**Supplementary Information**

**Phage mobility is a core determinant of phage-bacteria coexistence in biofilms**

Running title: Phage diffusion in bacterial biofilms

Matthew Simmons1, Knut Drescher2,3, Carey D. Nadell2,4\*†, Vanni Bucci1\*†

1 Department of Biology, Program in Biotechnology and Biomedical Engineering, University of Massachusetts Dartmouth, N. Dartmouth, MA 02747, USA

2 Max Planck Institute for Terrestrial Microbiology, D-35043 Marburg, Germany

3 Department of Physics, Philipps University Marburg, D-35032 Marburg, Germany

**4** Current Address: Department of Biological Sciences, Dartmouth College, Hanover, NH 03755, USA

**\* Equal contribution**

**† Correspondence to:** Carey.D.Nadell@dartmouth.edu**,** vanni.bucci@umassd.edu

**Supplementary Methods**

**Detailed Description of the Simulation Framework**

Computation of nutrient concentration profiles:

The framework can incorporate an arbitrary number of user-defined solutes, but here we only model a single limiting nutrient according to a simple reaction-diffusion system $∂\_{t}N=D\_{N}∇^{2}N+R(N)$, with *t* denoting time, $N$ the local nutrient concentration, $D\_{N}$the diffusivity of nutrients, $∂\_{t}N$ the change in nutrient concentration over time, and $∇^{2}$ the Laplace operator. Nutrient consumption by bacteria is defined by the Monod-type kinetic reaction $R\left(N\right)=-Y\frac{μ\_{S}N}{N+K\_{N}^{S}}S$, where $Y$ is the growth yield, $μ\_{S}$ is the bacterial maximum growth rate, $K\_{N}^{S}$ is the half saturation constant,$ S$is the bacterial active biomass density. The nutrient concentration is held constant at the upper boundary of the system, and there is no nutrient flux permitted through the bottom boundary (substratum). No nutrients are supplied from the substratum (Simulation parameters and their values are given in Table S1). This growth geometry is inspired by the standard flow channel biofilm growth assay, where biofilms are grown on a glass or plastic substratum over which a nutrient medium is flown.

 Similarly to previous work [1], we assume that diffusion of the nutrient solute is instantaneous relative to bacterial and phage-related processes, and as such it is solved at quasi-steady state, i.e. $∂\_{t}N=D∇^{2}N +R\left(N\right)=0$. Based on this assumption, we compute $N (x, y, t)$ by using Newton’s method to linearize $∂\_{t}N=0$ as:

$(D∇^{2}+f'(N)) Δ\_{N}= -f(N)$,

where

$f\left(N\right)=D∇^{2}N-Y\frac{μ\_{S}N}{N+K\_{N}^{S}}S$ ,

$f^{'}\left(N\right)=∂\_{N} f\left(N\right)=YS\frac{μ\_{S}K\_{N}^{S}}{(K\_{N}^{S}+N)^{2}}$, and

$Δ\_{N}=N\_{i+1}-N\_{i}$ ($i$ is the iteration counter of Newton’s method).

Once linearized, we use an algebraic multigrid solver [2] to compute the change in nutrient concentration at each Newton-iteration. Initial conditions for nutrient concentration are $N \left(x, y, t=0\right)=N\_{max}× (1-\frac{y}{y\_{max}})$, with $y \in (0,y\_{max})$ corresponding to the vertical coordinate of the system ($y=0$ at the system substratum, $y=y\_{max}$ at the system ceiling).

Computation of bacterial biomass dynamics:

The model variables that describe bacterial population dynamics over space and time are the active susceptible biomass $(S)$, the infected biomass $(I)$, and the inert biomass $(D)$. From the active biomass, $S$, the number of active bacteria$n\_{S}$ (used in single cell-dependent processes) is calculated as $n\_{S}(x,y)=S\left(x,y\right)×dl^{3}/m\_{S}$, where $m\_{S}$ denotes the specific mass of a bacterium. The order of operations for biofilm computations is the following:

1. **Biomass growth and decay.**
	1. Solve the bacterial growth equations for the active biomass, $∂\_{t}S=(\frac{μ\_{S}N}{N+K\_{N}^{S}}-δ\_{d}) S$, and the conversion from active to inert biomass, $∂\_{t}D=δ\_{d}S$, where $δ\_{d}$ is the biomass decay rate.
	2. Update active and inert biomass values in each grid element accordingly.
2. **Shoving.** For every grid element that meets the condition $S\geq S\_{max}$ (with $S\_{max}$ being the maximum active biomass concentration allowed in a grid element):
	1. Find the nearest empty grid point satisfying ($S+I+D\leq θ$)
	2. Compute a line towards the grid point determined in step (2-a) using the Bresenham algorithm.
	3. Shift intervening biomass along the line defined in step (2-b) such that biomass in the focal grid node can be split between the current node and a newly empty neighbor node.
3. **Erosion**.
	1. Calculate an erosion force – that is, biomass reduction due to the effect of shear stress – proportional to the square of the distance from the substratum, $F\_{e}=k\_{e} y^{2}$, then solve for the time to erosion ($t\_{e}$) using the fast marching method [3] . The amount of eroded biomass is proportional to the time to erosion divided by the simulation time step, $t\_{e}/dt$. This rule assumes active flow in the system and implements shear-dependent erosion that is strongest for biomass on the biofilm exterior.
	2. Any grid element with total biomass $S+I+D<\frac{2}{3}m\_{s}$ is set to an empty state.
4. **Sloughing**
	1. For every grid point, determine if there is a chain of neighboring biomass-occupied nodes continuous with the substratum. If any biomass is found unattached to the rest of the biofilm in this manner, it is removed.

Host cell lysis and phage propagation:

After biofilms have grown for a defined period of time (1.5 days in physical units for results presented in the main text), we implement phage exposure in a single pulse, introducing one phage particle to each empty grid element above the outer edge of the biofilm. This inoculation method was used to apply relatively strong phage exposure and minimize the influence of chance effects on the initial establishment of infection. The framework tracks infected bacteria and phages *via* infected bacterial biomass $(I)$, the number of phages ($n\_{p}$) and the remaining incubation time for each infected cell ($τ\_{r}$).

After infection, the order of operations for phage processes is the following:

1. **Infected cell lysis**. Decrement $τ\_{r}$ by $dt$ for each infected cell. If a cell lyses (those with $τ\_{r}\leq 0$.), the corresponding infected biomass is subtracted from the grid element and $β$ phages (the burst size) are added to that element.
2. **Phage diffusion** (see details below)
3. **Phage localization**. This step involves phage-cell adsorption, phage dispersal by advection, and phage decay.
	1. For each phage in a grid element with active biomass, calculate whether or not it is adsorbed to a host cell. The adsorption probability is determined as $1- e^{-γn\_{S}dt}$, where the rate is given by the product of infectivity $γ$, the local number of active susceptible bacteria $n\_{S}$, and the integration time step $dt$. Adsorptions cause a conversion of active bacterial biomass (*S*) to infected bacterial biomass (*I*), equal to the ratio of phage adsorptions to the node bacterial count, with a maximum of unity. The infected biomass is assigned an incubation period $τ\_{r}=τ$.
	2. Phages residing in the outer border of the biofilm that are not adsorbed to biomass, are immediately considered ‘detached’ and are removed from the system; this rule implements advective removal of phages that are not biofilm-attached.
	3. Phages decay by natural mortality, with probability $1-e^{-δ\_{p}dt}$, with $δ\_{p}$ being the phage decay rate.

Phage diffusion:

To compute the movement of phages via Brownian motion, we use the analytical solution to the diffusion equation of a Dirac delta function at each grid node to build a probability distribution from which to resample the phage locations.

