Phage and bacteria diversification through a prophage acquisition ratchet

Maxwell Anthenelli*[1,6], Emily Jasien*[2,6], Robert Edwards[5,6,7], Barbara Bailey[2,6], Ben Felts[2,6], Parag Katira[4,6], James Nulton[2,6], Peter Salamon[2,6], Forest Rohwer[3,6], Cynthia B. Silveira[3,6], and Antoni Luque†[2,6,7]

[1] *Department of Physics, San Diego State University, San Diego, USA.*
[2] *Department of Mathematics & Statistics, San Diego State University, San Diego, USA.*
[3] *Department of Biology, San Diego State University, San Diego, USA.*
[4] *Department of Mechanical Engineering, San Diego State University, San Diego, USA.*
[5] *Department of Computer Science, San Diego State University, San Diego, USA.*
[6] *Viral Information Institute, San Diego State University, San Diego, USA.*
[7] *Computational Science Research Center, San Diego State University, San Diego, USA.*

April 8, 2020

* Co-first authors.

† Corresponding author.

Abstract

Lysogeny is prevalent in the microbial-dense mammalian gut. This contrasts the classical view of lysogeny as a refuge used by phages under poor host growth conditions. Here we hypothesize that as carrying capacity increases, lysogens escape phage top-down control through superinfection exclusion, overcoming the canonical trade-off between competition and resistance. This hypothesis was tested by developing an ecological model that combined lytic and lysogenic communities and a diversification model that estimated the accumulation of prophages in bacterial genomes. The ecological model sampled phage-bacteria traits stochastically for communities ranging from 1 to 1000 phage-bacteria pairs, and it included a fraction of escaping lysogens proportional to the increase in carrying capacity. The diversification model introduced new prophages at each diversification step and estimated the distribution of prophages per bacteria using combinatorics. The ecological model recovered the range of abundances and sublinear relationship between phage and bacteria observed across eleven ecosystems. The diversification model predicted an increase in the number of prophages per genome as bacterial abundances increased, in agreement with the distribution of prophages on 833 genomes from marine and human-associated bacteria. The study of lysogeny presented here offers a framework to interpret viral and microbial abundances and reconciles the Kill-the-Winner and Piggyback-the-Winner paradigms in viral ecology.

Introduction

The human gut contains one of the highest concentrations of bacteria and phages—viruses that infect bacteria—across ecosystems (Knowles et al. 2016; Wigington et al. 2016; Parikka et al. 2017). This high concentration of microbes is sustained by the daily supply of nutrient-rich compounds received from food intake and microbial metabolism (Blaut 2011; Cotillard et al. 2013; Mirzaei and Maurice 2017). In other ecosystems—mostly aquatic environments—an increase of resources has been linked to bacterial growth and phage lytic life cycle. This phage strategy produces new phage particles upon infection and subsequently bursts the bacterial host (lysis). Combined with the bacterial growth, the lytic life cycle ensures a rapid turnover of nutrients charateristic of the kill-the-winner (KtW) dynamics (Maurice et al. 2011; Brum et al. 2016; Thingstad and Lignell 1997). In the nutrient-rich and microbial dense gut ecosystem, thus, one would expect a similar phage lytic strategy.

Yet the lysogenic life cycle seems to be prevalent in the gut, where upon infection the phage genome integrates in the bacterial host as a prophage, forming a phage-bacteria symbiont called lysogen. Markers of lysogeny have been observed in viral genomic and metagenomic data from healthy adults (Furuse et al. 1983; Letarov and Kulikov 2009; Reyes et al. 2010; Minot et al. 2011, 2013; Mirzaei and Maurice 2017; Beller and Matthijnssens 2019; Shkoporov and Hill 2019). Besides, the frequency of prophages in bacteria is positively correlated with the bacterial growth rate (Lauro et al. 2009; Touchon et al. 2016). This empirical observation in single cells aligns with having an increase of lysogeny in productive environments like the gut (Mirzaei and Maurice 2017; Kim and Bae 2018). Additionally, the virus-to-microbe ratio (VMR) in the gut is significantly lower than in environments with lower microbial densities (Knowles, Silveira, et al 2016; Wigington et al. 2016; Parikka et al. 2017). The combination of the factors mentioned above have led to the piggyback-the-winner (PtW) dynamics, which considers lysogeny as a substantial ecological strategy for phages in high microbial density environments (Knowles et al. 2016; Silveira and Rohwer 2016).

The prevalence of lysogeny in gut environments contrast with observations from marine studies over the years, where lysogeny had been characterized as an ecological strategy of phages to survive situations at limiting growth conditions (Weinbauer 2004). Prophages integrate in poor growth conditions and remain dormant until resources improve. Then the lytic pathway activates the production of new phage particles (Jiang and Paul 1998; Wilcox and Fuhrman 1994; Payet and Suttle 2013; Maurice et al. 2009; Paul and Weinbauer 2010; Maurice et al. 2011; Brum et al. 2016). This paradox of lysogeny at low versus high productive environments might stem from having overlooked the postive effects of lysogenic conversion in the bacterial host.

Prophages provide multiple positive attributes to the host (Brüssow et al. 2004; Howard-Varona et al. 2017). They confer immunity to the host cell against similar and dissimilar phages—superinfection exclusion mechanism—without necessarily compromising the host's fitness in laboratory experiments (Bondy-Denomy et al. 2016; Mavrich and Hatfull 2019). They can also provide metabolic pathways that improve the host's competitive edge in different conditions (Bossi et al. 2003; Edlin et al. 1975, 1977), supply regulatory proteins (Paul 2008), and promote the expression of promiscuous bacterial enzymes (Hultqvist et al. 2018). Superinfection exclusion provided by prophages also favors the acquisition of transducing virions that kill sensitive bacteria, increasing the gene transfer to lysogens (Haaber et al. 2016; Touchon et al. 2017). This includes antibiotic resistance genes (Colavecchio et al. 2017). Prophages also enhance the adaptation and fitness of lysogens by encoding bacterial virulence factors implicated in animal infection and immune response evasion (Costa et al. 2018; Ohnishi et al. 2001; Hayashi et al. 2001; Waldor and Mekalanos 1996; Fortier and Sekulovic 2013). In addition, the domestication of prophages in bacteria removes the capability of generating new viral particles while conserving phage genes (Bobay et al. 2014), a process linked to the evolutionary acquisition of important bacteria elements such as gene transfer agents (Lang et al. 2012), bacteriocins, and type VI secretion systems (Michel-Briand and Baysse 2002; Leiman et al. 2009).

Our hypothesis is that these beneficial aspects of lysogeny are particularly favorable at high production and bacterial concentration conditions. We propose that as microbial productivity increases, ecosystems transition through recurrent lysogenic-lytic cycles, as illustrated in Figure 1. In an established microbial community with lytic phage turnover, an increase in resources would favor the emergence of lysogens due to the superinfection exclusion defense mechanism and the other advantages conferred by prophages, in consonance with the piggyback-the-winner dynamics (Knowles et al. 2016; Silveira and Rohwer 2016). If the increase in resources is sustained, by virtue of the kill-the-winner dynamics (Thingstad et al. 2014; Våge et al. 2018), adapted virulent and temperate phages that can prey on abundant lysogens will rise, diversifying the community and increasing its richness (Martiny et al. 2014). As the microbial productivity and concentration increases, this recurrent lysogenic-lysis cycle would lead to a prophage ratchet, predicting an increase in the frequency of prophages in bacteria (Figure 1).

To test this hypothesis, we developed an ecological mathematical model with lytic and lysogenic phage-bacterial community compartments. This ecological model was coupled with a diversification model that estimated the accumulation of prophages in bacterial genomes. Using physiological values for the viral-bacterial traits extracted from the literature, the ecological model recovered viral and bacterial abundances observed across eleven ecosystems, and predicted a relationship between bacterial abundance and richness—defined from a viral predation standpoint. The diversification model predicted distributions of prophage per bacteria that agreed with abundances of prophages estimated bioinformatically in bacteria from aquatic environments and the human gut. It also provided predictions for the other ecosystems studied, which will further test the PtW-KtW cyclic framework introduced here.

Methods

In this study, the terms phages and bacteria refer to the model. The terms viral-like-particles (VLPs) and microbial cells refer to environmental data. The terms viruses and microbes are used when discussing the model and environmental data.

Lytic community compartment. The lytic compartment was modeled using a one-to-one network of phage and bacteria species. A bacterial species was defined ecologically based on being a prey of a different phage, and, for each community, the total number, n, of phage-bacteria species pairs was fixed. Phages in this community adopted the lytic cycle, without distinguishing between temperate and virulent phages. The model aimed to estimate the total phage and bacterial abundances in a community rather than the rank abundance of different taxonomic groups. The phage and bacterial life traits were approximated to be homogeneous in each community, but different communities had assigned different physiological values, which were explored stochastically (see the sampling section below). The net rate of each species in the community was given by

$$\begin{aligned} \overbrace{\frac{\mathrm{d}B_i}{\mathrm{d}t}}^{\text{bacterial rate}} &= \overbrace{r\left(1-\frac{B}{K}\right)B_i}^{\text{logistic growth}} - \overbrace{dB_iP_i}^{\text{lytic infection}} \quad , \\ \overbrace{\frac{\mathrm{d}P_i}{\mathrm{d}t}}^{\text{phage rate}} &= \overbrace{cdB_iP_i}^{\text{lytic burst}} - \overbrace{mP_i}^{\text{phage decay}} \quad , \end{aligned} \tag{1}$$

Here, B_i denoted each species of sensitive bacterium, and P_i each species of lytic phage. The index i identified each species and ranged from 1 to n, leading to $2n$ coupled equations per lytic community. The bacterial net production rate was the balance between the bacterial growth rate and the infection rate. In the bacterial growth term, the intrinsic growth rate r was reduced by a logistic factor that accounted for the total concentration of bacteria in the lytic community, $B = \sum_{i=1}^{n} B_i$, with respect to the carrying capacity, K. The functional response of phages praying on bacteria can introduce richer dynamics (Weitz and Dushoff 2008), but it is not usually considered in phage ecological models (Thingstad and Lignell 1997; Weitz et al. 2015) due to the lack of environmental parametrization values and its potential variability across different ecosystems (Hunsicker et al. 2011). The phage-bacteria infection mechanism was encoded as a standard mass action term, which has been reported to have a similar infection rate constant in different environments (Thingstad et al. 2014; Barr et al. 2015; Joiner et al. 2019). This reduced the number of parameters in the model. The infection rate was assumed to be proportional to the product of bacterial and viral species concentrations, that is, mass action for each phage-bacterial host pair, where d represented the infection rate constant. The viral net production was the balance of the lysis (burst size, c, times the infection rate) and viral decay (with a decay constant m).

