Quantification of lysogeny caused by phage coinfections in microbial communities from biophysical principles

Antoni Luque[1,2,3] and Cynthia Silveira[4]

[1] Department of Mathematics and Statistics, San Diego State University

[2] Viral Information Institute, San Diego State University

[3] Computational Science Research Center, San Diego State University

[4] Department of Biology, University of Miami

Key words: stochastic biophysical model, lysogeny, microbial abundance, multiplicity of infection, adsorption rates, commitment time

Words: 4,935

Figures: 5

References: 86

Abstract

Temperate phages can integrate in their bacterial host genome to form a lysogen, often modifying the phenotype of the host. Lysogens are dominant in the microbial-dense environment of the mammalian-gut. This observation contrasts with the long-standing hypothesis of lysogeny being favored in microbial communities with low densities. Here we hypothesized that phage coinfections—the most studied molecular mechanism of lysogeny in lambda phage—increases at high microbial abundances. To test this hypothesis, we developed a biophysical model of coinfection and stochastically sampled ranges of phage and bacterial concentrations, adsorption rates, lysogenic commitment times, and community diversity from marine and gut microbiomes. Based on lambda experiments, a Poisson process assessed the probability of lysogeny via coinfection in these ecosystems. In 90% of the sampled marine ecosystems, lysogeny stayed below 10% of the bacterial community. In contrast, 25% of the sampled gut communities stayed above 25% of lysogeny, representing an estimated nine trillion lysogens formed via phage coinfection in the human gut every day. The prevalence of lysogeny in the gut was a consequence of the higher densities and faster adsorption rates. In marine communities, which were characterized by lower densities and phage adsorption rates, lysogeny via coinfection was still possible for communities with long lysogenic commitments times. Our study suggests that physical mechanisms can favor coinfection and cause lysogeny at poor growth conditions (long commitment times) and in rich environments (high densities and adsorption rates).

41 **Importance**

42

43 Phage integration in bacterial genomes manipulate microbial dynamics from the oceans to
44 the human gut. This phage-bacteria interaction, called lysogeny, is well-studied in
45 laboratory models, but its environmental drivers remain unclear. Here we quantified the
46 frequency of lysogeny via phage coinfection—the most studied mechanism of lysogeny—by
47 developing a biophysical model that incorporated a meta-analysis of the properties of
48 marine and gut microbiomes. Lysogeny was found to be more frequent in high-productive
49 environments like the gut, due to higher phage and bacterial densities and faster phage
50 adsorption rates. At low cell densities, lysogeny via coinfection was possible for hosts with
51 long duplication times. Our research bridges the molecular understanding of lysogeny with
52 the ecology of complex microbial communities.

53

Introduction

Temperate phages are viruses of bacteria that, upon infection, can integrate in their host's genome as a prophage. This forms a symbiont with the bacterial host called lysogen. Half of bacterial genomes that have been sequenced contain prophages (Casjens, 2003; Fouts, 2006; Touchon *et al.*, 2016). Most lysogens display enhanced capabilities such as protection against other phage infections and additional metabolic pathways (Bondy-Denomy *et al.*, 2016; Jerlström Hultqvist *et al.*, 2018; Mavrich and Hatfull, 2019). Therefore, lysogeny has profound impacts in the structure and functioning of microbial communities. However, the drivers of lysogeny in microbial communities remain unclear.

The distribution of integrases, excisionases, and lambda-like repressor genes currently represents the best proxy to estimate the relative frequency of lysogeny in microbial communities (Luo *et al.*, 2020; Silveira *et al.*, 2020). The application of this approach to metagenomic analysis indicates that microbial-dense environments, such as the mammalian gut, are dominated by temperate phages and bacterial lysogens, compared with low-density marine ecosystems (Reyes *et al.*, 2010; Minot *et al.*, 2011, 2013; Beller and Matthijnssens, 2019; Shkoporov and Hill, 2019; Mirzaei *et al.*, 2020). In complete bacterial genomes from isolates, higher frequency of prophages is also observed in the gut (Anthenelli *et al.*, 2020). The gut-associated temperate phages mediate bacterial colonization dynamics in mouse models (Reyes *et al.*, 2013; Frazão *et al.*, 2019), and their distribution patterns have been associated with the presence of severe acute malnutrition (Reyes *et al.*, 2015).

77

78 Compared to animal-associated microbial communities, marine ecosystems have much
79 weaker lysogenic signatures (Knowles *et al.*, 2016; Luo *et al.*, 2020). Genomic analyses have
80 indicated that lysogeny is frequent in deep oligotrophic waters where bacterial abundances
81 are low, ranging from 10^4 to 10^6 cells per ml (Muck *et al.*, 2014; Luo *et al.*, 2017). The
82 increase in lysogeny in marine ecosystems with low productivity has been historically
83 hypothesized to serve a low-density refugium for phages during poor host growth (Paul,
84 2008). Yet, high intrinsic growth rate is the most important predictor of the frequency of
85 prophages in gut and marine bacteria with complete genomes sequenced (Lauro *et al.*,
86 2009; Touchon *et al.*, 2016). Lysogeny also increases in shallow marine environments
87 where abundances of bacteria rise above 10^6 cells/ml (Knowles *et al.*, 2016). The
88 observations of lysogeny at both high and low productivity conditions suggest that an
89 interplay between host physiology and density-dependence mechanisms modulate
90 lysogeny in complex communities (Thingstad, 2000; Weitz *et al.*, 2015; Knowles *et al.*,
91 2016).

92

93 Phage coinfections are the best understood mechanism for the establishment of lysogeny in
94 model phage and bacteria (Golding, 2016; Jimmy T. Trinh *et al.*, 2017). The percentage of
95 lysogenized cells increase with the multiplicity of infection (MOI) in lambda phage and *E.*
96 *coli* (Boyd, 1951; Lieb, 1953; Fry, 1959; Kourilsky, 1973; Herskowitz and Hagen, 1980).
97 Coinfections increase the expression of the cII repressor lytic pathway (phage replication)
98 and activate a cascade of genes responsible for phage integration (Kourilsky, 1975;

Ptashne, 1986; Cheng *et al.*, 1988). Yet, the extent to which coinfections occur in complex microbial communities, and their impact on lysogeny has not been quantified.

After its first introduction, the term MOI became known in the field as the initial ratio of phage particles (P_0) to bacterial cells (B_0) added to an experiment (MOI = P_0/B_0). This definition of MOI, however, does not capture the effective number of coinfections, which also depends on the chances of encounter between the phage and the host (Joiner *et al.*, 2019). A recent stochastic model shows that at the single-cell level, the average number of coinfections is primarily determined by the phage concentrations and phage decision times (Lang *et al.*, 2020). Since high productive environments have higher phage concentrations, here we hypothesize that the prevalence of lysogeny in these environments is a consequence of an increase in coinfections.

To test this hypothesis, a biophysical model was derived to incorporate the physical traits that determine phage coinfections (COI) and its associated probability of lysogeny. The model was simulated for a range of phage and bacterial abundances, community diversity, adsorption rates, and lysogenic commitments times from marine and mammalian gut communities. The availability of large amounts of public data from these two ecosystems allowed to test the hypothesis across a wide range of microbial densities.

