Hydrodynamic flow and concentration gradients in the gut enhance neutral bacterial diversity

Darka Labavić[1], Claude Loverdo[1,*,†] and Anne-Florence Bitbol[2,3,*,†]

[1]*Sorbonne Université, CNRS, Institut de Biologie Paris-Seine, Laboratoire Jean Perrin (UMR 8237), F-75005 Paris, France,*

[2]*Institute of Bioengineering, School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland,*
[3]*SIB Swiss Institute of Bioinformatics, CH-1015 Lausanne, Switzerland,*
[*]*Equal contribution,*
[†]*E-mail addresses: claude.loverdo@sorbonne-universite.fr; anne-florence.bitbol@epfl.ch.*

November 22, 2021

Abstract

The gut microbiota features important genetic diversity, and the specific spatial features of the gut may shape evolution within this environment. We investigate the fixation probability of neutral bacterial mutants within a minimal model of the gut that includes hydrodynamic flow and resulting gradients of food and bacterial concentrations. We find that this fixation probability is substantially increased compared to an equivalent well-mixed system, in the regime where the profiles of food and bacterial concentration are strongly spatially-dependent. Fixation probability then becomes independent of total population size. We show that our results can be rationalized by introducing an active population, which consists of those bacteria that are actively consuming food and dividing. The active population size yields an effective population size for neutral mutant fixation probability in the gut.

Significance statement

The human body harbors numerous and diverse bacteria, the vast majority of which are located in the gut. These bacteria can mutate and evolve within the gut, which is their natural environment. This can have important public health implications, e.g. when gut bacteria evolve antibiotic resistance. The gut features specific characteristics, including hydrodynamic flow and resulting gradients of food and bacterial concentrations. How do these characteristics impact the evolution and diversity of gut bacteria? We demonstrate that they can substantially increase the probability that neutral mutants reach high proportions and eventually take over the population. This is because only a fraction of gut bacteria is actively dividing. Thus, the specific environment of the gut enhances neutral bacterial diversity.

Introduction

In the human body, bacteria are approximately as numerous as human cells, and about 99% of these bacteria are located in the digestive tract [1]. The gut microbiota is very diverse, and collectively harbors more genes than there are human genes [2]. One source of this genetic diversity is evolution occurring within the gut, which is the natural environment of these bacteria. Such evolution can have important public health implications, as the gut can constitute a reservoir of antibiotic resistance, both in humans and in farm animals [3]. How does the environment in the gut affect the evolution of bacteria? A crucial feature of the gut is the flow of its contents along its main axis, and the associated gradients of concentration of food and bacteria. Going downstream along this axis, food is first ingested, then simple nutrients are absorbed by the body, next more complex molecules are broken down by bacteria, and eventually what remains of the food exits the system, together with many bacteria, which make up from a quarter to half of fecal mass [4]. These features yield a very particular spatial structure that can impact the evolution of bacteria.

Evolutionary models that investigate population spatial structure generally consider discrete patches of population with migrations between them, and the same environment in each of them [5–13]. Complex spatial structures are investigated through models on graphs where each individual [14–17] or each patch of population [18–21] occupies a node of the graph. Population structure can impact the rapidity of adaption [22–28] because local competition can allow the maintenance of larger genetic diversity. In simple population structures where migration is symmetric between patches [5, 6], the fixation probability of a mutant is unaffected by population structure [7, 8], unless extinctions of patches occur [11]. However, more complex population

structures with asymmetric migrations can impact the fixation probabilities of beneficial and deleterious mutants [13, 14, 21]. In the case of the gut, the flow can be viewed as yielding asymmetric migrations, but the system is continuous. In large-scale turbulent systems, hydrodynamic flow has been shown to strongly impact fixation probabilities and fixation times [29–31]. In addition, environmental gradients, e.g. of antibiotic concentration, can strongly impact evolution [32–35]. How do population structure, hydrodynamic flow and gradients shape the evolution of bacteria in the gut microbiota?

Here, we propose a minimal model of evolution of bacteria in the gut. Because most bacteria in the human digestive tract are located in the bulk of the colon lumen [1, 36], we focus on this compartment. Since most bacteria in the digestive tract have no self-motility [37, 38], we consider that they are carried passively with the digesta. The motion of the digesta is complex, but it was shown in Refs. [39, 40] that it can be approximated as a one-dimensional flow with net velocity and effective diffusion representing mixing. Within this model of the gut that includes hydrodynamic flow and resulting gradients of food and bacterial concentrations, we ask how the fixation probability of a neutral mutant compares with that in an equivalent well-mixed chemostat. We find that the structure of the gut can increase this fixation probability, specifically in the regime where the profiles of food and bacterial concentration are strongly spatially-dependent. In this regime, fixation probability becomes independent of total population size, in stark contrast with a well-mixed population, where fixation probability is inversely proportional to total population size [41, 42]. We show that this behavior can be understood by introducing the notion of active population, which corresponds to the fraction of the bacterial population that is actively consuming food and dividing.