The coexistence equilibrium concentrations for each agent in the community, $B_i^\star$ and $P_i^\star$, were obtained by solving analytically the steady-state of Eq. (1), that is, $dB_i/dt = 0$ and $dP_i/dt = 0$. Coexistence was required for the n species in the community, that is, $B_i^\star > 0$ and $P_i^\star > 0$. The stability conditions were obtained analytically using a linear approximation around the equilibrium values (Strogatz 2015). The Jacobian of the dynamical system was obtained at the coexistence equilibrium, and the determinant of the characteristic polynomial was transformed until extracting analytical expressions for all eigenvalues. The derivation is detailed in the Supplementary Information (S.1.1).

Lysogenic community compartment. This community was formed by n temperate phage-lysogen pairs. Each lysogen had an active integrated prophage that could spontaneously induced, producing temperate phages. Thus this community was formed by n temperate phage-lysogen pairs. This compartment represented bacteria that had recently incorporated a prophage and was escaping the lytic top-down control in the community through the superinfection exclusion mechanism observed in laboratory experiments Brüssow et al. 2004. Except indicated otherwise, in the models the term lysogen refers to these escaping lysogenic community. The set of differential equations for the net production rate of escaping lysogens (L_i) and temperate phages (T_i) was given by

$$\begin{aligned}
\overbrace{\frac{\mathrm{d}L_i}{\mathrm{d}t}}^{\text{lysogenic rate}} &= \overbrace{r\left(1-\frac{L}{K_L}\right)Li}^{\text{logistic growth}} - \overbrace{\beta L_i}^{\text{prophage induction}}, \\
\overbrace{\frac{\mathrm{d}T_i}{\mathrm{d}t}}^{\text{phage rate}} &= \overbrace{c\beta L_i}^{\text{lytic burst}} - \overbrace{mT_i}^{\text{viral decay}} - \overbrace{\chi d T_i L_i}^{\text{superinfection exclusion}},
\end{aligned} \tag{2}$$

The index i identified each species and ranged from 1 to n, leading to $2n$ coupled equations as in the lytic community. The net production rate of lysogens was the balance between the lysogenic growth rate and the spontaneous prophage induction rate. The lysogenic growth was the intrinsic growth rate times a logistic term that accounted for the fraction of the total concentration of lysogens in the community, $L = \sum_{i=1}^{n} L_i$, with respect to the carrying capacity in the lysogenic compartment, K_L. The phage net production in the compartment was given by the production of temperate phages upon spontaneous prophage induction minus the phage decay and the removal of viable viruses due to superinfection exclusion. A value of $\chi = 1/2$ accounted for a mix of surface defense mechanism with reversible phage-host binding ($\chi = 0$ limit) and defense mechanism inactivating phage DNA ($\chi = 1$ limit) Jasien 2017.

The equilibrium concentrations, $L_i^\star$ and $T_i^\star$, were obtained analytically by solving the algebraic equations that satisfied steady-state, $dL_i/dt = 0$ and $dT_i/dt = 0$, for the coexistence regime, that is, $L_i^\star > 0$ and $T_i^\star > 0$ for the n species pairs. The stability conditions were obtained analytically using a linear approximation around the equilibrium values (Strogatz 2015). The Jacobian of the dynamical system was obtained at the coexistence equilibrium, and the determinant of the characteristic polynomial was transformed until extracting analytical expressions for all eigenvalues. The derivation is detailed in the Supplementary Information (S.1.2).

Lytic-lysogenic coupling. The dynamics of the lytic-lysogenic model described in Eqs. (1) and (2) was studied for single species with additional terms accounting for direct interaction Jasien 2017. The superinfection exclusion mechanism led to a dominance of the lysogenic community. This was in agreement with laboratory experiments Chaudhry et al. 2019, but led to a relatively low phage concentration as well as phage-to-bacteria ratio compared to environmental data Wigington et al. 2016; Knowles et al. 2016; Parikka et al. 2017. To circumvent this issue, the lytic and lysogenic communities were treated in separated compartments, and the lytic community was given a preferential treatment. A given ecosystem was assumed to have an established lytic community with n phage-bacteria pairs. Sustaining this community in equilibrium required a carrying capacity $K > B^\star$. The remaining carrying capacity, $\Delta K = (K - B^\star)$, was assumed to provide potential resources for the community of lysogens to be established by escaping escaping top-down control due to the superinfection exclusion mechanism Brüssow et al. 2004. The available resources were assumed to promote the formation of new lysogens, and as they increased in abundance it was assumed that phages would evolve to apply top-down control, by virtue of the KtW motif (Våge et al. 2018). The model presented here did not have the resolution to explore the arms race between this emerging lysogens and evolving phages. Instead, it assumed that the carrying capacity that escaping lysogens had access to grow, K_L, was proportional to the remaining carrying capacity from the lytic compartment,

$$K_L = f(n, B^\star)\Delta K \quad . \tag{3}$$

The *escaping lysogenic* factor, f, was established using physical arguments that led to

$$f(n, B^\star) = n\,\frac{B^\star}{B_{max}} \quad . \tag{4}$$

The arguments were as follows. First, it was assumed that new emerging lysogens would be more likely to form in a richer community. This was based on the lysogenic-lytic cycle hypothesis and the associated speciation through prophage integration. This lead to the proportionality term n in the escaping lysogenic factor above. Second, it was assumed that an increase in the concentration of top-down controlled bacteria in the lytic compartment would increase the encounter rate with temperate phages and promote the emergence of new lysogens. This led to the proportionality term $B^\star$. The escaping lysogenic factor must be unitless, so the contribution of the concentration needs to be normalized. This was done dividing the concentration $B^\star$ by the maximum packing concentration of bacteria $B_{max} \approx 10^{12}$ bacteria/mL. This was calculated by assuming a bacterial volume of 1 μm^3 (Milo and Phillips 2015). The total bacterial concentration was the sum of the bacteria in both compartments:

$$M^\star = B^\star + L^\star \ . \tag{5}$$

The same applied to the total phage concentration:

$$V^\star = P^\star + T^\star \ . \tag{6}$$

Sampling of coexisting equilibria. Reference values for the intrinsic growth rate (r), infection rate (d), burst size (c), and decay rate (m) were obtained from Thingstad 2000 and Weitz et al. 2017, which published models that used established empirical values. The induction rate of prophages (β) was estimated from phage lambda (Rokney et al. 2008) and phage P1 (Rosner 1972). The range of carrying capacities were extrapolated from the microbial concentrations observed across ecosystems (Knowles et al. 2016). The reference value for each parameter can be found in Table 1. To account for environmental variability, each parameter was investigated for a range that spanned an order of magnitude higher and lower with respect to the reference value. The combination of parameters was explored using latin-hypercube sampling (LHS) (McKay et al. 1979; Weitz et al. 2017). For each given community, that is, a fix number of species n, 2500 equilibria points were sampled in three different scenarios: purely lytic, purely lysogenic, and lytic-lysogenic community. These equilibria included coexisting and non-coexisting solutions. The fraction of coexisting equilibria differed for different communities and scenarios, so 500 coexisting equilibria were randomly selected in each community for analysis. This protocol was applied to communities ranging from $n = 1$ to 1000 phage-bacteria pairs.

In the LHS, each parameter defined an orthogonal dimension on a hypercube. The range of each parameter was divided into an evenly spaced number of regions equal to the number of points sampled (2500 per fixed n and community model). Each sampling contained a value from a unique region of each parameter, and the value within this region was selected randomly from a uniform distribution. No partition was used more than once throughout the sampling process. All parameters were sampled in linear space except for the carrying capacity, which spanned over seven orders of magnitude (from 10^4 to 10^{11} bacteria/ml) and was sampled in a logarithmic scale on base 10. For lytic communities with fixed richness, this sampling strategy was analogous to that introduced in Weitz et al. 2017.

A sensitivity analysis was applied to assess how the variation of the parameter values impacted the total concentration of bacteria (M) and phages (V). The analysis was performed around the mid-range reference value for the intrinsic growth rate (r), infection rate constant (d), decay rate constant (m), burst size (c), induction rate constant (β), superinfection exclusion factor (χ), richness (n), and carrying capacity (K). Technical details are provided in the Supplementary Information, Section S.4 (Sensitivity analysis).

A variant of the model incorporating a grazing effect on bacteria was also explored. Grazers were introduced as a generic predator pressure on the bacterial population by the term ($-gB_i$) in the population of bacteria. The grazer pressure rate constant, g, was assumed to be half of the bacterial replication rate $r/2$ (Fuhrman and Noble 1995; Thingstad and Lignell 1997). This term did not affect the results significantly. Grazers competed with phages, reducing their population abundance only in scenarios with relatively low carrying capacity. The main impact was the amount of percentage of sampled points that displayed stable coexistence, which dropped from 90% to 70%. To reduce the number of parameters in the model, the term was not included in the final version of the model.

VLPs and microbial cell concentrations. Virus-like particles (VLP) and microbial cell counts were extracted from 22 studies covering eleven ecosystems as reported previously in Knowles et al. 2016. The meta-analysis included data from animal (Barr et al. 2013; Furlan 2009; Kim et al. 2011), coastal/estuarine (Bettarel et al. 2006; Bouvier and Maurice 2011; Hewson et al. 2001; Maurice et al. 2011, 2013; Parsons et al. 2015; Schapira et al. 2009), coral reef (Payet et al. 2014), deep ocean (Muck et al. 2014), drinking water (Rinta-Kanto et al. 2004), open ocean (Parsons et al. 2012), polar lakes (Laybourn-Parry et al. 2007; Lisle and Priscu 2004; Madan et al. 2005), sediment (Bettarel et al. 2006; Glud and Middelboer 2004; Mei and Danovaro 2004; Patten et al. 2008), soil (Amosse et al. 2013), soil pore water (Amosse et al. 2013), and temperate lake/river (Bouvier and Maurice 2011; Maurice et al. 2010) ecosystems. The studies used either flow cytometry or eplifuorescence to estimate cell and viral counts. Epifluorescence was used in 14 out of 22 studies for cell density (84% of the cell data) and 19 out of 22 studies for viral density (4% of the VLP data). The data on VLP and cell counts is provided in Source File 1. The raw image data were not available. The counts were obtained from tables and figures in the original manuscripts. The combined dataset was analyzed using the non-parametric method of smooth splines (*smooth.spline*) with cross-validation implemented by B. D. Ripley and Martin Maechler in the R language for statistical computing (Chambers et al. 1992; R Core Team 2018). Linear, least-squares regressions were calculated in R for logged VLPs and logged microbial concentrations as well as logged virus-to-microbe ratios (VMRs) and logged microbial concentrations. The logged data was in base 10. This analysis was applied to each ecosystem and all ecosystems combined. For the combined ecosystem analysis, the median values were used in the linear regression to reduce oversampling bias from individual ecosystems.