Methods

Physical model of the probability of lysogeny through phage coinfection. The average number of phage coinfections (COI) was derived from physical properties of phage and bacteria (Figure 1a). The rate of phage infections on a single bacterium can be estimated by solving the Smoluchowski coagulation equation (Joiner *et al.*, 2019). In steady state, this rate is the product of the phage concentration (P) and the phage adsorption rate (α), which depends on the mobilities and sizes of both the phage particle and the bacterium. The number of coinfections is defined by the number of phages infecting a cell within a given time window. This number is the product of the infection rate and the time window. In the case of lysogeny, this time corresponds to the lysogenic commitment time (τ) (Jimmy T. Trinh *et al.*, 2017). This led to the average phage coinfection equation (COI)

$$\mathrm{COI} = \mathrm{P} \cdot \alpha \cdot \tau \,. \qquad \text{Eq (1)}$$

The average coinfection (COI) was also expressed as a function of the phage-to-bacteria ratio, $\mathrm{COI} = \alpha \cdot \tau \cdot \mathrm{B} \cdot (\mathrm{P}/\mathrm{B})$, for one single phage-host pair to facilitate the comparison with lambda MOI experiments. Here, B was the bacterial concentration. COI was studied numerically as a function of the phage-to-bacteria ratio (P/B) for the range 0.01–100. The median values extracted for the lambda adsorption rate (α_0 = 5.6 10^{-7} ml/h), lysogenic commitment time (τ_0 = 0.1 h), and bacterial centration (B_0 = 5 10^8 cells/ml) were used as reference values (see Methods section on meta-analysis for lambda parameters). Two parameters were fixed at these reference values, and the third was explored over a range of values. These ranges were 10^5 to 10^{10} cells/ml for bacterial concentrations, 10^{-11} to 10^{-6} ml/h for the phage adsorption rates, and 10^{-3} to 10^2 h for the lysogenic commitment times.

A lysogen was formed when a cell was infected within the lysogenic commitment time by two or more phages from the same pair. This was based on the effect of cooperative infection by lambda phages on the production of lysogens (Zeng *et al.*, 2010; Golding, 2016; Jimmy T. Trinh *et al.*, 2017). Thus, the probability of lysogeny, p_{lys}, was determined by p_{lys}= 1 – p(0) – p(1), where p(k) was the probability of having k coinfections within the commitment time. The probability of coinfections was calculated by assuming that each infection was independent, and that the phage concentration (P), phage adsorption rates (α), and lysogenic commitment times (τ) remained constant. Following these assumptions, the probability of k coinfections was given by a Poisson process with average coinfection COI, Eq. (1), leading to the equation $p(k) = COI^k e^{-COI}/k!$. Thus, the probability of forming a lysogen via coinfection was

$$p_{lys} = 1 - e^{-COI} - COI\, e^{-COI}. \qquad \text{Eq (2)}$$

This probability of lysogeny was compared with the percentage of lysogeny obtained from lambda and *E. coli* MOI experiments (Fry, 1959; Kourilsky, 1973). The empirical data for percentage of lysogeny and MOI was fitted using non-linear least squares method for Hill-Langmuir cooperation models, $f(x) = ax/(b + x^n)$, with cooperation orders n=1, n=2, and n=3.

Meta-analysis of marine and gut microbiomes. *Adsorption rates.* The adsorption rates were obtained from prior experiments for 71 phage-host pairs from marine and gut microbiomes (Source data 1). The data consisted of values for tailed phages infecting *E. coli* (Schwartz, 1976; Joiner *et al.*, 2019), *Synechococcus sp.* (Waterbury and Valois, 1993; Mann, 2003; Stoddard *et al.*, 2007), *Prochlorococcus sp.* (Avrani *et al.*, 2011; Schwartz and Lindell,

2017; Beckett *et al.*, 2019), *Vibrio sp.* (Johnson, 1968; Levisohn *et al.*, 1987; Cohen *et al.*, 2013; Traits *et al.*, 2017), *Roseobacter sp.* (Huang *et al.*, 2010), and *Pseudoalteromonas sp.* (Deng *et al.*, 2013). A t-test (double tailed) compared the marine and gut values.

Lysogenic commitment times. The lysogenic commitment time (τ) was assumed to be 20% of the bacterial duplication time (Cortes *et al.*, 2017; Jimmy T. Trinh *et al.*, 2017). The ranges of bacterial duplication times at the community level were obtained *in situ* for marine ecosystems (Kirchman, 2016) and *in vivo* for mammalian gut ecosystems (Heinken *et al.*, 2013; Myhrvold *et al.*, 2015; Vandeputte *et al.*, 2016) (Source data 2).

Viral-like particles and cell abundances. Direct counts of viral-like particles (VLPs) and microbial cells were obtained for marine surface waters (Parsons *et al.*, 2012; Knowles *et al.*, 2016) and animal-associated microbiomes (Hoyles *et al.*, 2008; Lepage *et al.*, 2008; Furlan, 2009; Kim *et al.*, 2011; Barr *et al.*, 2013) (Source data 3). A t-test (double tailed) was applied to compare the concentrations of VLPs and cells between marine and animal ecosystems.

Phage and bacterial diversities. The rank-abundance curves of the frequency of phage genotypes representing unique viral contigs at 98% sequence identity were reconstructed from the median slope and intercept of power-law functions fitted to 192 marine viromes and 1,158 human-associated viromes (Cobián Güemes *et al.*, 2016) (Source data 4). The rank-abundance curves of the frequency of bacterial species in marine communities were obtained from operational taxonomic unit (OTU) tables constructed by clustering

189 universal, protein-coding, single-copy phylogenetic marker genes into metagenomic OTUs
190 (which can be interpreted as species-level clusters) from the Tara Oceans dataset (Source
191 data 5) (Mende *et al.*, 2013; Sunagawa *et al.*, 2013). For animal-associated bacterial
192 microbiomes, rank-abundance curves were constructed using OTU tables obtained by
193 mapping metagenomic reads from 11,850 human gut-associated metagenomes to 92,143
194 metagenome-assembled genomes (Almeida *et al.*, 2019). Consensus rank abundance curves
195 were obtained by averaging the frequency of bacteria in the same rank across the
196 metagenomes of marine and gut associated microbiomes, respectively.
197
198 **Quantification of lysogeny through phage coinfection in communities.** The biophysical
199 COI model, Eqs. (1) and (2), was applied to predict the probability of lysogeny in marine
200 and gut ecosystems as a result of coinfections. The model generated stochastic
201 communities that sampled empirical phage and bacteria concentrations, relative
202 abundance of the top 100 members of the community, phage adsorption rates, and
203 lysogenic commitment times obtained from the meta-analysis of marine and gut
204 ecosystems described above.
205
206 *Stochastic sampling.* The model generated 100,000 stochastic communities for both marine
207 and gut ecosystems using latin hypercube sampling (LHS). For each ecosystem, the range of
208 bacterial concentrations, adsorption rates, and lysogenic commitment times were divided
209 in 100,000 equidistributed values in logarithmic scale (base 10). 100,000 random values
210 were selected from each range and combined without repeating any value, that is, all