1. First, we calculate the distance $p\_{ij}$ to each grid element $i$ from each grid element $j$ that contains at least one phage. To model implicitly the effect of bacterial matrix on phage movement, the effective distance between two neighboring elements is assumed to depend on the presence of bacterial biomass (see detailed explanation below). Specifically, when moving from:
	1. an empty grid node into a neighboring empty grid node, the effective distance is equal to $dl$;
	2. an empty grid node into a biomass-occupied grid node, the effective distance is equal to $dl$;
	3. a biomass-occupied node into an empty node, the effective distance is equal to $dl × Z\_{P}$;
	4. a biomass-occupied grid node into another biomass-occupied node, the effective distance is equal to $dl × 2 × Z\_{P}$ (see note below).
2. For each grid element $j$ that contains at least one phage, we perform the following steps:
	1. Estimate the probability of diffusion into each neighboring grid element $i$ by computing the analytical solution of the diffusion equation of a Dirac function [4], using distances calculated in step (1).
	2. Multiply the resulting distribution by the number of phages that are present in element $j$.
	3. Sum the phage distributions obtained by performing step (2a-b) in each grid element, element-wise, and then normalize across all grid elements.
	4. Sample the total number of phages in the system from the distribution calculated in (c) and deposit phages in the system accordingly.

**Notes on diffusion of particles across the biofilm/liquid interface**

The term $Z\_{P}$, which we call the phage impedance, captures the degree to which biofilms block diffusion of phage particles. Our default implementation assumes that it is easier for phages to cross from the liquid into the biofilm than *vice versa*. We incorporated this feature because theory predicts that – to reach any particular point of distance $dl$ away from a starting position – a random walker in two dimensions must travel $dl^{2}$ steps in open space, with unit step length [5]. For a random walk in a maze, however, to reach any particular point the same distance $dl$ away, the walker must instead move on the order of $dl^{3} $steps. Taking two adjacent grid cells, one constituted by a random maze (e.g. biofilm biomass) and the other one clear (e.g. liquid phase), in order to reach the boundary between the two grids that is distance $dl$ away, the phage starting from the liquid bulk must only walk $dl^{2} $steps, whereas the phage exiting from the biomass-containing grid will need to take $dl^{3} $steps [6]. Consequently, in a fixed amount of time, the number of phages that start in the liquid bulk and diffuse into the biofilm is expected be greater than the number of phages inside the biofilm 'escaping' to the liquid bulk, all else being equal. While the default implementation of our simulations includes this effect, we also relax this assumption ad compare the original models with those in which diffusion is equally likely in either direction across biofilm/liquid interface (see Main Text Figure 4, and Supplementary Figure S4).

**Model and code benchmarking against non-spatial ODE system**

To check that the core mechanisms of our simulations are sound, we performed an analysis in which a well-mixed version of the phage-biofilm framework (see below for details) was compared against a traditional system of ordinary differential equations describing the dynamics of substrate, bacteria and phages.

ODE model: The system of ordinary differential equations solved for comparison to our well-mixed implementation of the phage-biofilm framework is the following:

$$\left\{\begin{array}{c}\frac{dN}{dt}=δ\_{N}\left(N-N\_{0}\right)-Y\frac{μ\_{S}N}{N+K\_{N}^{S}}S \\\frac{dS}{dt}=\left(\frac{μ\_{S}N}{N+K\_{N}^{S}}-γP-δ\_{D} \right)S\\\frac{dI}{dt}=γPS -αI\\\frac{dP}{dt}=αβI-γ νPS-δ\_{P}P\end{array}\right.$$

where, in addition to the parameters already introduced to describe the framework (see main text), we introduce $δ\_{N}$ as the inflow rate for the growth substrate, $α$ the infected cell lysis rate, $β$ the yield of phages per unit infected biomass, and $ν$ the conversion factor describing the amount of phage biomass lost during infection of a susceptible bacterium.

We then proceeded to a partial non-dimensionalization of the system by introducing the following quantities $\overbar{μ}\rightarrow Yμ$, $\overbar{N}\rightarrow N/N\_{0}$, $\overbar{S}\rightarrow \frac{S}{S\_{max}},$ $\overbar{I}\rightarrow \frac{I}{S\_{max}}$, $\overbar{P}\rightarrow \frac{P}{P\_{0}}V$,$ \overbar{γ}\rightarrow γ \frac{P\_{0}}{V}$, $ \overbar{ν}\rightarrow ν\frac{S\_{max}}{P\_{0}}V$, $\overbar{β}\rightarrow β\frac{S\_{max}}{P\_{0}}V$. Here, $P\_{0}$ is the mass of an individual phage, $S\_{max}$ is the maximum concentration of active biomass and $V$ is the total volume of the system. The non-dimensionalization by means of these three quantities was chosen in order for the ODE model to be parameterized identically to the biofilm framework, allowing us to use exactly the same constants as they have identical meaning in both contexts.