Ecological model predictions. From the sampled communities in the model, linear statistical models were obtained for the richness, n, as a function of the median phage and bacterial concentrations using least-squares method. This statistical model was used to estimate the effective richness for each ecosystem studied using the median values for the VLPs and microbial cell abundances. For each ecosystem, the final richness was the average of the richness obtained from VLPs and microbial cell abundances. The percentage of bacteria escaping top-down control as lysogens was studied as a function of the community richness for the lytic-lysogenic model. A Hill function of order one (Michaelis-Menten) was fitted using a linear regression (least-squares method) on a double-reciprocal plot (Linewaver-Burk) to obtain a statistical model of the percentage of escaping lysogeny as a function of community richness. This statistical model was combined with the average richness obtained for each ecosystem to predict the percentage of the community escaping top-down control through lysogeny.

Diversification model for the distribution of prophages per bacteria. A combinatoric model was developed to estimate the number of prophages per bacteria accumulated in different ecosystems. The integration of a prophage and its associated phenotypic advantages to the lysogen represent a substantial genetic variant with respect the original bacterium. Thus, in the diversification model, the integration of a prophage defined the formation of a new bacterial strain. At each diversification step, a newly evolved temperate phage was introduced in the community. Its integration as a prophage defined new bacterial strains. In the first evolutionary step ($z = 1$), the model contained a bacterial community composed of a single bacterial strain ($s = 1$) free of prophages ($q = 0$). In step two ($z = 2$), a temperate phage was introduced. Its integration as a prophage on an undetermined fraction of the initial bacaterial strain formed a community with two strains ($s = 2$): the original free-prophage bacterial strain ($q = 0$) and a new strain containing one prophage ($q = 1$). In the third diversification step ($z = 3$), a newly evolved temperate phage was added to the community. Its integration as a prophage promoted the formation of new strains. This led to three total strains ($s = 3$) and two alternative scenarios or states: in state 1, the prophage integrated in a fraction of the free prophage strain, leading to a community with a free-prophage strain ($q = 0$) and two strains with a different prophage each (two $q = 1$ strains); in state 2, the prophage integrated in a fraction of the single-prophage strain from the prior diversification step, leading to a community with a free-prophage strain ($q = 0$), a one-prophage strain ($q = 1$), and a newly formed two-prophage strain ($q = 2$). This diversification process of generating strains was repeated subsequently for an arbitrary number of diversification steps, z, containing at each step $s = z$ number of strains per state. A recursion relationship was derived to calculate the probability of having q prophages per bacteria in a system with z strains, $P(z, q)$.

Distribution of prophage per bacteria from low and high productive environments. The environmental distribution of prophages per bacteria was assessed from a low (marine surface) and high (human gut) productive environments. Phispy (version 3.2) (Akhter et al. 2012) was ran against a curated database of bacterial genomes from the Pathosystems Resource Integration Center (PATRIC). The process of curation focused on adding environmental annotation and removing plasmids, incomplete genomes, and metagenomes. A total of 83 genomes from marine surface bacteria and 750 from human gut were analyzed. Prophage genes were identified based on sequence similarity and genomic signatures as described previously (Akhter et al, 2012). A group of 30 or more prophage genes in a genome defined a single prophage in a bacterial genome. See Source Files 2 and 3.

Diversification model predictions. The average number of prophages per bacteria obtained in marine surface and human gut environments was combined with the estimated richness from marine and animal gut communities to calibrate the prophage diversification and lytic-lysogenic community models. This facilitated the prediction of prophage frequency per bacteria in different ecosystems. The theoretical number of strains or diversification steps, z, was estimated by obtaining the theoretical average of prophages per bacteria, $< q(z) >= \sum_{q=0}^{q=z-1} qP(z, q)$, nearest to the average number of prophages obtained, respectively, for marine surface and human gut bacteria. The two theoretical diversification steps, z, were related with the effective number, n, of phage-bacteria pairs in marine surface and human gut assuming a power function model, $z(n) = an^b$. This determined the values for a and b. The parameters obtained were $a = 27.96$ and $b = 2.04$. This power function model was applied to estimate the number of strains (diversification steps), z, as a function of the community richness, n, for each ecosystem studied. The number of strains (or steps) obtained determined the probability of number of prophages per bacteria, $P(z, q)$, for the eleven ecosystems studied.

Results

In the lytic community, Eq. (1), the equilibrium concentrations for each species of bacteria and phage were respectively, $B_i^\star = m/(c\,d)$ and $P_i^\star = (r/d)(1 - B^\star/K)$, where i ranged from 1 to the total number of species in the community n. The total bacterial and phage equilibrium concentrations were, respectively,

$$B^\star = n\,\frac{m}{cd} \tag{7}$$

and

$$P^\star = n\,\frac{r}{d}\left(1 - \frac{B^\star}{K}\right) \ . \tag{8}$$

The total bacterial concentration, $B^\star$, was determined by the phage properties, that is, phage decay rate m, burst size c, and infection rate d. This was due to the top-down control applied by lytic phages. The bacterial population was also proportional to the richness of phage-bacteria pairs in the community n. The total phage concentration, $P^\star$, was a function of the bacterial growth rate r, the infection rate d, and the fraction of underutilized resources in the lytic community $(1 - B^\star/K)$. The total phage concentration had a quadratic dependence on the phage-bacterial host pair richness, n. The community richness that maximized the concentration of phages was $n_{P_{max}} = 2Km/(c\,d)$. The analytical expressions for the $2n - 2$ eigenvalues derived from the linear stability analysis indicated that the total phage and bacteria concentrations were stable for carrying capacities larger than the total bacterial population, $K > B^\star$ (see Supplementary section S.1.1). The individual species displayed an oscillatory behavior with an angular frequency $\omega = \sqrt{rm(1 - B^\star/K)}$. This was a consequence of having multiple one-to-one phage-bacteria pairs sharing the same resources, that is, carrying capacity, K.

In the lysogenic community, Eq. (2), the equilibrium distribution of lysogens was determined by the initial concentrations of lysogenic species, $L_i(t)/L(t) = L_i(0)/L(0)$. The total concentration of lysogens was

$$L^\star = \left(1 - \frac{\beta}{r}\right) K_L \ . \tag{9}$$

The total concentration of lysogens was proportional to the carrying capacity in the lysogenic compartment, K_L, and it was reduced by the fraction of spontaneous prophage induction per replication (β/r). Thus, contrary to bacteria in the lytic compartment, the lysogenic community displayed a bottom-up control. This was a consequence of the superinfection exclusion mechanism, which removed the lytic predator pressure on the lysogens. The concentration of temperate phage species in the lysogenic community was $T_i^\star = c\beta L_i^\star/(m + \chi d L_i^\star)$. This led to a total concentration of temperate phages, $T^\star$, as a function of the distribution of lysogenic species. But the concentration of temperate phages generated from the lysogenic community had a small contribution to the pool of phages due to the small spontaneous induction rate. Two extreme cases illustrate this finding. A homogeneous distribution of lysogens, $L_i^\star = L_j^\star$, led to the total temperate phage concentration $T^\star = c\beta L^\star/(m\,n + \chi d L^\star)$. A community dominated by a single lysogenic species, $L_1^\star \approx L^\star$ and $L_{i>1}^\star \approx 0$, led to $T^\star = c\beta L^\star/(m + \chi d L^\star)$. Both expressions were Hill functions of order one with respect the lysogenic community, reaching the same plateau as the population of lysogens increased, regardless of the rank abundance of lysogenic species in the community,

$$T^\star \approx \frac{c\beta}{\chi d} \ . \tag{10}$$

For the reference values given in Table 1, this was $T^\star \sim 2 \cdot 10^2$ phage/ml. Thus, the contribution of temperate phages from the active lysogenic community was small, and the choice of the specific distribution of the lysogenic community did not affect the results of the model significantly. The stability analysis was performed for the most general case, that is, with no specific distribution of lysogens (see Supplementary section S.1.2). There was n negative eigenvalues and $n-1$ null eigenvalues, and the total concentration of lysogens remain stable for spontatenous induction rates smaller than the bacterial growth rate, $\beta < r$, which was the case for the empirical values, Table 1.

In the lytic-lysogenic community model, the total bacterial ($M^\star$) and phage concentrations ($V^\star$), which combined the lytic and lysogenic communities sampled stochastically, covered, respectively, the range 10^4–10^{10} bacteria/ml and 10^5–10^{10} phages/ml, in consonance with the ranged observed across the eleven ecosystems studied (Figures 2**a** and 2**b**). The phage concentration increased sublinearly as a function of the bacterial concentration with an exponent $\alpha = 0.79 \pm 0.04$ (SE) ($\mathrm{R}^2 = 0.99$, $\mathrm{p} < 0.001$, $n_p = 7$), in agreement with the exponent observed environmentally ($\alpha = 0.84 \pm 0.11$ (SE), $\mathrm{R}^2 = 0.86$, $\mathrm{p} < 0.001$, $n_p = 11$). This power-function relating viral and microbial abundances was a first order approximation. A more refined statistical analyses presented in the Supplementary Information indicated that the exponent of the viral-microbial relationship becomes smaller with microbial concentration for both the model and the environmental data (Figures S.1, S.2, and S.3).

The phage-to-bacteria ratios (PBR) in the model ranged from 10^2 to 10^{-2} phage/bacterium (Figures 2**c**). This aligned with thevirus-to-microbe ratios (VMR) observed environmentally (Figure 2**d**). PBR decreased as a function of bacterial concentration following a power function with exponent $\alpha_{PBR} = -0.21 \pm 0.04$ (SE) ($\mathrm{R}^2 = 0.83$, $\mathrm{p} < 0.001$, $n_p = 7$), which was consistent with the trend of environmental virus-to-microbe ratios ($\alpha_{VMR} = -0.16 \pm 0.11$ (SE), $\mathrm{R}^2 = 0.86$, $\mathrm{p} < 0.001$). As expected, the trend in PBR (or VMR) was related with the phage-bacteria (or viral-microbial) exponent by $\alpha_{PBR} \approx \alpha - 1$, since $\mathrm{PBR} = V/M$. In communities with fixed richness, the stochastic sampling produces locally a sublinear trend that is intensified as lysogeny increases (Figure S.1). It is worth noting that no parameters were fitted in the model, except to the values used for the minimum and maximum richness (n).