211 equidistributed values were sampled, following the standard LHS implementation (McKay
212 *et al.*, 1979; Weitz *et al.*, 2017; Anthenelli *et al.*, 2020)
213
214 *Parameter ranges.* The ranges of bacterial concentrations used were $3.78 \cdot 10^4$ – $6.75 \cdot 10^6$
215 bacteria/ml (marine) and $3.45 \cdot 10^5$ – $7.60 \cdot 10^9$ bacteria/ml (gut). The ranges of phage
216 concentrations used were $1.45 \cdot 10^5$ – $3.80 \cdot 10^7$ phages/ml (marine) and $5.09 \cdot 10^6$ –
217 $1.05 \cdot 10^{10}$phages/ml (gut). The ranges of phage adsorption rates used were $7.2\ 10^{-10}$–3.7
218 10^{-7} ml/h for marine and $5.9\ 10^{-8}$–$1.2\ 10^{-6}$ for gut. The ranges of lysogenic commitment
219 times used were 11 h – 808 h for marine and 2.74 h –7.27 h for gut.
220
221 *Assumed relationships.* The phage concentration (P) was assumed to follow a power
222 function relationship with the bacterial concentration (B)(Knowles *et al.*, 2016; Anthenelli
223 *et al.*, 2020): $P(B) = a\,(B/B_u)^b$, where the bacterial concentration was given in units of B_u =
224 bacteria/ml. The prefactor a and exponent b were obtained by fitting the power function to
225 the viral and microbial counts obtained in the marine and gut meta-analysis. A linear
226 regression fit was applied using least-squares method to the log-log data in base 10. The
227 parameters obtained were $a = 10^{2.50}$ and $b = 0.712$ for marine and $a = 10^{5.35}$ and $b = 0.388$
228 for gut. To reproduce the noise observed in empirical viral communities, the value P(B)
229 was weighted by a normal distribution, N(mean, S.D.), with mean 1 and standard deviation
230 0.05, that is, $P = P(B)*N(1,0.05)$. The final value of the phage concentration was constrained
231 within the empirical phage range, that is, $P_{min} \leq P \leq P_{max}$. The community model also
232 assumed that the most dominant phages infected the most dominant bacteria (Minot *et al.*,

2011, 2013; Coutinho *et al.*, 2017). This led to a phage-host network where the phage of rank *i* infected the bacteria with the same rank *i*.

Results

Relationship between COI and phage-to-bacteria ratios. The model introduced here predicted the average number of phage coinfections (COI) occurring across a gradient of phage-to-bacteria ratios for single pairs (Figure 1a). To illustrate the relationship of COI with different physical parameters, the typical laboratory values for lambda were used as a reference, and only one value was varied at a time covering typical environmental ranges (see meta-analysis section for environmental ranges). Coinfection (COI) increased with phage-to-bacteria ratio (P_i/B_i) (Figure 1). The level of COI was higher for higher bacterial concentration (Figure 1b), phage adsorption rate constant (Figure 1c), and lysogenic commitment time (Figure 1d). The model showed that an average of two or more lambda coinfections (COI $\geqslant$ 2) were unlikely to occur at bacterial densities below 10^6 cells/ml, even for phage-to-bacteria ratios above 10 (bottom right values in Figure 1b). The typical adsorption rates for lambda were near the upper bound of coinfections generated per phage-to-bacteria ratio (Figure 1c). Due to the short lysogenic commitment times of lambda, the average coinfections per phage-to-bacteria ratio unit were below the environmental values (Figure 1d).

The probability of lysogeny as a function of average coinfections (COI) was compared with lambda-*E. coli* MOI experiments, where MOI was defined operationally as the initial phage-

to-bacteria ratio. The percentage of lysogenized cells increased as a function of MOI and was best described by a sigmoidal Hill-Langmuir equation of order n = 2, compared to orders n = 1 and n = 3 (Figure 2a). This equation implied that two phages cooperated in producing lysogeny, in agreement with single-molecule experiments (Jimmy T Trinh *et al.*, 2017). This empirical model was functionally similar to the predicted probability of lysogeny from the average coinfection Poisson model, Eq. (2), which assumed that at least two co-infections were necessary to produce lysogeny (Figure 2b). The discrepancy between the maximum percentage of lysogeny for the MOI and COI models was due to the fact that the MOI model was a function of the initial phage-to-bacteria ratio instead of the effective number of coinfections per cell.

Meta-analysis of COI physical parameters from marine and animal ecosystems. To apply the biophysical COI model to microbial communities and estimate lysogeny generated by phage coinfections, the ranges of phage adsorption rates, lysogenic commitment times, and phage-bacteria pair abundances were determined for marine and animal ecosystems. The range of adsorption rates was $7.2\ 10^{-10}$ – $3.7\ 10^{-7}$ ml/h for marine phages infecting *Prochlorococcus, Roseobacter, Pseudoalteromonas, Synechococcus, and Vibrio* (Figure 3a). The range of adsorption rates was $5.9\ 10^{-8}$ – $1.2\ 10^{-6}$ ml/h for gut phages infecting *E. coli*. The median adsorption rates for gut phages was one order of magnitude higher (median $4.21\ 10^{-7}$ ml/h) than the median for the marine phages (median $3.43\ 10^{-8}$ ml/h, Figure 3c, t-test p = $7.23\ 10^{-10}$). The ranges of estimated lysogenic commitment times were much longer in marine communities, 11–808 h, than in mammalian gut, 2.74–7.27 h. This was a consequence of the long duplication times of marine communities in their

natural environment (Source data 5). The phage and bacteria pair abundances were determined by combining the total and relative abundances of phage and bacteria in each ecosystem. The total abundances were at least two orders of magnitude higher in animal-associated mucosa compared to the free-living communities of surface marine environments (t-test $p = 7.02\ 10^{-15}$ for phage, Figure 3b, and p-value = $4.17\ 10^{-7}$ for bacteria, Figure 3c). For phages, the ranges were $1.4\ 10^{5}$–$3.7\ 10^{7}$ phages/ml (marine) and $5.1\ 10^{6}$–$1.1\ 10^{10}$ phages/ml (animal), while bacteria ranged from $3.8\ 10^{4}$ to $6.8\ 10^{6}$ cells/ml (marine) and from $3.5\ 10^{5}$ to $7.7\ 10^{9}$ cells/ml (animal) (Figures 3b and 3c). The most abundant phage genotype (P_1) in marine environments comprised only 0.8% of the total phage community, while in the gut, the dominant phage comprised just over 1% of the community (Figure 3d). In the bacterial community, this pattern was inverted, with the dominant bacterial species (B_1) reaching 19% in marine environments, but only 15% in the gut (Figure 3e).

Estimated lysogeny by coinfection in microbial communities. The community model assumed a direct phage-host network, where each phage rank infected the same rank level in the bacterial community, that is, P_i infected B_i (Figure 4a). The biophysical COI model quantified the percentage of lysogeny generated by phage coinfections for each pair by stochastically sampling the parameter ranges from the meta-analysis of marine and gut ecosystems. The percentage of lysogeny caused by coinfections increased with total bacterial density (Figures 4b and S.1). Lysogeny was more frequent in the gut, where 25% of the simulated communities displayed at least 25% of bacteria becoming lysogens by coinfection (Figure 4c and Tables S.1 and S.2). Instead, 90% of the marine communities

displayed 10% or less bacteria becoming lysogens by coinfection (Figure 4c and Tables S.1 and S.3). Among communities with a percentage of lysogeny larger than 1%, the most abundant phage-host pairs contributed an average of 67±12% (S.D.) for marine and 51±16% for gut to the total lysogeny (Figures 4d and S.2). This was significantly higher than the contribution from the second most abundant phage-host pair, which yielded 13±1% for marine and 15±2% for gut. For communities with lysogeny above 1%, the most abundant phage-host rank displayed median average coinfections of COI = 1.00 for marine (Figures 4e and S.2 and Table S.4) and COI = 2.35 for gut (Figure 4f and S.2 and Table S.5).

Physical parameters contributing to the formation of lysogens in communities.

Communities with at least 1% lysogeny caused by coinfection were analyzed to extract the distribution of physical parameters associated to lysogeny. The distribution of bacterial abundances favoring lysogeny in marine communities was skewed towards high densities with a median of 1.5 10^6 cells/ml (Figure 5a and Table S.4). In gut communities, low bacterial abundances did not contribute to lysogeny, displaying a first quartile at 3.8 10^5 cells/ml (Figure 5a and Table S.5). Phage concentrations in marine communities were also skewed towards higher densities (median of 9.1 10^6 phages/ml) but more centered compared to the bacterial density distribution (Figure 5b and Table S.4). In the gut, low phage concentrations did not contribute to lysogeny, displaying a first quartile at 2.8 10^8 phages/ml (Figure 5b and Table S.5).