Model and methods

Because the majority of bacteria in the human digestive tract are in the colon [36], we focus on this compartment. Within the colon, there are marked differences between bacteria associated to mucus and bacteria in the digesta, i.e. in the bulk of the colon lumen [36]. The latter constitute the majority of bacteria in the colon. Indeed, the surface area of the large intestine, including its folds, is about 2 square meters [43], while the mucus layer is about $100 - 300$ μm thick [44], and typically comprises a few 10^8 bacteria per milliliter in healthy samples [45], which leads to an order of magnitude of 10^{11} bacteria associated to mucus. This number is small compared to the total colon content, which is around 10^{14} bacteria [1]. Since mucus-associated bacteria constitute a small minority in the colon, and since their spatial structure and migration patterns are not well characterized, we focus on the bacteria present in the bulk of the colon lumen, and do not model the mucus layer. Henceforth, we refer to the colon lumen by "gut" for simplicity.

The dynamics of wild-type and mutant bacteria and food in the gut is described through three concentration fields, of food F, wild-type bacteria B and mutant bacteria M, based on the description of the coupled dynamics of food and bacteria (without mutants) developed in Ref. [39]. The gut is represented by a tube of length L and cross-section with surface area S (Figure 1A). In addition to this cylindrical symmetry, we neglect radial variations, and are left with a one dimensional model along the x axis, specifically a segment of length L. We assume a constant inflow of nutrients at the entrance of this gut segment and no inflow of bacteria. At the exit of the gut, we assume that there is a free outflow of both nutrients and bacteria. The dynamics is affected by the constant flow velocity v, by the mixing due to different mechanisms e.g. peristaltic movement, which is modelled by effective diffusion with diffusion coefficient D, and by the harvesting of the food by bacteria, which is described by a Hill-type function with Monod constant k, and is coupled to their growth which has maximal rate r. This leads to the following coupled partial differential equations:

$$\frac{\partial F}{\partial t} = D\frac{\partial^2 F}{\partial x^2} - v\frac{\partial F}{\partial x} - \frac{r}{\alpha}\frac{(B+M)F}{k+F}, \tag{1a}$$

$$\frac{\partial B}{\partial t} = D\frac{\partial^2 B}{\partial x^2} - v\frac{\partial B}{\partial x} + r\frac{BF}{k+F}, \tag{1b}$$

$$\frac{\partial M}{\partial t} = D\frac{\partial^2 M}{\partial x^2} - v\frac{\partial M}{\partial x} + r\frac{MF}{k+F}, \tag{1c}$$

with boundary conditions

$$-D\frac{\partial[F;B;M]}{\partial x}(x=0) + v[F;B;M](x=0) = [vF_{\text{in}};0;0]\,, \tag{2a}$$

$$-D\frac{\partial[F;B;M]}{\partial x}(x=L) = [0;0;0], \tag{2b}$$

where $[F;B;M]$ denotes a vector. Here vF_{in} is the food inflow at the entrance of the gut segment, while α denotes the yield of the conversion from food to bacteria. Note that there is zero inflow of bacteria, in agreement with

observations that bacterial concentration in the smaller intestine is orders of magnitude smaller [36, 39, 40]. The boundary conditions at $x = L$ cancel the diffusive flux, corresponding to free outflow toward the downstream part of the colon.

In our study of the fate of mutants appearing in the gut, initial conditions are

$$F(t = 0, x) = F^*(x), \tag{3a}$$

$$B(t = 0, x) = B^*(x), \tag{3b}$$

$$M(t = 0, x) = \begin{cases} M_0, |x - x_{\mathrm{M}}| \leq \Delta x/2, \\ 0, |x - x_{\mathrm{M}}| > \Delta x/2, \end{cases} \tag{3c}$$

where F^* and B^* represent the steady state of system 1 without mutant bacteria, while $x_{\mathrm{M}} \in (0, L)$ is the position in the gut where the mutant appears, while Δx is a short length, taken equal to the spatial discrete step in our numerical resolutions, and $M_0 \ll B(x_{\mathrm{M}})$ is the initial local concentration of mutant at this location. In practice, M_0 is set through $N_{\mathrm{M}} = M_0 S \Delta x$, where S is the surface area of the section of the gut, so that the total number N_{M} of mutants introduced in the system is always the same, and our results do not depend on Δx as long as it is small compared to the length scale over which concentrations vary.