The sensitivity analysis indicated that richness (n) was the most relevant parameter in the model, with a sensitivity factor $S_{M,n} \approx 1.5$ in the bacteria population and $S_{V,n} \approx 1$ in the phage population (Figure S.6). This was followed by the infection rate constant (d), which displayed a sensitivity factor $S \approx 1$ in both populations. The decay rate (m) and burst size (c) (both phage traits) where also sensitive in the bacteria population (sensitive factor $S \approx 1$), while the growth rate affected the phage population (sensitive factor $S \approx 1$). The carrying capacity was the last relevant parameter in the bacteria population (sensitive factor $S \approx 0.5$). The other parameters did not have a relevant impact in the community (sensitivity factor, $S \approx 0$). For further details, see Supplementary Information S.4.

The purely lytic community model, lacking the lysogenic compartment, recovered similar ranges of phage and bacteria concentrations (Figure S.5**a**). As in the case of the lytic-lysogenic community model, this was due to the variation of richness, n, across communities. The phage concentration followed a linear relationship with the microbial concentration ($\alpha = 1.00 \pm 0.01$ (SE), $\mathrm{R}^2 = 1.00$, $\mathrm{p} < 0.001$). Accordingly, the average phage-to-bacteria (PBR) ratio was constant, PBR ≈ 10 (Figure S.5**b**), that is, it did not vary with respect the microbial concentration ($\alpha_{PBR} = 0.00 \pm 0.01$ (SE), $\mathrm{R}^2 = 0.02$, $\mathrm{p} = 0.736$), in disagreement with the trend observed environmentally for the virus-to-microbe ratio (Figure 2**b**). Lysogeny, thus, was necessary in the model to recover a sublinear relationship between phage and bacterial concentrations across ecosystems.

The percentage of lysogeny, $L[\%]$ in the lytic-lysogenic community increased as a function of the total bacterial concentration, $M^\star$, saturating near 10^{10} bacteria/ml (Figure 3**a**). The parameters obtained for the fitted Hill function of order one were $a = 97.1 \pm 0.7$ and $b = (2.86 \pm 8) \cdot 10^8$ bacteria/ml (see details in the caption of Figure 3). The model indicated that the median percentage of lysogeny increased abruptly at bacterial concentrations larger than 10^6 bacteria/mL. In particular, the median percentage of lysogeny was 1% at $\sim 3.0 \cdot 10^6$ bacteria/mL, 10% at $\sim 3.3 \cdot 10^7$ bacteria/mL, 50% at $\sim 3.0 \cdot 10^8$ bacteria/mL, and 90% at $\sim 3.6 \cdot 10^9$ bacteria/mL.

To facilitate the comparison with empirical data, the estimated phage-bacteria pair richness obtained for each ecosystem was expressed relative to the ecosystem with the lowest predicted richness, that is, deep ocean (Figure 4**a**). The relative richness obtained across ecosystems was $\sim$ 1 (deep ocean), $\sim$ 7 (open ocean), $\sim$ 12 (soil pore water) , $\sim$ 16 (drinking water), $\sim$ 18 (coral reef), $\sim$ 64 (polar lakes), $\sim$ 85 (temperate lakes), $\sim$ 88 (coastal estuaries), $\sim$ 104 (soil), $\sim$ 248 (animal), and $\sim$ 978 (sediment). The phage-bacteria host pair richness of a community increased with the microbial and viral concentrations in ecosystems (Figure 2**b**).

The percentage of escaping lysogens predicted across ecosystems was (Figure 4**b**) ∼0.0073% (deep ocean), ∼0.11% (open ocean), ∼0.15% (soil pore water) , ∼0.20% (drinking water), ∼0.16% (coral reef), ∼0.37% (polar lakes), ∼0.78% (temperate lakes), ∼1.5% (coastal estuaries), ∼0.5% (soil), ∼3.7% (animal), and ∼27% (sediment). The percentage of escaping lysogeny increased with richness, but some ecosystems, such as coral reefs and soils, displayed a relatively lower lysogeny with respect the contiguous ecosystems. This was due to their higher virus-to-microbe ratios.

In the diversification model, the number of bacterial strains with q prophages after z evolution steps (strains in the system) was related to the number of strains containing $q-1$ prophages in the prior evolution step plus $(z-1)!$ times the number of bacterial strains having q in the prior step, that is, $R(z, q) = R(z-1, q-1) + (z-1)!R(z-1, q)$. The total number of strains in the ensamble at step z was $z!$ The probability of having q prophages per bacteria in a given evolutionary step was

$$P(z, q) = \frac{R(z, q)}{z!} = \frac{1}{z}P(z-1, q-1) + \frac{z-1}{z}P(z-1, q) \ . \tag{11}$$

This determined the theoretical distribution of prophages per bacteria as a function of the total number of strains (evolutionary steps) z.

The empirical distributions of prophage per bacteria were obtained for superficial sea water (5 ± 2 (SD) prophage/bacteria, 83 genomes) and human gut (14 ± 5 (SD) prophage/bacteria, 750 genomes) (Figure 5**a**). The theoretical distributions were centered by fitting the prophage evolution generation number (z) to recover the average number of prophage per bacteria obtained empirically, leading to $z_1 = 182$ (marine surface) and $z_2 = 1.41 \cdot 10^6$ (human gut). The theoretical distributions of prophage per bacteria recovered the same distribution shape of the empirical data. By applying the effective richness (n) derived from the ecology model, the evolutionary model predicted the distributions of prophage per bacteria for each ecosystem (Figure 5**b**). The average number of prophages (± SD) per bacteria predicted for each ecosystem was (Figure 5**c**): open ocean (3.1 ± 1.6), soil pore water (4.0 ± 1.8), drinking water (5 ± 2), coral reef (5 ± 2), temperate lakes/rivers (8 ± 3), polar lakes (8 ± 3), coastal/estuarine (8 ± 3), soil (9 ± 3), animal (10 ± 3), and sediment (13 ± 4).

Finally, the lytic-lysogenic community model was then reinterpreted using an effective single lytic virus-host community with variable burst size. This aimed to faciliate the comparison of the lytic-lysogenic community model with the original PtW model (Knowles et al. 2016), which treated lysogeny implicitely in a single phage-bacteria species system by reducing burst size as a function of eutrophic conditions. The equilibrium values and stability of the lytic compartment were recovered by renormalizing the infection rate by the underlying community richness: $d \to d/n$ (see Eq. S.17). Note that the absence of the explicit phage-bacteria network removed the oscillatory modes obtained in the explicit lytic community. The lysogenic compartment was absorbed by the effective burst size $c_{eff} = (d/n)/m(B^\star + L^\star)$. This led to an effective burst size that decreased with bacterial concentration with a power law function with exponent $\alpha = -0.19 \pm 0.04$ (SE) ($R^2 = 0.82$, $p < 0.001$) (Figure 3**b**).

Discussion

Over decades, lysogeny has been interpreted as an ecological strategy for phages to find refuge in their bacterial hosts during starvation and low production conditions (Weinbauer 2004; Paul and Weinbauer 2010; Maurice et al. 2009; Brum et al. 2016). The human gut and other host-associated microbiomes, however, are dominated by lysogenic bacteria and temperate phages despite being highly productive and microbial-dense environments (Minot et al. 2011; Reyes et al. 2010; Minot et al. 2013; Mirzaei and Maurice 2017; Kim and Bae 2018). To explain this paradox, here we hypothesized that the benefits acquired by the bacterial host during phage integration favor the emergence of lysogens at high microbial abundances. It is propose that lysogens escape the lytic top-down control of phages, and eventually phages evolve to lytically infect the escaping lysogens, forming a lysogeny-lysis cycle. The recurrence of this lysogeny and lysis cycle as resources increase lead to a diversification of the viral and microbial communities (Figure 1). This PtW-KtW cyclic hypothesis was tested by coupling a lytic-lysogenic ecological model and a prophage acquisition evolutionary model.

In the ecological model, lysogeny increased as a function of resources, density, and richness (Figure 3). The model recovered microbial and viral concentrations observed empirically across eleven ecosystems, ranging from 10^4 cells/ml and 10^5 phages/ml to 10^{10} cells/ml and 10^{10} phages/ml (Wigington et al. 2016; Knowles et al. 2016; Parikka et al. 2017). It also recovered the empirical sublinear trend between viral and microbial concentrations (exponent ~ 0.7). In the model, the fraction of lysogens increased from 1% at $\sim 10^6$ bacteria/mL to 90% at $\sim 10^9$ bacteria/mL. This was a consequence of lysogenic bacteria escaping top-down control through superinfection exclusion (Brüssow et al. 2004; Touchon et al. 2017). Thus, in environments with $\sim 10^6$ bacteria/mL, the signal associated to lysogeny is predicted to be weak, in consonance with metagenomic observations in marine systems (Knowles et al. 2016; Luo et al. 2020). In environments with higher microbial density, instead, the lysogeny signal is predicted to be strong, a result consistent with observations in animal gut (Minot et al. 2011; Reyes et al. 2010; Minot et al. 2013; Mirzaei and Maurice 2017; Kim and Bae 2018).

Richness was defined as the number of phage-bacterial host pairs in the community richness, n. The increase in richness was associated to an increase of phage and bacteria concentrations. This is consistent with genomic data across gradients of microbial abundance, showing an increase of phage-host pair diversification (Coutinho et al. 2019). The purely lytic community model recovered similar ranges of phage and bacterial concentrations (Figure S.5a). In this scenario, however, the viral concentration across communities with different richness displayed a linear trend (exponent ~ 1) with a constant average virus-to-microbe ratio of ~ 10 (Figure S.5b)—in disagreement with the sublinear trend observed environmentally (Figure 2). For communities with fixed richness, however, the lytic model displayed a sublinear trend. A prior lytic model with fixed richness showed that this is a consequence of the stochastic sampling of the phage-bacteria traits (Weitz et al. 2017). The purely lytic community model, thus, displays a Simpson's paradox, where the local sublinear trend observed in communities with fixed richness differs from the global trend when comparing communities across richness (Simpson 1951; Knowles et al. 2017). This paradox is a consequence of both viral and microbial concentrations scaling linearly with richness, Eqs. (8) and (7).

At each step of the diversification model, an evolved temperate phage was assumed to be able to integrate as a prophage, defining new bacterial strains in the community. The average number of prophages per bacteria increased with the number of diversification cycles, in consonance with the distributions of prophages observed in the genomes of marine and human gut bacteria (Figure 5a). The calibration of the model using the mean prophage number per bacteria in these two ecosystems predicted 10,000 times more PtW-KtW cycles in the human gut than in marine surface communities, consistent with the increase in the frequency of prophages as a function of the environment productivity (Lauro et al. 2009; Lee and Patterson 2002; Bakenhus et al. 2017; Touchon et al. 2016). The microbial community colonization of the human gut in early childhood is extremely dynamic, showing sequential changes in dominant phage and bacteria (Breitbart et al. 2008; Sharon et al. 2013; Lim et al. 2015; Mirzaei and Maurice 2017). The PtW-KtW cycles introduced here would eventually lead to the stable, richer, and lysogenic communities observed in adulthood (Minot et al. 2013; Reyes et al. 2010; Shkoporov and Hill 2019). The predicted distribution of prophages per bacteria in the other ecosystems studied can be tested emperically in future work (Figures 5b and 5c).