Phage adsorption rates in marine communities producing lysogeny were skewed towards high values with a median of 1.1 10^{-7} ml/h (Figure 5c and Table S.4). In the gut, instead, the

full range of adsorption rates contributed to communities with lysogeny (Figure 5c and Table S.5). The lysogenic commitment time in marine communities producing lysogeny was again skewed towards long window times, with a median of 262 h (Figure 5d and Table S.4). For the gut, the full range of lysogenic commitment times contributed to communities with lysogeny, but larger values had higher likelihood of contribution, displaying a median lysogenic commitment time of 0.92 h (Figure 5d and Table S.5).

Discussion

The stochastic biophysical COI model estimated an increase of coinfections and percentage of lysogeny in high-productive mammalian gut microbial communities (Figure 4). This higher frequency of coinfections explained the dominance of lysogenic bacteria observed in animal mucosa microbiomes (Reyes *et al.*, 2010; Minot *et al.*, 2013; Silveira and Rohwer, 2016; Shkoporov and Hill, 2019; Anthenelli *et al.*, 2020; Mirzaei *et al.*, 2020). The model also indicated that the formation of lysogens through coinfection was compatible with the observation of low virus-to-microbe ratios (VMR < 1) in high-density animal mucosa (Kim *et al.*, 2011; Mirzaei *et al.*, 2020) (Figures 3 and 5). This was possible because coinfections were not required to occur simultaneously. Instead, they occurred within the lysogenic commitment time, which was assumed to be proportional to the duplication time of bacteria *in vivo* (Heinken *et al.*, 2013; Myhrvold *et al.*, 2015; Vandeputte *et al.*, 2016).

The lysogenic commitment time played a paramount role in the findings of the model because the duplication time of gut bacteria *in vivo* is significantly slower than the

348 duplication time of gut bacterial isolates in pure cultures (Heinken *et al.*, 2013; Myhrvold *et*
349 *al.*, 2015; Vandeputte *et al.*, 2016). For laboratory *E. coli* duplication times, the model
350 predicted average number of coinfections almost an order of magnitude lower than *in vivo*
351 (Figure 1d). This result aligned with the low number of coinfections observed even at
352 higher MOIs in a recent agent-based stochastic model that used similar parameters to
353 lambda and *E. coli* (Lang *et al.*, 2020). The influence of the lysogenic commitment time was
354 even more pronounced in marine communities (Figure 5d). Levels of lysogeny above 10%
355 were only possible for a small fraction of communities that displayed long lysogenic
356 commitment times, which compensated the lower adsorption rates and phage and bacteria
357 abundances in marine communities (Figure 3 and 5).
358
359 The relationship between the lysogenic commitment time and duplication time in the
360 biophysical COI model was based on lambda-*E. coli* single-cell experiments, which
361 identified the relationship $\tau \sim 20\%$ of the duplication time (Cortes *et al.*, 2017; Jimmy T.
362 Trinh *et al.*, 2017). This pattern arises from the voting phenomenon, where each
363 coinfecting phage genome independently votes for lytic or lysogenic commitment (Zeng *et*
364 *al.*, 2010). In cells that undergo lysogeny phages cooperate for host cell resources (Jimmy T.
365 Trinh *et al.*, 2017). Late phage genome arrivals contribute less to the cell-level decision, and
366 the time window for the contribution of the second phage is proportional to the
367 transcription level of cII repressor, which varies with the host growth rate (Cortes *et al.*,
368 2017). Future studies addressing the relationship between the lysogenic commitment time
369 and host duplication time in other phages would facilitate the refinement of the model and
370 further test its validity in different ecosystems.

The stochastic community simulations assumed that the most dominant phages infected the most dominant bacteria (Minot *et al.*, 2011, 2013; Coutinho *et al.*, 2017). This assumption allowed to bridge the environmental data on total viral and bacterial abundances with the species distributions from genomic data (Figure 3). It resulted into a higher contribution from the dominant ranks to lysogeny (Figure 4d). While the positive relationship between phage and host abundance has been shown for some dominant species in marine and gut samples (Deng *et al.*, 2013; Zhao *et al.*, 2013; Džunková *et al.*, 2019), the prediction of phage hosts from genomic data is a current challenge, and the host of most phages identified through metagenomics remain unknown (Edwards *et al.*, 2016; Gregory *et al.*, 2019; Silveira *et al.*, 2020). The reconstruction of accurate phage-bacterial infection networks will in the future improve the models' predictive power (Labonté *et al.*, 2015; Marbouty *et al.*, 2017).

Only 10% of marine communities simulated displayed levels of lysogeny above 10%. These communities were characterized by high phage and bacterial concentrations for marine microbiomes, displaying medians of 1.3 10^7 phages/ml and 2.0 10^6 cells/ml (Figures 5a and 5b). This finding was consistent with the weak signals of lysogeny observed in marine ecosystems (Knowles *et al.*, 2016; Luo *et al.*, 2020). The average percentage of lysogens formed by coinfection in the community decreased by almost two orders of magnitude as total bacterial concentrations reduced from 10^6 to 10^5 cells/ml (Figure S.1b). Phage coinfection, thus, could not explain an increase of lysogeny at low cell densities. Viral metagenomic studies from deep oceans with microbial abundances ranging from 10^4 to 10^5

394 cells per ml showed an increase in the genomic evidence for lysogeny compared to more
395 productive surface waters (Coutinho *et al.*, 2019; Luo *et al.*, 2020). In these environments,
396 lysogeny has been proposed to serve as low-density refugium for temperate phages in
397 conditions of poor host growth and scarce resources for viral particle production (Jiang
398 and Paul, 1994; Maurice *et al.*, 2010; Paul and Weinbauer, 2010; Brum *et al.*, 2015). In the
399 biophysical COI results, lysogeny in these low cell density communities was only possible at
400 very long commitment times. Additional mechanisms might be necessary to support the
401 increase of lysogeny at these low concentrations. One explanation could be the favored
402 phage integration in starved cells as observed in lambda due to the reduced degradation of
403 the cII repressor (Shotland *et al.*, 1997).
404
405 Metabolic shifts might also be reinforcing the formation of lysogens at high cell
406 abundances. High energy availability modifies the bacterial central carbon metabolism by
407 decoupling growth rates from growth efficiencies (Russell and Cook, 1995; Haas *et al.*,
408 2016; Silveira *et al.*, 2019). The metabolic decoupling, which is common is hypoxic and
409 anaerobic environments, decreases the ATP yield turning the intracellular environment
410 more favorable for lysogeny. As a result, high-productive environments may favor lysogeny
411 by both phage coinfections and the decrease in ATP yield. The incorporation of metabolic
412 mechanisms was beyond the scope of the current biophysical model but will improve
413 predictions of lysogeny in the future.
414
415 **Conclusion**
416

417 The stochastic biophysical COI model proposed here identified the ranges of physical
418 parameters that drive phage coinfections in complex microbial communities. The model
419 predicted high frequency of lysogeny caused by phage coinfections in the high-density
420 mammalian gut. This finding was a consequence of high phage and bacterial densities and
421 fast phage adsorption rates in comparison with marine communities. Longer lysogenic
422 commitment times *in vivo* when compared to laboratory isolates also contributed to high-
423 lysogeny in the gut. The simulated marine communities showed lower frequency of
424 lysogeny by coinfections. Those communities that displayed a high fraction of lysogeny
425 were characterized by long lysogenic commitment times. Our findings bridge the main
426 molecular mechanism causing lysogeny in laboratory systems with metagenomic
427 observations of lysogeny in complex microbial communities.
428