The partial differential equations in Eqs. 1 with boundary conditions in Eqs. 2 and initial conditions in Eqs. 3 were solved numerically (Supplementary material Section S1, and code publicly available [46]).

Results

Spatial dependence of the steady-state bacterial concentration

Our aim is to study the fate of neutral mutants appearing in the gut, starting from initial conditions where the concentrations of food and wild-type bacteria are at steady state (see Eqs. 3). Therefore, we start by describing the steady-state profiles of food and wild-type bacteria in the mutant-free gut.

Steady-state solutions of the spatial model described by Eqs. 1 can strongly depend on the spatial coordinate x for some values of flow velocity v, and effective diffusion constant D, as exemplified by Figure 1B. Such strong spatial dependence is relevant in the ascending colon [40], which is our focus here. These strongly spatial profiles resemble Fisher waves, and indeed, the steady-state equation describing our system has the same form as the equation satisfied by a traveling wave in a Fisher-KPP equation [47–49]. However, the velocity v is here an imposed parameter, in contrast to a traveling wave velocity. Moreover, the nontrivial stationary solutions satisfying the boundary conditions Eqs. 2 are in a different parameter regime compared to Fisher waves (see Supplementary material Section S3). We quantify the spatial dependence of the concentration profiles through the difference between food concentration at the entrance and at the exit of the gut, normalized by the incoming food concentration F_{in}, namely $[F(0) - F(L)]/F_{\mathrm{in}}$. A heat map of this quantity is depicted in the $(v, D)-$parameter space in Figure 1C. We observe diverse levels of spatial dependence, ranging from strongly spatial profiles to quasi-flat ones, where the system is almost well-mixed and resembles a chemostat, or where bacteria are washed out by the flow [39, 40] (see Figure S1 for examples of concentration profiles across these regimes). There are two washout limits here. First, for large diffusion coefficients, if the flow timescale is smaller than the replication timescale, bacteria exit the system before reproducing. Second, for small diffusion coefficients, on the timescale of one replication, if the characteristic length of flow is larger than that of diffusion, bacteria are washed out (see Supplementary material Section S4 and Figure S5).

To compare our spatial system to a well-mixed one, we consider a chemostat [50] with the same total number of bacterial reproductions N_{R} per unit time as in the spatial system, which is

$$N_{\mathrm{R}} = S \int_0^L B(x)\rho(x)dx, \tag{4}$$

where S is the surface area of the gut section, while $\rho(x)$ is the reproduction rate of bacteria, which can be expressed using food concentration as in Eqs. 1:

$$\rho(x) = r\frac{F(x)}{k + F(x)}. \tag{5}$$

This reproduction rate strongly depends on the spatial coordinate in the spatial regime of the concentration profiles, see Figure 1B. Once the total number of reproductions is matched, it is possible to impose an additional matching condition, and we consider three possibilities for it in the Supplementary material (see Section S5). These matching conditions allow to set the parameters characterizing the chemostat matching the spatial