The accuracy of the model could be improved by adding the molecular mechanisms leading to the formation of new lysogens, which has been modeled at single strain-level but proven hard to incorporate into ecological scenarios (Steward and Levin 1984; Wang and Goldenfeld 2010; Maslov and Sneppen 2017; Wahl et al. 2018; Weitz et al. 2019; Chaudhry et al. 2019). The accuracy could be also improved by incorporating by incorporating specific predator and immune system pressure (Thingstad and Lignell 1997; Thingstad 2000; Winter et al. 2010; Dwayne et al. 2017; Caron et al. 2017; Talmy et al. 2019), variable energy sources (Thingstad and Lignell 1997; Weitz et al. 2015), and ecosystem-specific ranges for phage-bacteria traits and nested viral-microbial networks (Flores et al. 2011; Thingstad et al. 2014; Gao et al. 2017; Hendricks 1972; Kirchman 2016; De Paepe and Taddei 2006).

Conclusion

The ecological-evolutionary community model introduced here assumed a recurrent cycle of lysogeny and lysis. Productive microbial conditions promoted lysogeny (piggyback-the-winner) and the selection of phages able to control lysogens protected by superinfection exclusion (kill-the-winner). The model recovered viral and microbial concentrations observed across 11 ecosystems and predicted an increase in the frequency of prophages in bacteria, consistent with observations in marine ecosystems and the mammalian gut. The prophage-acquisition ratchet explains the observation of high lysogeny in rich microbial environments.

References

S. Akhter, R.K. Aziz, and R.A. Edwards. Phispy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic Acids Res.*, 40:e126, 2012.

J Amosse, Y. Bettarel, C. Bouvier, T. Bouvier, T.T. Duc, T.D. Thu, and P. Jouquet. The flows of nitrogen, bacteria and viruses from the soil to water compartments are influenced by earthworm activity and organic fertilization (compost vs. vermicompost). *Soil Biol. Biochem.*, 66:197–203, 2013.

Insa Bakenhus, Leon Dlugosch, Sara Billerbeck, Helge-Ansgar Giebel, Felix Milke, and Meinhard Simon. Composition of total and cell-proliferating bacterioplankton community in early summer in the north sea–roseobacters are the most active component. *Frontiers in microbiology*, 8:1771, 2017.

J.J. Barr, R. Auro, M. Furlan, K.L. Whiteson, L.M. Erb, J. Pogliano, A. Stotland, R. Wolkowicz, A.S. Cutting, K.S. Doran, P. Salamon, M. Youle, and F. Rohwer. Bacteriophage adhering to mucus provide a non–host-derived immunity. *Proc. Natl. Acad. Sci. U.S.A.*, 110:10771–10776, 2013.

J.J. Barr, R. Auro, N. Sam-Soon, S. Kassegne, G. Peters, N. Bonilla, M. Hatay, S. Mourtada, B. Bailey, M. Youle, B. Felts, A. Baljon, J. Nulton, P. Salamon, and F. Rohwer. Subdiffusive motion of bacteriophage in mucosal surfaces increases the frequency of bacterial encounters. *Proc. Natl. Acad. Sci. U.S.A.*, 112:13675–13680, 2015.

L. Beller and J. Matthijnssens. What is (not) known about the dynamics of the human gut virome in health and disease. *Current opinion in virology*, 37:52–57, 2019.

Y. Bettarel, M. Bouvy, C. Dumont, and T. Sime-Ngando. Virus-bacterium interactions in water and sediment of west african inland aquatic systems. *Appl. Environ. Microbiol.*, 72:5274—-5282, 2006.

M. Blaut. Ecology and physiology of the intestinal tract. In *Between Pathogenicity and Commensalism*, pages 247–272. Springer, 2011.

L.-M. Bobay, M. Touchon, and E.P.C. Rocha. Pervasive domestication of defective prophages by bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, 111:12127–12132, 2014.

J. Bondy-Denomy, J. Qian, E. R. Westra, A. Buckling, D. S. Guttman, A. R. Davidson, and K. L. Maxwell. Prophages mediate defense against phage infection through diverse mechanisms. *ISME J.*, 10:2854–2866, 2016.

L. Bossi, J.A. Fuentes, G. Mora, and N. Figueroa-Bossi. Prophage contribution to bacterial population dynamics. *J. Bacteriol.*, 185:6467–6471, 2003.

T. Bouvier and C.F.A. Maurice. A single-cell analysis of virioplankton adsorption, infection, and intracellular abundance in different bacterioplankton physiologic categories. *Microb. Ecol.*, 62: 669—-678, 2011.

M. Breitbart, M. Haynes, S. Kelley, F. Angly, R. A Edwards, B. Felts, J.M. Mahaffy, J. Mueller, J. Nulton, S. Rayhawk, B. Rodriguez-Brito, P. Salamon, and F. Rohwer. Viral diversity and dynamics in an infant gut. *Res. Microbiol.*, 159(5):367–373, 2008.

J.R. Brum, B.L. Hurwitz, O. Schofield, H.W. Ducklow, and M.B. Sullivan. Seasonal time bombs: dominant temperate viruses affect southern ocean microbial dynamics. *ISME J*, 10:437, 2016.

H. Brüssow, C. Canchaya, and W.-D. Hardt. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Ref.*, 68:560–602, 2004.

D.A. Caron, H. Alexander, A.E. Allen, J.M. Archibald, E.V. Armbrust, C. Bachy, C.J. Bell, A. Bharti, S.T. Dyhrman, S.M. Guida, K.B. Heidelberg, J.Z. Kaye, J. Metzner, S.R. Smith, and A.Z. Worden. Probing the evolution, ecology and physiology of marine protists using transcriptomics. *Nat. Rev. Microbiol.*, 15:6–20, 2017.

John M Chambers, Trevor J Hastie, et al. *Statistical models in S*, volume 251. Wadsworth & Brooks/Cole Advanced Books & Software Pacific Grove, CA, 1992.

W. Chaudhry, N. Vega, A.R. Govindan, R. Garcia, E. Lee, I.C. McCall, and B.R. Levin. The population and evolutionary dynamics of bacteriophage: Why be temperate revisited. *bioRxiv*, page 824235, 2019.

M.J. Claesson, O. O'Sullivan, Q. Wang, J. Nikkil´o, J.R. Marchesi, H. Smidt, W.M. de Vos, R. Paul Ross, and P.W. O'Toole. Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS One*, 4:e6669, 2009.

A. Colavecchio, B. Cadieux, A. Lo, and L.D. Goodridge. Bacteriophages contribute to the spread of antibiotic resistance genes among foodborne pathogens of the enterobacteriaceae family – a review. *Front. Microbiol.*, 8:1108, 2017.

J.K. Cornuault, E. Moncaut, V. Loux, A. Mathieu, H. Sokol, M. Petit, and M. De Paepe. The enemy from within: a prophage of roseburia intestinalis systematically turns lytic in the mouse gut, driving bacterial adaptation by crispr spacer acquisition. *ISME J.*, 14(3):771–787, 2020.

A.R. Costa, R. Monteiro, and J. Azeredo. Genomic analysis of acinetobacter baumannii prophages reveals remarkable diversity and suggests profound impact on bacterial virulence and fitness. *Sci. Rep.*, 8:15346, 2018.

A. Cotillard, S.P. Kennedy, L.C. Kong, E. Prifti, N. Pons, E. Le Chatelier, M. Almeida, B. Quinquis, F. Levenez, N. Galleron, et al. Dietary intervention impact on gut microbial gene richness. *Nature*, 500(7464):585–588, 2013.

F.H. Coutinho, R. Rosselli, and F. Rodríguez-Valera. Trends of microdiversity reveal depth-dependent evolutionary strategies of viruses in the mediterranean. *mSystems*, 4(6), 2019.

M. De Paepe and F. Taddei. Viruses' life history: towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS Biol*, 4:e193, 2006.

R.R. Dwayne, C.Y. Leung, M. Henry, E. Morello, D. Singh, J.P. Di Santo, J.S. Weitz, and L. Debarbieux. Synergy between the host immune system and bacteriophage is essential for successful phage therapy against an acute respiratory pathogen. *Cell Host Microbe*, 22:38–47, 2017.

G. Edlin, L. Lin, and R. Kudrna. Lambda lysogens of *E. coli* reproduce more rapidly than non-lysogens. *Nature*, 1975:735–737, 1975.

G. Edlin, L. Lin, and R. Bitner. Reproductive fitness of p1, p2, and mu lysogens of *Escherichia coli*. *J. Virol.*, 1977:560–564, 1977.

C. O. Flores, J. R. Meyer, S. Valverde, L. Farr, and J. S. Weitz. Statistical structure of host-phage interactions. *Proc. Natl. Acad. Sci. U.S.A.*, 108:11309–11310, 2011.

L.-C. Fortier and O. Sekulovic. Importance of prophages to evolution and virulence of bacterial pathogens. *Virulence*, 4:354–365, 2013.

J.A. Fuhrman and R.T. Noble. Viruses and protists cause similar bacterial mortality in coastal seawater. *Limnol. Oceanogr.*, 40:1236–1242, 1995.

M. Furlan. Viral and microbial dynamics in the human respiratory tract. Master's thesis, San Diego State University, USA, 2009.

K. Furuse, S. Osawa, J. Kawashiro, R. Tanaka, A. Ozawa, S. Sawamura, Y. Yanagawa, T. Nagao, and I. Watanabe. Bacteriophage distribution in human faeces: continuous survey of healthy subjects and patients with internal and leukaemic diseases. *Journal of General Virology*, 64(9): 2039–2043, 1983.

N.L. Gao, C. Zhang, Z. Zhang, S. Hu, M.J. Lercher, X.-M. Zhao, P. Bork, Z. Liu, and W.-H. Chen. MVP: a microbe–phage interaction database. *Nucleic Acids Res.*, 46:D700–D707, 2017.

R.N. Glud and M. Middelboer. Virus and bacteria dynamics of a coastal sediment: implication for benthic carbon cycling. *Limnol. Oceanogr.*, 49:2073—-2081, 2004.

A. Golomidova, E. Kulikov, A. Isaeva, A. Manykin, and A. Letarov. The diversity of coliphages and coliforms in horse feces reveals a complex pattern of ecological interactions. *Appl. Environ. Microbiol.*, 73(19):5975–5981, 2007.