Figures

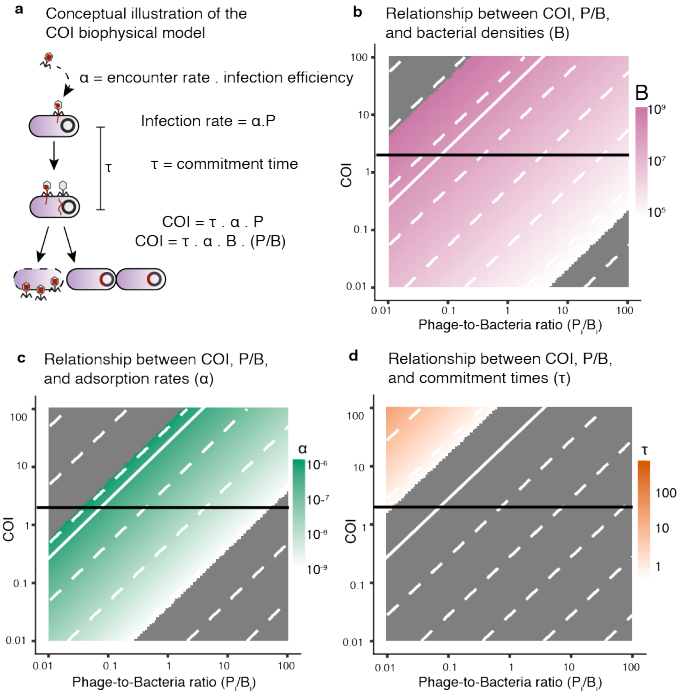


Figure 1. Relationship between number of coinfections and phage-to-bacteria ratios. **a**) Illustration of the derivation and parameter description for the biophysical coinfection (COI) model, Eq. (1). Relationship between COI and phage-to-bacteria ratios as a function of bacterial abundances (**b**), adsorption rates (**c**), and lysogenic commitment times (**d**). The color gradients cover the environmental ranges of values for each of these parameters: bacterial concentrations (pink), adsorption rates (green), and lysogenic commitment times (orange). (**b–d**) The dotted white lines indicate constant values of the parameters in the gradient scale. The grey areas correspond to values beyond the environmental ranges of bacterial concentrations, adsorption rates, and commitment times obtained from the meta-analysis of marine and gut ecosystems. The horizontal black line indicates COI = 2. The solid white line indicates the median values for the bacterial concentration (B_0), adsorption rates (α_0), and lysogenic commitment time (τ_0) for lambda.

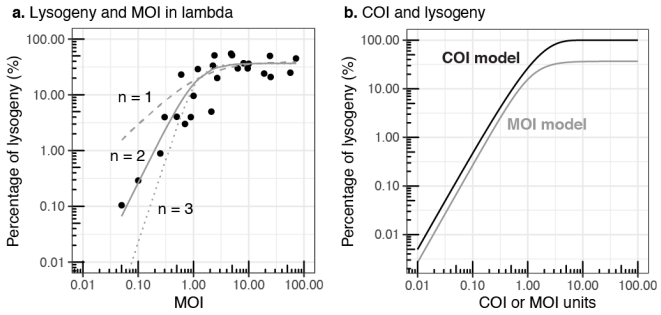
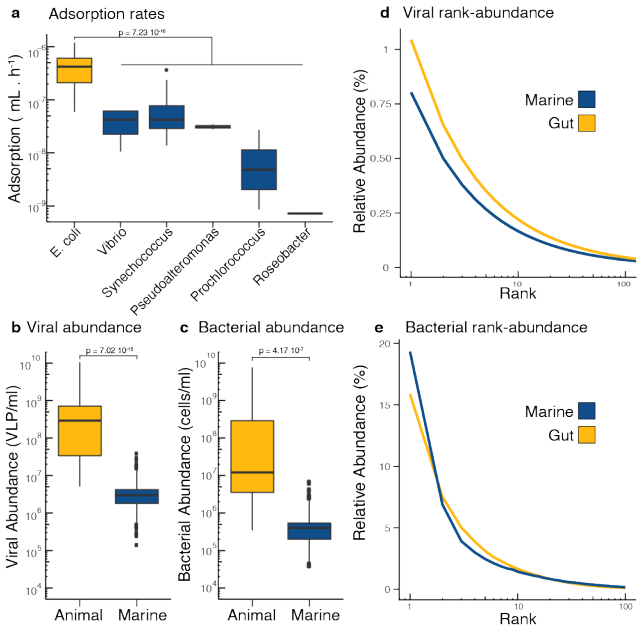


Figure 2. Comparison of percentage of lysogeny for lambda MOI and COI model. a) Relationship between the percentage of lysogeny and MOI (initial phage-to-bacteria ratio) in lambda-E*. coli* experiments (Fry, 1959; Kourilsky, 1973). The lines correspond to fitted Hill-Langmuir cooperation models of order n =1 (dashed), n=2 (solid), and n=3 (dotted). **b)** Percentage of lysogeny estimated from the coinfection probability model as a function of COI, Eq. (2) (solid black line). Hill-Langmuir model of order n=2 from panel **a**) as a function of MOI (grey solid line).

454



455 **Figure 3. Meta-analysis of COI parameters in marine and animal microbiomes. a)**
456 Adsorption rates from single phage-host pairs of bacterial hosts derived from the
457 mammalian gut (yellow) and marine environments (blue). The same color coding applies
458 to all panels. **b)** Viral abundances and **c)** bacterial abundances per ml in animal-associated
459 and marine environments. **d)** Rank abundance curve displaying the top 100 ranks of phage
460 genotypes identified in metagenomic data from human gut and marine samples. **e)** Rank
461 abundance curve displaying the top 100 ranks of bacterial genotypes identified in
462 metagenomic data from human gut and marine samples. The references for each original
463 study and the values for each datapoint are provided as Source Data 1 to 5.