Our new system of differential equations becomes the following,

$$\left\{\begin{array}{c}\frac{d\overbar{N}}{dt}=1-\overbar{N}-\overbar{μ}\frac{S\_{max}}{N\_{0}}\frac{\overbar{N}}{\overbar{N}+\frac{K\_{N}^{S}}{N\_{0}}}\overbar{S}\\\frac{d\overbar{S}}{dt}=\left(\frac{\overbar{μ}Y\overbar{N}}{\overbar{N}+K\_{N}^{S}/N\_{0}}-\overbar{γ}\overbar{P}-δ\_{D} \right)\overbar{S}\\\frac{d\overbar{I}}{dt}=\overbar{γ}\overbar{P}\overbar{S} -α\overbar{I}\\\frac{d\overbar{P}}{dt}=α\overbar{β}\overbar{I}-\overbar{γ} \overbar{ν}\overbar{P}\overbar{S}-δ\_{P}\overbar{P}\end{array}\right.$$

such that $\overbar{γ}$ now exactly matches the infectivity constant used in the phage-biofilm framework, and $S\_{max} $and $N\_{0}$ which are both explicitly considered in the framework.

To constrain $\overbar{β}$ to have the same effect of the burst size parameter ($β=100)$ in the phage-biofilm framework (see Main Text), we use the following equation: $\overbar{β}=100×\frac{S\_{MAX}}{m\_{S}}V$ where $m\_{S}$ is the mass of a single bacterial cell. Similarly, we set $\overbar{ν}=1×\frac{S\_{MAX}}{m\_{S}}V$ given that it takes the equivalent of one phage to infect one dimensionless unit of biomass.

The ode model was solved with Python *scipy* using the Runge-Kutta integration algorithm of order 4 ([https://docs.scipy.org/doc/scipy-0.19.0/reference/generated/scipy.integrate.ode.html)](https://docs.scipy.org/doc/scipy-0.19.0/reference/generated/scipy.integrate.ode.html%29).

Parameters for the ODE model are identical to those in Table 1 (Main Text) for the phage-biofilm framework with the difference that $\overbar{β}=100×8e5$ and set $\overbar{ν}=1×8e5$ as explained above. We set also $α=0.02/day$ which is equivalent to having an incubation time $t=28.8 mins$, as was used in the phage-biofilm framework described in the previous section here, and in the main text.

To benchmark the biofilm framework against the ODE model, the framework was simplified as follows, implementing a well-mixed system:

* Single 1X1 grid
* Phage diffusion off
* Phage detachment off
* Biofilm detachment off
* Erosion off
* Shoving off
* Substrate diffusion is calculated using ode above, with no dependency on local spatial concentration.



**Figure S1.** Example benchmark comparison of a non-spatial version of our biofilm-phage simulation framework (black lines) with a standard ODE model (red line) using the same parameters for bacterial and phage proliferation. Biofilms are grown for 1.5 simulated days before introduction of phages. This is one of many test runs we performed, and we note that under the same parameter constraints used for the spatial simulations in the main text, the only outcome if the well-mixed system is bacterial extinction. The slight delay in population decline of the simulations relative to the ODE model is due to the phage incubation phase, which is implemented in the simulations but not in the ODE model (for which infected cells are instantaneously converted to new phage biomass).



**Figure S2.** A comparison of the phage-biofilm interaction distributions for Zp = 10, 15, with variable phage burst sizes (figure format as in Figure 3 of the main text). This analysis revealed that stable state distributions were not qualitatively altered past phage burst sizes of approximately 100, which is at or below the reported mean burst size for many lytic phages in the literature. A fixed burst size of 100 was used for simulations in the mean text and elsewhere in the supplement.