J. Haaber, J.J. Leisner, M.T. Cohn, A. Catalan-Moreno, J.B. Nielsen, H. Westh, J.R. Penadés, and H. Ingmer. Bacterial viruses enable their host to acquire antibiotic resistance genes from neighbouring cells. *Nature communications*, 7(1):1–8, 2016.

T. Hayashi, K. Makino, M. Ohnishi, K. Kurokawa, K. Ishii, K. Yokoyama, C.G. Han, E. Ohtsubo, K. Nakayama, T. Murata, M. Tanaka, T. Tobe, T. Iida, H. Takami, T. Honda, C. Sasakawa, N. Ogasawara, T. Yasunaga, S. Kuhara, T. Shiba, M. Hattori, and H. Shinagawa. Complete genome sequence of enterohemorrhagic escherichia coli o157:h7 and genomic comparison with a laboratory strain k-12. *DNA Res.*, 9:11–22, 2001.

C.W. Hendricks. Enteric bacterial growth rates in river water. *Appl. Environ. Microbiol.*, 24: 168–174, 1972.

I. Hewson, J.M. O'Neil, J.A. Fuhrman, and W.C. Dennison. Virus-like particle distribution and abundance in sediments and overlying waters along eutrophication gradients in two subtropical estuaries. *Limnol. Oceanogr.*, 46:1734—-1746, 2001.

C. Howard-Varona, K.R. Hargreaves, S.T. Abedon, and M.B. Sullivan. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *ISME J.*, 11:1511–1520, 2017.

J.J. Hultqvist, O. Warsi, A. Soderholm, M. Knopp, U. Eckhard, E. Vorontsov, M. Selmer, and D.I. Andersson. A bacteriophage enzyme induces bacterial metabolic perturbation that confers a novel promiscuous function. *Nat. Ecol. Evol.*, doi: https://doi.org/10.1038/s41559-018-0568-5, 2018.

M.E. Hunsicker, L Ciannelli, K.M. Bailey, J.A. Buckel, J. Wilson White, J.S. Link, T.E. Essington, S. Gaichas, T.W. Anderson, R.D. Brodeur, K.S. Chan, K. Chen, G. Englund, K.T. Frank, V. Freitas, M.A. Hixon, T. Hurst, D.W. Johnson, J.F. Kitchell, D. Reese, G.A. Rose, H. Sjodin, W.J. Sydeman, H.W. van der Veer, K. Vollset, and S. Zador. Functional responses and scaling in predator-prey interactions of marine fishes: contemporary issues and emerging concepts. *Ecol. Lett.*, 14:1288–1299, 2011.

E. Jasien. Modeling the transition between lytic and lysogenic behavior in a bacteria-phage system. Master's thesis, San Diego State University, USA, 2017.

S.C. Jiang and J.H. Paul. Significance of lysogeny in the marine environment: Studies with isolates and a model of lysogenic phage production. *Microbial Ecology*, 35:235–243, 1998.

K.L. Joiner, A. Baljon, J. Barr, F. Rohwer, and A. Luque. Impact of bacteria motility in the encounter rates with bacteriophage in mucus. *Sci. Rep.*, 9:16427, 2019.

M. S. Kim and J. W. Bae. Lysogeny is prevalent and widely distributed in the murine gut microbiota. *ISME J.*, 12:1127–1141, 2018.

M.-S. Kim, E.-J. Park, S.W. Roh, and J.-W. Bae. Diversity and abundance of single-stranded dna viruses in human feces. *Environ. Microbiol.*, 77:8062—-8070, 2011.

D.L. Kirchman. Growth rates of microbes in the oceans. *Ann. Rev. Mar. Sci.*, 8:285–309, 2016.

B. Knowles, C. B. Silveira, K. Barrot, V. A. Cantu, A. G. Cobián-Güemes, F. H. Coutinho, E. A. Dinsdale, B. A. Felts, K. A. Furby, E. E. George, K. T. Green, G. B. Gregoracci, A. F. Haas, J. M. Haggerty, E. R. Hester, N. Hisakawa, L. W. Kelly, Y. W. Lim, M. Little, T. Luque, A. McDole-Somera, L. S. McNair, K. de Oliveira, S. D. Quistad, Robinett N. L., Sala E., Salamon P., Sanchez S. E., Sandin E., Silva G. G. Z., Smith J., Sullivan C., Thompson C., M. J. A. Vermeij, M. Youle, B. Young, C. Zgliczynski, R. Brainard, R. A. Edwards, J. Nulton, F. Thompson, and F. Rowher. Lytic to temperate switching of viral communities. *Nature*, 531:466–470, 2016.

B. Knowles, B. Bailey, L. Boling, M. Breitbart, A. Cobian-Guemes, J. del Campo, R. Edwards, B. Felts, J. Grasis, A. F. Haas, P. Katira, L. Wegley Kelly, A. Luque, J. Nulton, L. Paul, G. Peters, N. Robinett, S. Sandin, A. Segall, C. Silveira, M. Youle, and F. Rohwer. Variability and host density independence in inductions-based estimates of environmental lysogeny. *Nat. Microbiol.*, 2:17064, 2017.

J. Ladau, T.J. Sharpton, M.M. Finucane, G. Jospin, S.W. Kembel, J. O'Dwyer, A.F. Koeppel, J.L. Green, and K.S. Pollard. Global marine bacterial diversity peaks at high latitudes in winter. *ISME J.*, 7:1669–1677, 2013.

A. S. Lang, O. Zhaxybayeva, and J. T. Beatty. Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Microbiol.*, 10:472–482, 2012.

F.M. Lauro, D. McDougald, T. Thomas, T.J. Williams, S. Egan, S. Rice, M.Z. DeMaere, L. Ting, H. Ertan, J. Johnson, S. Ferriera, A. Lapidus, I. Anderson, N. Kyrpides, A.C. Munk, C. Detter, C.S. Han, M.V. Brown, F.T. Robb, S. Kjelleberg, and R. Cavicchioli. The genomic basis of trophic strategy in marine bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, 106:15527–15533, 2009.

J. Laybourn-Parry, W.A. Marshall, and N.J. Madan. Viral dynamics and patterns of lysogeny in saline antarctic lakes. *Polar Biol.*, 30:351–358, 2007.

WJ Lee and DJ Patterson. Abundance and biomass of heterotrophic flagellates, and factors controlling their abundance and distribution in sediments of botany bay. *Microbial ecology*, 43: 467–481, 2002.

P.G. Leiman, M. Basler, U.A. Ramagopal, J.B. Bonanno, J.M. Sauder, S. Pukatzki, S.K. Burley, S.C. Almo, and J.J. Mekalanos. Type vi secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc. Natl. Acad. Sci. U.S.A.*, 106:4154–4159, 2009.

A. Letarov and E. Kulikov. The bacteriophages in human-and animal body-associated microbial communities. *Journal of applied microbiology*, 107(1):1–13, 2009.

E.S. Lim, Y. Zhou, G. Zhao, I.K. Bauer, L. Droit, I.M. Ndao, B.B. Warner, P.I. Tarr, D. Wang, and L.R. Holtz. Early life dynamics of the human gut virome and bacterial microbiome in infants. *Nat. Med.*, 21(10):1228, 2015.

J.T. Lisle and J.C. Priscu. The occurrence of lysogenic bacteria and microbial aggregates in the lakes of the mcmurdo dry valleys, antarctica. *Microb. Ecol.*, 47:427—-439, 2004.

C.A. Lozupone, J.I. Stombaugh, J.I. Gordon, J.K. Jansson, and R. Knight. Diversity, stability and resilience of the human gut microbiota. *Nature*, 489:220–230, 2012.

E. Luo, J.M. Eppley, A.E. Romano, D.R. Mende, and E.F. DeLong. Double-stranded dna virioplankton dynamics and reproductive strategies in the oligotrophic open ocean water column. *ISME J.*, pages 1–12, 2020.

N.J. Madan, W.A. Marshall, and J. Laybourn-Parry. Virus and microbial loop dynamics over an annual cycle in three contrasting antarctic lakes. *Freshw. Biol.*, 50:1291–1300, 2005.

J.B.H. Martiny, L. Riemann, M.F. Marston, and M. Middelboe. Antagonistic coevolution of marine planktonic viruses and their hosts. *Annu. Rev. Mar. Sci.*, 6:393–414, 2014.

S. Maslov and K. Sneppen. Population cycles and species diversity in dynamic kill-the -wnnier model of microbial ecosystems. *Sci. Rep.*, 7:39642, 2017.

C. Maurice, C.F.and Bouvier, R. deWit, and T Bouvier. Linking the lytic and lysogenic bacteriophage cycles to environmental conditions, host physiology and their variability in coastal lagoons. *Environ. Microbiol.*, 15:2463–2475, 2013.

C. F. Maurice, T. Bouvier, J. Comte, F. Guillemette, and P. A. Del Giorgio. Seasonal variations of phage life strategies and bacterial physiological states in three northern temperate lakes. *Environmental Microbiology*, 12:628–641, 2009.

C.F. Maurice, T. Bouvier, J. Comte, F. Guillemette, and P.A. DelGiorgio. Seasonal variations of phage life strategies and bacterial physiological states in three northern temperate lakes. *Environ. Microbiol.*, 12:628–641, 2010.

C.F. Maurice, D. Mouillot, Y. Bettarel, R. De Wit, H. Sarment, and T. Bouvier. Disentangling the relative influence of bacterioplankton phylogeny and metabolism on lysogeny in reservoirs and lagoons. *ISME J.*, 5:831–842, 2011.

T.N. Mavrich and G.F. Hatfull. Evolution of superinfection immunity in cluster a mycobacteriophages. *mBio*, 10:e00971–19, 2019.

M.D. McKay, R.J. Beckman, and W.J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21: 239–245, 1979.

M.L. Mei and R. Danovaro. Virus production and life strategies in aquatic sediments. *Limnol. Oceanogr.*, 49:459—-470, 2004.

Y. Michel-Briand and C. Baysse. The pyocins of pseudomonas aeruginosa. *Biochimie*, 84:499–510, 2002.

R. Milo and R. Phillips. *Cell Biology by the numbers.* Garland Science, 1 edition, 2015.

S. Minot, R. Sinha, J. Chen, H. Li, S.A. Keilbaugh, G.D. Wu, J.D. Lewis, and F.D. Bushman. The human gut virome: inter-individual variation and dynamic response to diet. *Genome research*, 21(10):1616–1625, 2011.

S. Minot, A. Bryson, C. Chehoud, G.D. Wu, J.D. Lewis, and F.D. Bushman. Rapid evolution of the human gut virome. *Proceedings of the National Academy of Sciences*, 110(30):12450–12455, 2013.