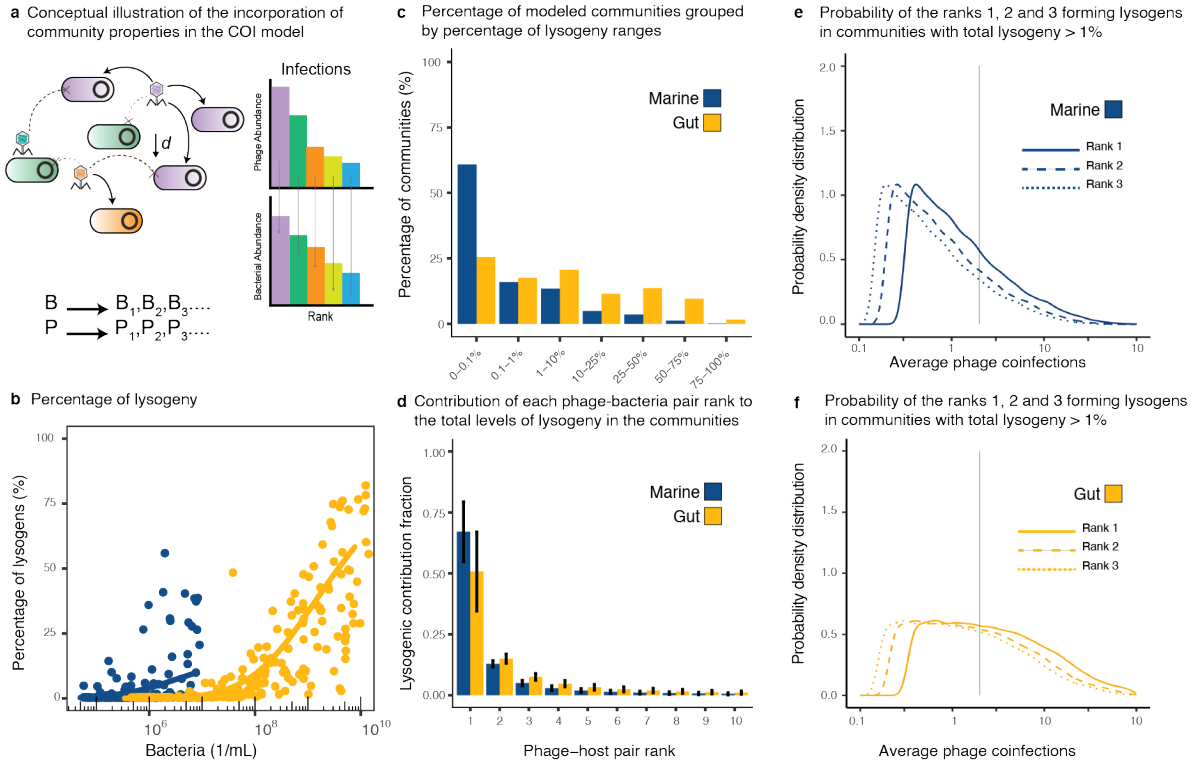


Figure 4. Lysogeny caused by coinfections in communities. a) Conceptual figure describing the implementation of the biophysical COI model with stochastic sampling of community parameters. Each Phage P_i infected each bacterium B_i according to their ranks in the community. The chances of encounter decreased with phage-host rank as a function of their absolute abundances in the community. **b)** Percentage of lysogenic cells in the bacterial community predicted to be formed through phage coinfections as a function of the total bacterial density. The data points are a subsample of 200 stochastic models out of 100,000 sampled models for each ecosystem. The solid lines represent generalized additive models (GAM) fitted to the full data set for marine (blue) and gut (yellow) ecosystems. **c)** Percentage of sampled communities displaying different ranges of lysogeny. **d)** Average contribution (with error bars corresponding to S.D.) from the top phage-bacteria pair ranks to the lysogenic pool. **e–f)** Probability distributions of average phage coinfections (COI) for the top ranks in communities with lysogeny above 1% for marine and gut ecosystems. The vertical line indicates COI = 2.

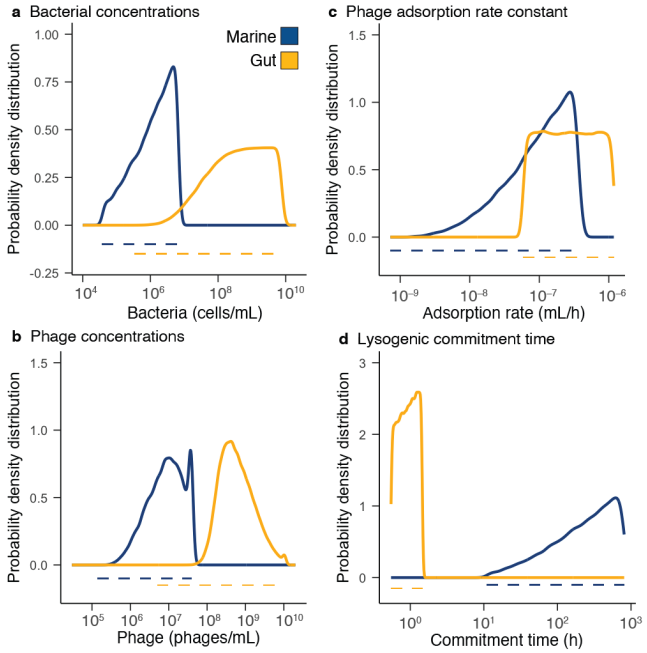


Figure 5. Ranges of parameter values leading to lysogeny caused by coinfection. Probability density distributions (solid lines) for bacterial abundances **(a)**, phage densities **(b)**, adsorption rates **(c)**, and lysogenic commitment times **(d)** in marine (blue) and gut (yellow) communities displaying lysogeny levels above 1%. The dotted lines indicate the range of parameters explored in the model and obtained from the meta-analysis of each ecosystem.

References

Almeida, A., Mitchell, A.L., Boland, M., Forster, S.C., Gloor, G.B., Tarkowska, A., et al. (2019) A new genomic blueprint of the human gut microbiota. *Nature* **568**: 499.

Anthenelli, M., Jasien, E., Edwards, R., Bailey, B., Felts, B., Katira, P., et al. (2020) Phage and bacteria diversification through a prophage acquisition ratchet. *BioRxiv* **in review**:

Avrani, S., Wurtzel, O., Sharon, I., Sorek, R., and Lindell, D. (2011) Genomic island variability facilitates Prochlorococcusvirus coexistence. *Nature*.

Barr, J.J., Auro, R., Furlan, M., Whiteson, K.L., Erb, M.L., Pogliano, J., et al. (2013) Bacteriophage adhering to mucus provide a non–host-derived immunity. *Proc Natl Acad Sci* **110**: 10771–10776.

Beckett, S.J., Demory, D., Coenen, A.R., Casey, J.R., Follett, C.L., Dugenne, M., et al. (2019) A day in the life of Prochloroccocus: Diel ecological oscillations of cyanobacteria, viruses and grazers in the North Pacific Subtropical Gyre. In *2019 ESA Annual Meeting (August 11--16)*. ESA.

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

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

Boyd, J.S.K. (1951) Observations on the relationship of symbiotic and lytic bacteriophage. *J Pathol Bacteriol* **63**: 445–457.

Brum, J.R., Hurwitz, B.L., Schofield, O., Ducklow, H.W., and Sullivan, M.B. (2015) Seasonal

time bombs: dominant temperate viruses affect Southern Ocean microbial dynamics. *ISME J* 1–13.

Casjens, S. (2003) Prophages and bacterial genomics: What have we learned so far? *Mol Microbiol* **49**: 277–300.

Cheng, H.H., Muhlrad, P.J., Hoyt, M. a, and Echols, H. (1988) Cleavage of the cII protein of phage lambda by purified HflA protease: control of the switch between lysis and lysogeny. *Proc Natl Acad Sci* **85**: 7882–7886.

Cobián Güemes, A.G., Youle, M., Cantú, V.A., Felts, B., Nulton, J., and Rohwer, F. (2016) Viruses as winners in the game of life. *Annu Rev Virol* **3**: 197–214.

Cohen, Y., Pollock, F.J., Rosenberg, E., and Bourne, D.G. (2013) Phage therapy treatment of the coral pathogen Vibrio coralliilyticus. *Microbiol Open* **2**: 64–74.

Cortes, M.G., Trinh, J.T., Zeng, L., and Balázsi, G. (2017) Late-Arriving Signals Contribute Less to Cell-Fate Decisions. *Biophys J* **113**: 2110–2120.

Coutinho, F.H., Rosselli, R., and Rodríguez-Valera, F. (2019) Trends of Microdiversity Reveal Depth-Dependent Evolutionary Strategies of Viruses in the Mediterranean. *mSystems* **4**:.

Coutinho, F.H., Silveira, C.B., Gregoracci, G.B., Thompson, C.C., Edwards, R.A., Brussaard, C.P., et al. (2017) Marine viruses discovered via metagenomics shed light on viral strategies throughout the oceans. *Nat Commun* **8**: 15955.

