* **20**: 981–989.

Scanlan PD, Hall AR, Blackshields G, Friman VP, Davis MR Jr, Goldberg JB *et al.* (2015). Coevolution with bacteriophages drives genome-wide host evolution and constrains the acquisition of abiotic-beneficial mutations. *Mol Biol Evol* **32**: 1425–1435.

Shahbaba B. (2012). *Biostatistics with R*. Springer: New York, USA.

Söding J. (2005). Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**: 951–960.

Studier FW. (1975). Gene 0.3 of bacteriophage T7 acts to overcome the DNA restriction system of the host. *J Mol Biol* **94**: 283–295.

Sullivan MB, Waterbury JB, Chisholm SW. (2003). Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. *Nature* **424**: 1047–1051.

Sullivan MB, Coleman ML, Weigele P, Rohwer F, Chisholm SW. (2005). Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. *PLoS Biol* **3**: e144.

Suttle CA, Chan AM. (1993). Marine cyanophages infecting oceanic and coastal strains of *Synechococcus*: abundance, morphology, cross-infectivity and growth characteristics. *Mar Ecol Prog Ser* **32**: 99–109.

Thompson LR, Zeng Q, Kelly L, Huang KH, Singer AU, Stubbe J *et al.* (2011). Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc Natl Acad Sci USA* **108**: E757–E764.

Wang K, Chen F. (2008). Prevalence of highly host-specific cyanophages in the estuarine environment. *Environ Microbiol* **10**: 300–312.

Waterbury JB, Valois FW. (1993). Resistance to co-occurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. *Appl Environ Microbiol* **59**: 3393–3399.

Winter C, Bouvier T, Weinbauer MG, Thingstad TF. (2010). Trade-offs between competition and defense specialists among unicellular planktonic organisms: the 'killing the winner' hypothesis revisited. *Microbiol Mol Biol Rev* **74**: 42–57.

Woolhouse ME, Webster JP, Domingo E, Charlesworth B, Levin BR. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat Genet* **32**: 569–577.

Zwirglmaier K, Jardillier L, Ostrowski M, Mazard S, Garczarek L, Vaulot D *et al.* (2008). Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. *Environ Microbiol* **10**: 147–161.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

© The Author(s) 2017

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)