M.K. Mirzaei and C.F Maurice. Ménage à trois in the human gut: interactions between host, bacteria and phages. *Nature Reviews Microbiology*, 15(7):397–408, 2017.

S. Muck, T. Griessler, N. Köstner, A. Klimiuk, C. Winter, and G. Herndl. Fracture zones in the mid atlantic ridge lead to alterations in prokaryotic and viral parameters in deep-water masses. *Front. Microbiol.*, 5:264, 2014.

M. Ohnishi, K. Kurokawa, and T. Hayashi. Diversification of escherichia coli genomes: are bacteriophages the major contributors? *Trends. Microbiol.*, 9:481–485, 2001.

K. J. Parikka, M. Le Romancer, N. Wauters, and S. Jacquet. Deciphering the virus-to-prokaryote ratio (vpr): insights into virus-host relationships in a variety of ecosystems. *Biol. Rev.*, 92: 1081–1100, 2017.

R.J. Parsons, M. Breitbart, M.W. Lomas, and C.A. Carlson. Ocean time-series reveals recurring seasonal patterns of virioplankton dynamics in the northwestern sargasso sea. *ISME J.*, 6: 273–284, 2012.

R.J. Parsons, C.E. Nelson, C.A. Carlson, C.C. Denman, Andersson A.J., A.L. Kledzik, K.L. Vergin, S.P. McNally, A.H. Treusch, and S.J. Giovannoni. Marine bacterioplankton community turnover within seasonally hypoxic waters of a subtropical sound: Devil's hole, bermuda. *Environ. Microbiol.*, 17:3481—-3499, 2015.

N.L. Patten, J.G. Mitchell, M. Middelboe, B.D. Eyre, L. Seuront, P.L. Harrison, and R.N. Glud. Bacterial and viral dynamics during a mass coral spawning period on the great barrier reef. *Aquatic Microbial Ecology*, 50(3):209–220, 2008.

J.H. Paul. Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J.*, 2:579–589, 2008.

J.H. Paul and M. Weinbauer. Detection of lysogeny in marine environments. In S. Wilhelm, M. Weinbauer, and C. Sutlle, editors, *Man. Aquat. Viral Ecol.* American Society of Limnology and Oceanography, Wacom, TX, 2010.

J.P. Payet and C.A. Suttle. To kill or not to kill: The balance between lytic and lysogenic viral infection is driven by trophic status. *Limnol. Oceanogr.*, 58:465–474, 2013.

J.P. Payet, R. McMinds, D.E. Burkepile, and R.L. Vega Thurber. Unprecedented evidence for high viral abundance and lytic activity in coral reef waters of the south pacific ocean. *Front. Microbiol.*, 5:493, 2014.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018. URL https://www.R-project.org/.

M. Rajilić-Stojanović, H. Smidt, and W.M. de Vos. Diversity of the human gastrointestinal tract microbiota revisited. *Environ. Microbiol.*, 9:2125–2136, 2007.

A. Reyes, M. Haynes, N. Hanson, F.E. Angly, A.C. Heath, F. Rohwer, and J.I. Gordon. Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature*, 466(7304):334–338, 2010.

J.M. Rinta-Kanto, M.J. Lehtola, T. Vartiainen, and P.J. Martikainen. Rapid enumeration of virus-like particles in drinking water samples using sybr green i-staining. *Water Res.*, 38:2614—-2618, 2004.

A. Rokney, O. Kobiler, A. Amir, D.L. Court, J. Stavans, S. Adhya, and A.B. Oppenheim. Host repsonses influence on the induction of lambda prophage. *Molecular Microbiology*, 68:29–36, 2008.

J.L. Rosner. Formation, induction, and curing of bacteriophage p1 lysogens. *Virology*, 48:679–689, 1972.

M. Schapira, M.-J. Buscot, S.C. Leterme, T. Pollet, C. Chapperon, and L. Seuront. Distribution of heterotrophic bacteria and virus-like particles along a salinity gradient in a hypersaline coastal lagoon. *Aquat. Microb. Ecol.*, 54:171—-183, 2009.

I. Sharon, M.J. Morowitz, B.C. Thomas, E.K. Costello, D.A. Relman, and J.F. Banfield. Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. *Genome Res*, 23(1):111–120, 2013.

A.N. Shkoporov and C. Hill. Bacteriophages of the human gut: the "known unknown" of the microbiome. *Cell host & microbe*, 25(2):195–209, 2019.

C.B. Silveira and F. Rohwer. Piggyback-the-winner in host-associated microbial communities. *npj Biofilms and Microbiomes*, 2:16010, 2016.

E.H. Simpson. The interpretation of interaction in contingency tables. *J. Royal Stat. Soc. Ser. B*, 13(2):238–241, 1951.

F. M. Steward and B. R. Levin. The population biology of bacterial viruses: Why be temperate. *Theor. Pop. Biol.*, 26:93–117, 1984.

S.H. Strogatz. *Nonlinear dynamics and chaos*. Westview Press, 2 edition, 2015.

D. Talmy, S.J. Beckett, D.A.A. Taniguchi, C.P.D. Brussaard, J.S. Weitz, and M.J. Follows. An empirical model of carbon flow through marine viruses and microzooplankton grazers. *Environ. Microbiol.*, 21:2171–2181, 2019.

F.T. Thingstad. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol. Oceanogr.*, 45:1320–1328, 2000.

T. F. Thingstad, S. Vage, J. E. Storesund, R.-A. Sandaa, and J. Giske. A theoretial analysis of how strain-specific viruses can control microbial species diversity. *Proc. Natl. Acad. Sci. U.S.A.*, 111:7813–7818, 2014.

T.F. Thingstad and R. Lignell. Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquat. Microb. Ecol.*, 13:19–27, 1997.

M. Touchon, A. Bernheim, and E.P.C. Rocha. Genetic and life-history traits associated with the distribution of prophages in bacteria. *ISME J.*, 10:2744–2754, 2016.

M. Touchon, J.A.M. de Sousa, and E.P.C. Rocha. Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. *Curr. Opin. Microbiol.*, 38:66–73, 2017.

S. Våge, G. Bratbak, J. Egge, M. Heldal, A. Larsen, S. Norland, M.L. Paulsen, B. Pree, R.-A. Sandaa, E.F. Skjoldal, T.M. Tsagaraki, L. Øvreås, and T.F. Thingstad. Simple models combining competition, defence and resource availability have broad implications in pelagic microbial food webs. *Ecol. Lett.*, 21:1440–1452, 2018.

L.M. Wahl, M.I. Betti, D.W. Dick, T. Pattenden, and A.J. Puccini. Evolutionary stability of the lysis-lysogeny decision: Why be virulent? *Evolution*, 73–1:92–98, 2018.

J.K. Waldor and J.J. Mekalanos. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science*, 272:1910–1914, 1996.

Z. Wang and N. Goldenfeld. Fixed points and limit cycles in the population dynamics of lysogenic viruses and their hosts. *Phys. Rev. E*, 82:011918, 2010.

M. Weinbauer. Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.*, 28:127–181, 2004.

J. S. Weitz and J. Dushoff. Alternative stable states in host-phage dynamics. *Theor. Ecol.*, 1: 13–19, 2008.

J. S. Weitz, C. A. Stock, S. W. Wilhelm, L. Bourouiba, M. L. Coleman, A. Buchan, M. J. Follows, J. A. Fuhrman, L. F. Jover, J. T. Lennon, M. Middelboe, D. L. Sonderegger, C. A. Suttle, B. P. Taylor, T. Frede Thingstad, W. H. Wilson, and E. Wommack. A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes. *ISME J.*, 9:1352–1364, 2015.

J. W. Weitz, S. J. Beckett, J. R. Brum, Cael B. B., and J. Dushoff. Lysis, lysogeny, and virus-microbe ratios. *Nature*, 549:E1–E3, 2017.

J.S. Weitz, G. Li, H. Gulbudak, M.H. Cortez, and R.J. Whitaker. Viral invasion fitness across a continuum from lysis to latency. *Virus Evol.*, 5:vez006, 2019.

Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos.* Springer-Verlag, New York, NY, second edition, 2003.

C.H Wigington, D. Sonderegger, C.P.D Brussaard, A. Buchan, J.F. Finke, J.A. Fuhrman, J.T. Lennon, M. Middelboe, C.A. Suttle, C. Stock, W.H. Wilson, K.E. Wommack, S.W. Wilhelm, and J.S. Weitz. Re-examination of the relationship between marine virus and microbial cell abundances. *Nat. Microbiol.*, 1:15024, 2016.

R.M. Wilcox and J.A. Fuhrman. Bacterial viruses in coastal seawater: lytic rather than lysogenic production. *Mar. Ecol. Prog. Ser.*, 114:35–45, 1994.

C. Winter, T. Bouvier, M.G. Weinbauer, and T.F. Thingstad. Trade-offs between competition and defense specialists among unicellular organisms: The “killing the winner” hypothesis revisited. *Microbiol. Mol. Biol. Rev.*, 74:42–57, 2010.

849 Figures and Tables

Parameter	Description	Environmental reference value
r	Bacterial growth constant	$\frac{1}{24}$ hr^{-1}
m	Viral decay constant	$\frac{1}{6}$ hr^{-1}
β	Induction rate of lysogen	$\frac{1}{24} \cdot 10^{-6}$ hr^{-1}
d	Infection rate	10^{-8} mL bacteria^{-1} hr^{-1}
c	Burst size	20
χ	Lysogen immunity	$\frac{1}{2}$
K	Carrying capacity	10^4–10^{11} bacteria mL$^-1$

Table 1: Parameters of the ecological model and reference empirical values. See the Methods section for the range explored for each parameter and literature sources.