Deng, L., Gregory, A., Yilmaz, S., Poulos, B.T., Hugenholtz, P., and Sullivan, M.B. (2013) Contrasting life strategies of viruses that infect photo- and heterotrophic bacteria, as revealed by viral tagging. *MBio* **3**:.

Džunková, M., Low, S.J., Daly, J.N., Deng, L., Rinke, C., and Hugenholtz, P. (2019) Defining the

533 human gut host–phage network through single-cell viral tagging. *Nat Microbiol* **4**:
534 2192–2203.
535 Edwards, R.A., McNair, K., Faust, K., Raes, J., and Dutilh, B.E. (2016) Computational
536 approaches to predict bacteriophage–host relationships. *FEMS Microbiol Rev* **40**: 258–
537 272.
538 Fouts, D.E. (2006) Phage_Finder: Automated identification and classification of prophage
539 regions in complete bacterial genome sequences. *Nucleic Acids Res* **34**: 5839–5851.
540 Frazão, N., Sousa, A., Lässig, M., and Gordo, I. (2019) Horizontal gene transfer overrides
541 mutation in Escherichia coli colonizing the mammalian gut. *Proc Natl Acad Sci* **116**:
542 17906–17915.
543 Fry, B.A. (1959) Conditions for the infection of *Escherichia coli* with lambda phage and for
544 the establishment of lysogeny. *Microbiology* **21**: 676–684.
545 Furlan, M. (2009) Viral and microbial dynamics in the human respiratory tract.
546 Golding, I. (2016) Single-Cell Studies of Phage λ: Hidden Treasures Under Occam's Rug.
547 *Annu Rev Virol* **3**: 453–472.
548 Gregory, A.C., Zayed, A.A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., et al.
549 (2019) Marine DNA Viral Macro- and Microdiversity from Pole to Pole. *Cell* **177**: 1109-
550 1123.e14.
551 Haas, A.F., Fairoz, M.F.M.F.M., Kelly, L.W.L.W.L.W., Nelson, C.E.C.E., Wei, Y., Knowles, B., et al.
552 (2016) Global microbialization of coral reefs. *Nat Microbiol* **1**: 16042.
553 Heinken, A., Sahoo, S., Fleming, R.M.T., and Thiele, I. (2013) Systems-level characterization
554 of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes* **4**: 28–40.
555 Herskowitz, I. and Hagen, D. (1980) The lysis-lysogeny decision of phage lambda: explicit

556 programming and responsiveness. *Annu Rev Genet* **14**: 399–445.

557 Hoyles, L., Honda, H., Logan, N.A., Halket, G., La Ragione, R.M., and McCartney, A.L. (2008)
558 Isolation of Bacillus clausii, Bacillus licheniformis and other bacilli from human faecal
559 samples. *INRA–Rowett 2008 Gut Microbiome Funct Interact with Host Impact Environ*.

560 Huang, C., Zhang, Y., and Jiao, N. (2010) Phage resistance of a marine bacterium,
561 Roseobacter denitrificans OCh114, as revealed by comparative proteomics. *Curr*
562 *Microbiol* **61**: 141–147.

563 Jerlström Hultqvist, J., Warsi, O., Söderholm, A., Knopp, M., Eckhard, U., Vorontsov, E., et al.
564 (2018) A bacteriophage enzyme induces bacterial metabolic perturbation that confers
565 a novel promiscuous function. *Nat Ecol Evol* **2**: 1321–1330.

566 Jiang, S.C. and Paul, J.H. (1994) Seasonal and diel abundance of viruses and occurrence of
567 lysogeny/bacteriocinogeny in the marine environment. *Mar Ecol Prog Ser* **104**: 163–
568 172.

569 Johnson, R.M. (1968) Characteristics of a marine Vibrio-bacteriophage system. *J Arizona*
570 *Acad Sci* **5**: 28–33.

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

573 Kim, M.-S., Park, E.-J., Roh, S.W., and Bae, J.-W. (2011) Diversity and abundance of single-
574 stranded DNA viruses in human feces. *Appl Environ Microbiol* **77**: 8062–70.

575 Kirchman, D.L. (2016) Growth rates of microbes in the oceans. *Annu Rev Mar Sci* **8**: 285–
576 309.

577 Knowles, B., Silveira, C.B.C., Bailey, B.B.A., Barott, K., Cantu, V.V.A., Cobián-Güemes, A., et al.
578 (2016) Lytic to temperate switching of viral communities. *Nature* **531**: 466–470.

579 Kourilsky, P. (1975) Lysogenization by bacteriophage lambda. II -Identification of genes
580 involved in the multiplicity dependent processes. *Biochimie* **56**: 1511–1516.
581 Kourilsky, P. (1973) Lysogenization by bacteriophage lambda. *Mol Gen Genet* **122**: 183–
582 195.
583 Labonté, J.M., Swan, B.K., Poulos, B., Luo, H., Koren, S., Hallam, S.J., et al. (2015) Single-cell
584 genomics-based analysis of virus–host interactions in marine surface
585 bacterioplankton. *ISME J* **9**: 2386–2399.
586 Lang, M., Pleška, M., and Guet, C.C. (2020) Population dynamics of decision making in
587 temperate bacteriophages. *bioRxiv* 2020.03.18.996918.
588 Lauro, F.M., McDougald, D., Thomas, T., Williams, T.J., Egan, S., Rice, S., et al. (2009) The
589 genomic basis of trophic strategy in marine bacteria. *Proc Natl Acad Sci* **106**: 15527–
590 15533.
591 Lepage, P., Colombet, J., Marteau, P., Sime-Ngando, T., Doré, J., and Leclerc, M. (2008)
592 Dysbiosis in inflammatory bowel disease: a role for bacteriophages? *Gut* **57**: 424–425.
593 Levisohn, R., Moreland, J., and Nealson, K.H. (1987) Isolation and characterization of a
594 generalized transducing phage for the marine luminous bacterium Vibrio fischeri MJ-1.
595 *J Gen Microbiol* **133**: 1577–1582.
596 Lieb, M. (1953) The establishment of lysogenicity in *Escherichia coli*. *J Bacteriol* **65**: 642.
597 Luo, E., Aylward, F.O., Mende, D.R., and DeLong, E.F. (2017) Bacteriophage distributions and
598 temporal variability in the ocean's interior. *MBio* **8**: e01903-17.
599 Luo, E., Eppley, J.M., Romano, A.E., Mende, D.R., and DeLong, E.F. (2020) Double-stranded
600 DNA virioplankton dynamics and reproductive strategies in the oligotrophic open
601 ocean water column. *ISME J* 1–12.

602 Mann, N.H. (2003) Phages of the marine cyanobacterial picophytoplankton. *FEMS Microbiol*
603 *Rev* **27**: 17–34.
604 Marbouty, M., Baudry, L., Cournac, A., and Koszul, R. (2017) Scaffolding bacterial genomes
605 and probing host-virus interactions in gut microbiome by proximity ligation
606 (chromosome capture) assay. *Sci Adv* **3**: e1602105.
607 Maurice, C.F., Bouvier, T., Comte, J., Guillemette, F., and Del Giorgio, P.A. (2010) Seasonal
608 variations of phage life strategies and bacterial physiological states in three northern
609 temperate lakes. *Environ Microbiol* **12**: 628–41.
610 Mavrich, T.N. and Hatfull, G.F. (2019) Evolution of Superinfection Immunity in Cluster A
611 Mycobacteriophages. *Am Soc Microbiol*.
612 McKay, M.D., Beckman, R.J., and Conover, W.J. (1979) Comparison of three methods for
613 selecting values of input variables in the analysis of output from a computer code.
614 *Technometrics* **21**: 239–245.
615 Mende, D.R., Sunagawa, S., Zeller, G., and Bork, P. (2013) Accurate and universal delineation
616 of prokaryotic species. *Nat Methods*.
617 Minot, S., Bryson, A., Chehoud, C., Wu, G.D., Lewis, J.D., and Bushman, F.D. (2013) Rapid
618 evolution of the human gut virome. *Proc Natl Acad Sci* **110**: 12450–12455.
619 Minot, S., Sinha, R., Chen, J., Li, H., Keilbaugh, S.A., Wu, G.D., et al. (2011) The human gut
620 virome: inter-individual variation and dynamic response to diet. *Genome Res* **21**:
621 1616–1625.
622 Mirzaei, M., Khan, M.A.A., Ghosh, P., Taranu, Z.E., Taguer, M., Ru, J., et al. (2020)
623 Bacteriophages Isolated from Stunted Children Can Regulate Gut Bacterial
624 Communities in an Age-Specific Manner. *Cell Host Microbe*.