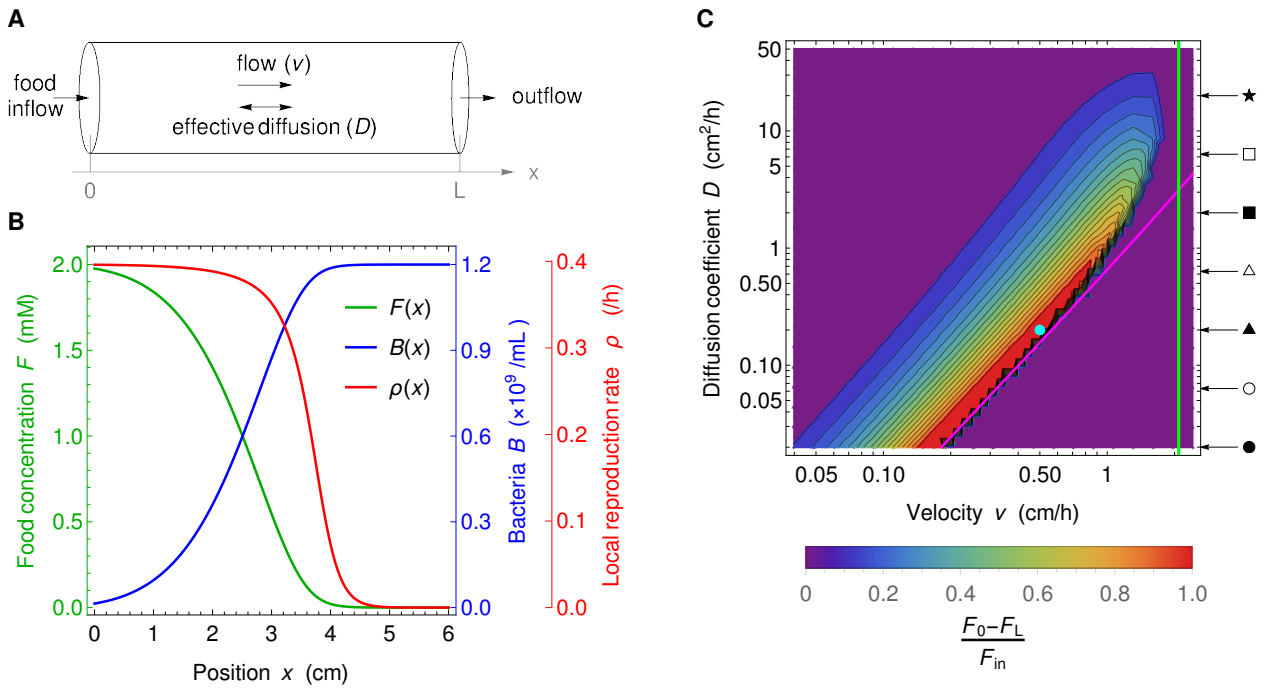

Figure 1: Model of the gut and associated spatial gradients. **A**: Schematic representation of the gut model investigated. We consider a cylinder with length L and neglect concentration variations in the radial direction, thus simplifying the system to one dimension along the x axis. Transport is modeled as flow with constant velocity v and effective diffusion coefficient D. At the upstream boundary $x = 0$ we consider constant food inflow, and no bacteria inflow, while at the downstream boundary $x = L$ we consider zero diffusive outflow. **B**: Concentration F of food, amount B of bacteria, and reproduction rate ρ of bacteria versus the coordinate x along the gut. Curves are numerical solutions of Eq. 1 for $D = 0.2\,\mathrm{cm}^2/\mathrm{h}$, $v = 0.5\,\mathrm{cm/h}$, $k = 0.1\,\mathrm{mM}$, $r = 0.42\,\mathrm{h}^{-1}$, $v\,F_{\mathrm{in}} = 1\,\mathrm{mM\,cm/h}$, $\alpha = 6.13 \times 10^8$ bacteria/mM and no mutant bacteria. The section area, S, is taken to be $1\,\mathrm{cm}^2$ in the entire paper, and the length L is 6 cm, as in the mini-gut of Ref. [39], in the main text, but is varied in the Supplementary material (see Section S10). The parameters are chosen such that they fall in range of parameters compared to the experiments in [39] and that the concentration profile is dependent of the spatial coordinate. The depicted concentrations represent the state of the system after numerically integrating partial differential Eqs. 1 for time t =500 h which is sufficient to reach the steady state. **C**: Heat map of the level of spatial dependence of the concentration profiles, quantified by $[F(0) - F(L)]/F_{\mathrm{in}}$, versus v and D. High values of $[F(0) - F(L)]/F_{\mathrm{in}}$ (red) mean strong gradients in the gut. Magenta and green lines represent washout limits, resp. $D = v^2(k/F_{\mathrm{in}} + 1)/(4r)$ and $v = rL/(k/F_{\mathrm{in}} + 1)$. Below the magenta line and on the right side of the green line, there are no bacteria in the gut at steady state, while in the purple region on the top left hand side, the system is well-mixed, leading to an almost uniform but non-zero concentration of bacteria in the gut. Parameter values (except v and D) are the same as in panel B. The values of v and D used in panel B are indicated by a circular cyan marker. Arrows and symbols on the right hand-side of the heat map indicate the diffusion coefficient values employed in Figure 4 with the same symbols.

system, namely its dilution rate, food inflow and volume. In all cases, we observe that matching chemostats feature extreme values for some of these parameters (see Figure S6), which arises from the very small outflow of food in the spatial system (see Supplementary material Section S5). These results emphasize that the large intestine is a highly efficient system for converting unabsorbed nutrients into bacteria.