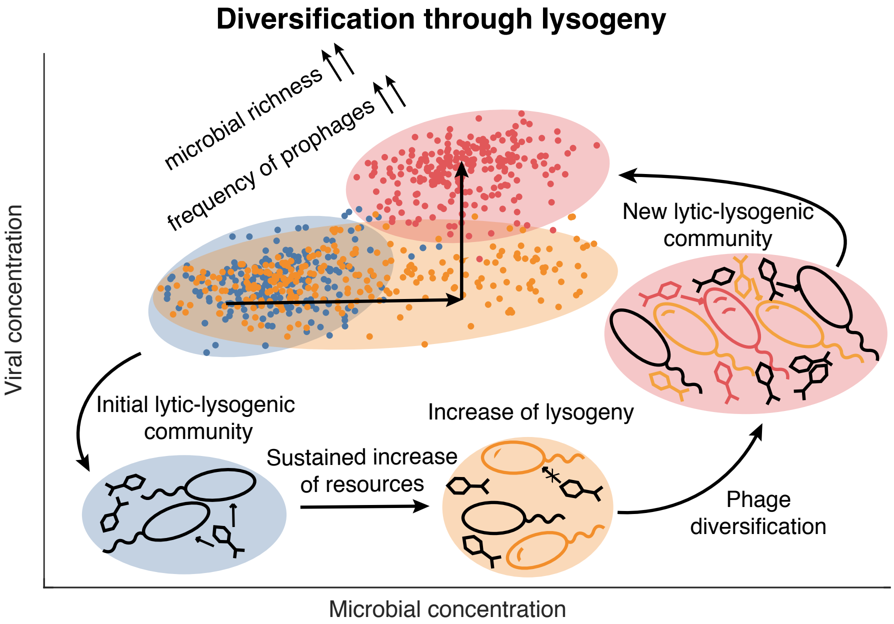

Figure 1: **Diversification through lysogeny.** The figure illustrates the central hypothesis of the research presented here. Upon increasing resources, an established viral-microbial community (blue) transitions to a new community (orange), where the microbial concentration increased due to the opportunistic growth of newly formed lysogens that escaped lytic phage top-down control by the superinfection exclusion mechanism. If the increase of resources is sustained, phages able to top-down control these newly form lysogens will eventually be selected and lead to a new mature viral-microbial community (red). This process increases the number of viral-host pairs (richness) and number of propahges per bacteria.

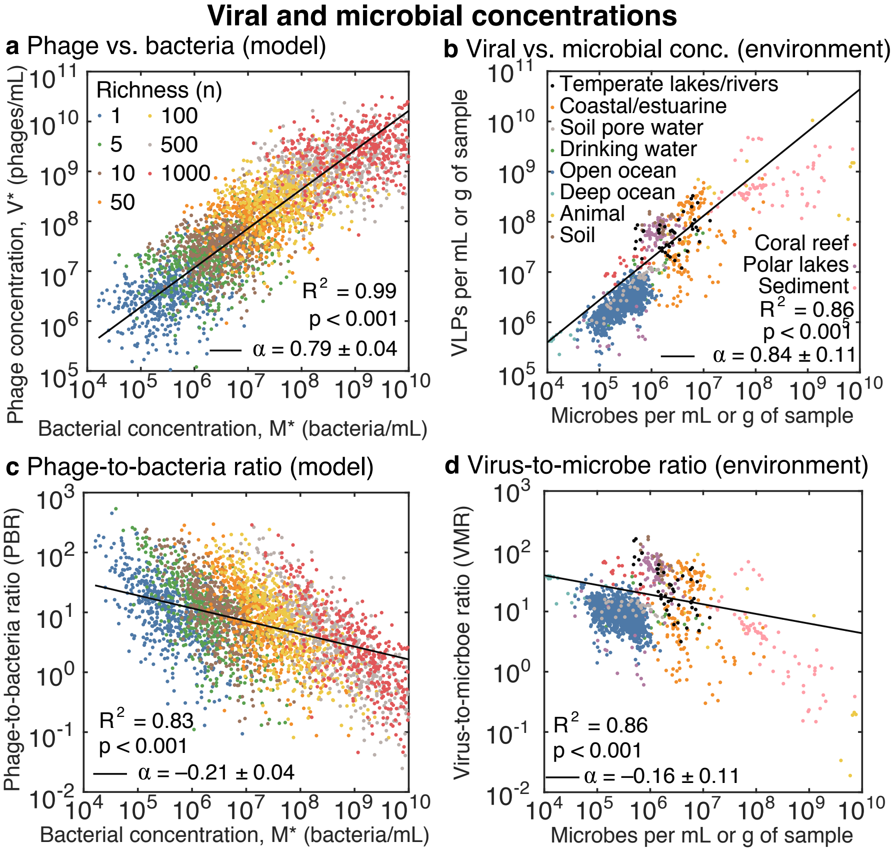

Figure 2: **Viral and microbial concentrations.** **a**, Phage equilibrium concentrations versus bacterial concentrations in the ecological lytic-lysogenic model. The colors distinguish communities with different richness, that is, number of bacteria species or phage-bacteria host pairs. **b**, Environmental concentration of virus-like particles (VLPs) versus microbial concentrations from eleven ecosystems (see legend for colors). Each ecosystems is plotted separately in Figure S.4. **c**, Phage-to-bacteria (PBR) versus bacterial concentration obtained in the model (same color coding as in panel **a**). **d**, Environmental virus-to-microbe ratio (VMR) versus microbial concentration for the ecosystems listed in panel **b**. **a-d.** The solid lines correspond to the linear regression of the logged data. The slope ($\alpha \pm$ SE), p-value ($\alpha \neq 0$), and the coefficient of determination (R^2) are displayed.

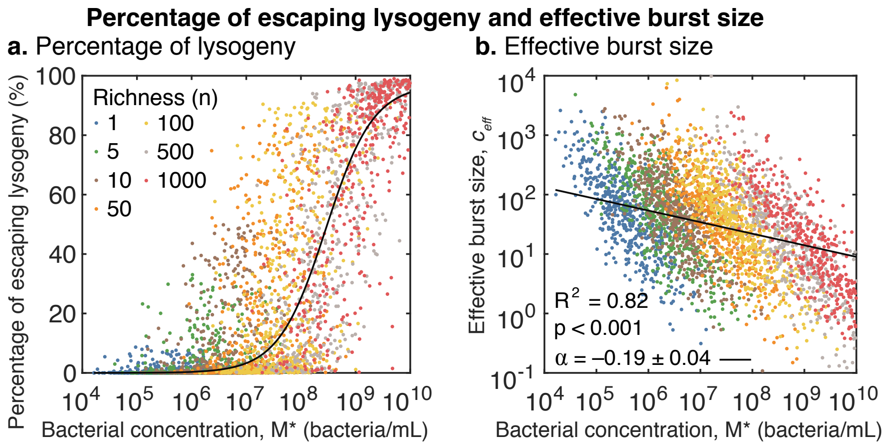

Figure 3: **Percentage of lysogeny and effective burst size.** **a**, Sampled percentage of active lysogeny as a function of bacterial concentration for the lytic-lysogenic community model. The solid line corresponds to the best fitted sigmoidal function $L[\%] = aB^\star/(b + B^\star)$ for the medians of the percentage of lysogeny, $L[\%]$, and total bacterial concentrations, $B^\star$, for each community, n. The parameters were obtained using the nonlinear least-squares Gauss-Newton algorithm implemented in the Nonlinear Least Squares (nls) function in R. The values obtained for the parameters corresponded to $a = 97.1 \pm 0.7$ and $b = (2.86 \pm 8) \cdot 10^8$ bacteria/ml. **b**, Effective burst size (c_{eff}) as a function of bacterial concentration for the interpretation of the lytic-lysogenic community model as an effective single phage-bacteria community pair model. **a-b**, The colors correspond to communities with richness ranging from $n = 1$ to $n = 1000$.

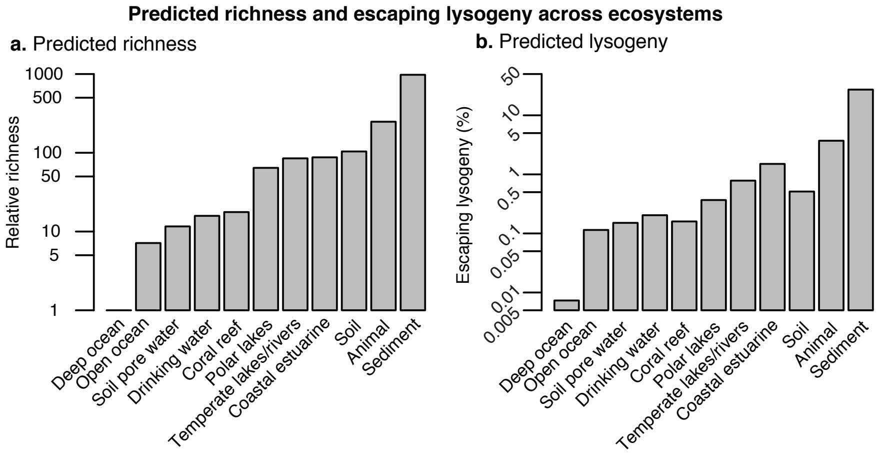

Figure 4: **Predicted richness and escaping lysogeny across ecosystems. a,** Relative community richness predicted for different ecosystems using the lytic-lysogenic community model. The values are given relative to the Deep Ocean data (lowest richness). **b,** Percentage of escaping lysogeny predicted by the model for the different ecosystems.

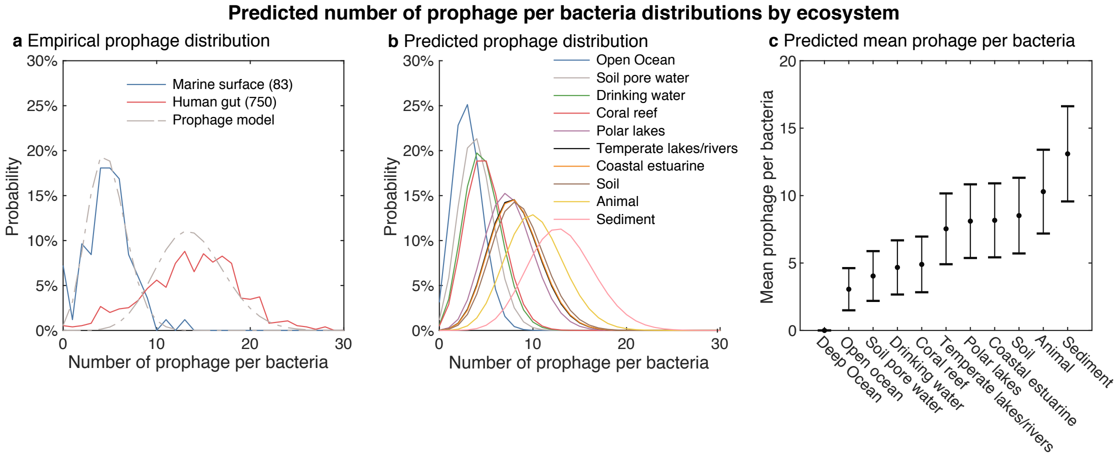

Figure 5: **Evolutionary model predictions.** **a**, Percentage of number of prophage per bacteria found in marine surface ecosystems (blue line) and human gut samples (red line), as well as the prophage distributions obtained from the prophage accumulation evolutionary model (grey dashed line). **b**, Predicted probability distributions for the number of prophages per bacteria in each ecosystem using the prophage accumulation evolutionary model. **c**, Mean number of prophage per bacteria predicted in each ecosystem. The error bars correspond to the associated standard deviations.