625 Muck, S., Griessler, T., Köstner, N., Klimiuk, A., Winter, C., and Herndl, G.J. (2014) Fracture
626 zones in the Mid Atlantic Ridge lead to alterations in prokaryotic and viral parameters
627 in deep-water masses. *Front Microbiol* **5**: 264.
628 Myhrvold, C., Kotula, J.W., Hicks, W.M., Conway, N.J., and Silver, P.A. (2015) A distributed
629 cell division counter reveals growth dynamics in the gut microbiota. *Nat Commun* **6**: 1–
630 10.
631 Parsons, R.J., Breitbart, M., Lomas, M.W., and Carlson, C.A. (2012) Ocean time-series reveals
632 recurring seasonal patterns of virioplankton dynamics in the northwestern Sargasso
633 Sea. *ISME J* **6**: 273–84.
634 Paul, J.H. (2008) Prophages in marine bacteria: dangerous molecular time bombs or the key
635 to survival in the seas? *ISME J* **2**: 579–89.
636 Paul, J.H. and Weinbauer, M. (2010) Manual of Aquatic Viral Ecology. 30–33.
637 Ptashne, M. (1986) A genetic switch: Gene control and phage lambda.
638 Reyes, A., Blanton, L. V, Cao, S., Zhao, G., Manary, M., Trehan, I., et al. (2015) Gut DNA
639 viromes of Malawian twins discordant for severe acute malnutrition. *Proc Natl Acad*
640 *Sci* **112**: 11941–11946.
641 Reyes, A., Haynes, M., Hanson, N., Angly, F.E., Heath, A.C., Rohwer, F., and Gordon, J.I. (2010)
642 Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* **466**:
643 334–338.
644 Reyes, A., Wu, M., McNulty, N.P., Rohwer, F.L., and Gordon, J.I. (2013) Gnotobiotic mouse
645 model of phage-bacterial host dynamics in the human gut. *Proc Natl Acad Sci U S A*
646 **110**: 20236–20241.
647 Russell, J.B. and Cook, G.M. (1995) Energetics of bacterial growth: balance of anabolic and

catabolic reactions. *Microbiol Rev* **59**: 48–62.

Schwartz, D.A. and Lindell, D. (2017) Genetic hurdles limit the arms race between Prochlorococcus and the T7-like podoviruses infecting them. *ISME J*.

Schwartz, M. (1976) The adsorption of coliphage lambda to its host: Effect of variations in the surface density of receptor and in phage-receptor affinity. *J Mol Biol* **103**: 521–536.

Shkoporov, A.N. and Hill, C. (2019) Bacteriophages of the Human Gut: The "Known Unknown" of the Microbiome. *Cell Host Microbe* **25**: 195–209.

Shotland, Y., Koby, S., Teff, D., Mansur, N., Oren, D. a, Tatematsu, K., et al. (1997) Proteolysis of the phage lambda CII regulatory protein by FtsH (HflB) of Escherichia coli. *Mol Microbiol* **24**: 1303–1310.

Silveira, C.B., Coutinho, F.H., Cavalcanti, G.S., Benler, S., Doane, M.P., Dinsdale, E.A., et al. (2020) Genomic and ecological attributes of marine bacteriophages encoding bacterial virulence genes. *BMC Genomics* **21**: 126.

Silveira, C.B., Luque, A., Roach, T.N.F., Villela, H., Barno, A., Green, K., et al. (2019) Biophysical and physiological processes causing oxygen loss from coral reefs. *Elife* **8**:.

Silveira, C.B. and Rohwer, F.L. (2016) Piggyback-the-Winner in host-associated microbial communities.

Stoddard, L.I., Martiny, J.B.H., and Marston, M.F. (2007) Selection and characterization of cyanophage resistance in marine Synechococcus strains. *Appl Environ Microbiol* **73**: 5516–5522.

Sunagawa, S., Mende, D.R., Zeller, G., Izquierdo-Carrasco, F., Berger, S.A., Kultima, J.R., et al. (2013) Metagenomic species profiling using universal phylogenetic marker genes. *Nat Methods*.

671 Thingstad, T.F. (2000) Elements of a theory for the mechanisms controlling abundance,
672 diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol*
673 *Oceanogr* **45**: 1320–1328.
674 Touchon, M., Bernheim, A., and Rocha, E.P.C. (2016) Genetic and life-history traits
675 associated with the distribution of prophages in bacteria. *ISME J* **10**: 2744–2754.
676 Traits, P., Castillo, D., Alvise, P.D., Xu, R., Zhang, F., and Middelboe, M. (2017) Comparative
677 Genome Analyses of Vibrio anguillarum Strains Reveal a Link with. **2**: 1–14.
678 Trinh, Jimmy T., Székely, T., Shao, Q., Balázsi, G., and Zeng, L. (2017) Cell fate decisions
679 emerge as phages cooperate or compete inside their host. *Nat Commun* **8**: 1–13.
680 Trinh, Jimmy T, Székely, T., Shao, Q., Balázsi, G., and Zeng, L. (2017) Cell fate decisions
681 emerge as phages cooperate or compete inside their host.
682 Vandeputte, D., Falony, G., Vieira-Silva, S., Tito, R.Y., Joossens, M., and Raes, J. (2016) Stool
683 consistency is strongly associated with gut microbiota richness and composition,
684 enterotypes and bacterial growth rates. *Gut* **65**: 57–62.
685 Waterbury, J.B. and Valois, F.W. (1993) Resistance to co-occurring phages enables marine
686 Synechococcus communities to coexist with cyanophages abundant in seawater. *Appl*
687 *Environ Microbiol* **59**: 3393–3399.
688 Weitz, J.S., Beckett, S.J., Brum, J.R., Cael, B.B., and Dushoff, J. (2017) Lysis, lysogeny and
689 virus–microbe ratios. *Nature* **549**: E1–E3.
690 Weitz, J.S., Stock, C.A., Wilhelm, S.W., Bourouiba, L., Coleman, M.L., Buchan, A., et al. (2015) A
691 multitrophic model to quantify the effects of marine viruses on microbial food webs
692 and ecosystem processes. *ISME J* **9**: 1352–1364.
693 Zeng, L., Skinner, S.O., Zong, C., Sippy, J., Feiss, M., and Golding, I. (2010) Decision Making at

694 a Subcellular Level Determines the Outcome of Bacteriophage Infection. *Cell* **141**: 682–
695 691.
696 Zhao, Y., Temperton, B., Thrash, J.C., Schwalbach, M.S., Vergin, K.L., Landry, Z.C., et al. (2013)
697 Abundant SAR11 viruses in the ocean. *Nature* **494**: 357–60.
698