**Figure S3**. Steady states of biofilm-phage population dynamics as a function of nutrient availability and phage infection rate. Each point in each heatmap summarizes the outcome distributions of ~30 simulations, corresponding to ~3 replicates for 11 different initial conditions. Different initial conditions were obtained by varying the time – and thus the biofilm population size – at which phages were introduced to the system (between 0.1 and 3 days after the start of biofilm growth). The plots show the distribution of simulation outcomes for the combination of nutrient availability and per-host-encounter phage infection probability specified on the vertical and horizontal axes, respectively. Phage extinction (biofilm survival) is denoted by blue, biofilm-phage coexistence is denoted by tan, and biofilm death is denoted by orange. The sweep was performed for two values of phage impedance *Zp* (10 and 15), also examined in detail in Figure 3 of the main text. We found very good agreement with the distribution of steady states presented in the main text, confirming that our results are robust to variation in initial conditions.



**Figure S4.** A comparison of the phage-biofilm interaction distributions for Zp = 2, 5, 10, 15, with symmetric and asymmetric diffusion of phages across the biofilm/liquid interface. Default simulations in the paper assume that phages can enter the biofilm phase more easily than they can exit, due to the longer effective random walk path that occurs through meshes, relative to open liquid (as shown in the left-hand column of heatmaps here). When this assumption is relaxed, and phages diffuse equally easily from the biofilm to the liquid and *vice versa*, then the system tends to be pushed more strongly toward coexistence of bacteria and phages for values of Zp of 10 and higher.

**Supplementary Video Files**

**Video SV1:** A simulation corresponding to biofilm eradication when the biofilm is growing as a uniform front. Uninfected bacteria are shown in red, infected bacteria are shown in blue, and phages are shown in black. For this and other videos, the color gradient is analogous to that in Figure 1 of the main text.

**Video SV2:** A simulation corresponding to biofilm eradication after biofilms have produced tower-like structures and a spatially heterogeneous front. This pattern of biofilm morphology occurs more readily as nutrients become scarce; here phage-induced death occurs mostly at the flanks of these towers, which are sloughed from the biofilm front. Uninfected bacteria are shown in red, infected bacteria are shown in blue, and phages are shown in black.

**Video SV3:** A simulation corresponding to long-term coexistence of bacteria and phages, in which both populations remain at constant or nearly unchanging size. Uninfected bacteria are shown in red, infected bacteria are shown in blue, and phages are shown in black.

**Video SV4:** A simulation corresponding to long-term coexistence of bacteria and phages, with oscillating population size for bacteria and phages. This pattern corresponds to growth, phage-induced sloughing, and re-growth of biofilm towers over time. Uninfected bacteria are shown in red, infected bacteria are shown in blue, and phages are shown in black.

**Video SV5:** A simulation corresponding to phage extinction due to low bacterial growth and consequent low likelihood of phage-bacterial encounter. Phages are not able to infect due to the low density of available bacterial targets and are therefore swept away due to erosion. Uninfected bacteria are shown in red, infected bacteria are shown in blue, and phages are shown in black.

**Video SV6:** A simulation corresponding to phage extinction due to extremely rapid biofilm growth and phage expulsion from the biofilm biomass; this result occurred only when biofilm growth rate was increased to a point that is empirically unrealistic. Uninfected bacteria are shown in red, infected bacteria are shown in blue, and phages are shown in black.

**References**

[1] Xavier, J.B., Picioreanu, C. & van Loosdrecht, M.C.M. 2005 A framework for multidimensional modelling of activity and structure of multispecies biofilms. *Environ Microbiol* **7**, 1085-1103. (doi:10.1111/j.1462-2920.2005.00787.x).

[2] N, B., L, O. & J, S. 2013 PyAMG: Algebraic Multigrid Solvers in Python.

[3] Xavier, J.d.B., Picioreanu, C. & van Loosdrecht, M.C.M. 2005 A general description of detachment for multidimensional modelling of biofilms. *Biotechnology and bioengineering* **91**, 651-669.

[4] Lapidus, L. & Pinder, G.F. 2011 *Numerical solution of partial differential equations in science and engineering*, John Wiley & Sons.

[5] McCrea, W. & Whipple, F. 1940 Random Paths in Two and Three Dimensions. *Proceedings of the Royal Society of Edinburgh* **60**, 281-298.

[6] Motwani, R. & Raghavan, P. 1995 *Randomized Algorithms*. Cambridge, UK, Cambridge University Press.