Dynamics and fate of neutral mutants appearing in the gut

Let us now consider neutral mutants that spontaneously appear in the gut at steady state. Mutants may appear at any position along the gut, which can feature strong spatial heterogeneities (see Figure 1). How does the initial position of these mutants affect their dynamics and their steady-state concentration?

The initial local concentration of mutants is assumed to be much smaller than that of the wild type at the position x_{M} where the mutants appear (see Eqs. 3), as we aim to describe the fate of a single mutant or a few mutants, but in the framework of the continuous description of the gut. The early dynamics of mutant concentration is governed by the fluid dynamics in the gut. Indeed, the position x with the highest mutant concentration at a given time t initially follows the $x = x_{\mathrm{M}} + vt$ line, while the time t for which the mutant concentration is maximal for a given position x initially follows the $t = (x - x_{\mathrm{M}})^2/(2D)$ curves (see Supplementary material, Figure S7). This is consistent with the infinite space solution of the diffusion equation obtained from Eqs. 1 when ignoring reproduction. Hence, transport by convective and diffusive flow allows the early spread of the mutants in the gut. Afterwards, coupling with the reproduction term and the boundary conditions yields more complex dynamics.

Because neutral mutant concentration satisfies the same partial differential equation as wild-type bacteria concentration (see Eqs. 1), the steady-state concentrations of mutant and wild type bacteria satisfy $M(x)/B(x) = C$, where C depends on the initial conditions but not on x. In other words, the steady-state concentration profile of neutral mutants versus position x along the gut is the same as for wild-type bacteria, but with an overall rescaling. The magnitude of this rescaling (i.e. the value of C) depends on the initial mutant quantity, and on the position x_{M} where mutant bacteria appear. The latter dependence on x_{M} is strong in the regime where spatial dependence is strong in the mutant-free system (see Figure 1C), as shown in Figure 2A and Figures S8 and S9 in the Supplementary material. If the number of mutants that appear is held constant, then mutants make up a much larger steady-state fraction of bacteria if they appeared close to the entrance of the gut than if they appeared close to its exit, because they have more opportunity to spread and divide in the gut.

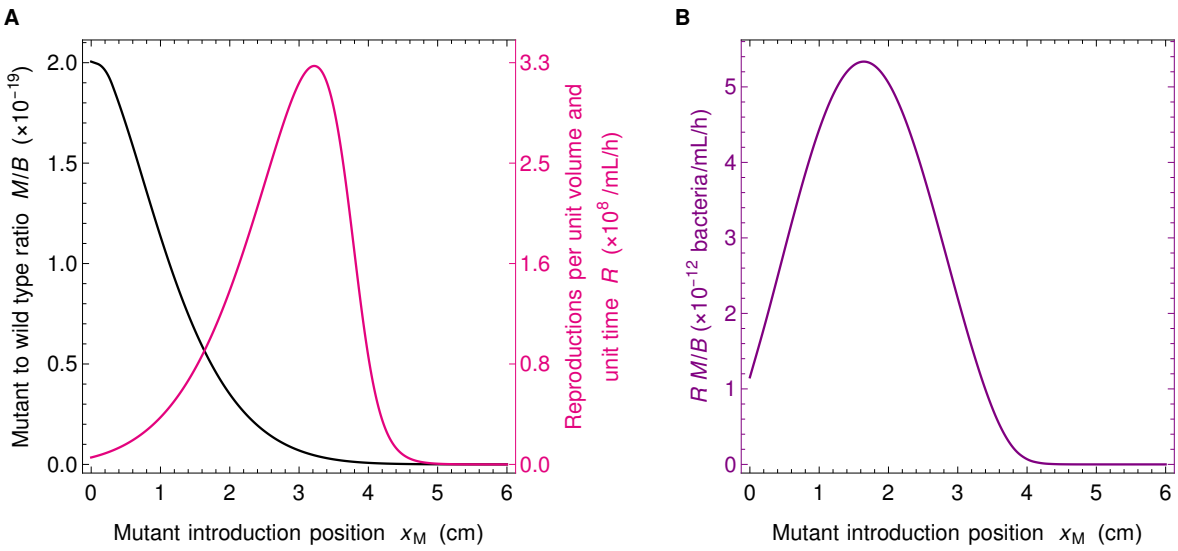

Figure 2: Fate of neutral mutants appearing at various locations in the gut. A: Steady-state ratio M/B of mutant to wild-type bacteria concentrations, and number of reproduction events R per unit volume and unit time versus position x_{M} of the mutant introduction. The ratio M/B yields the fixation probability of a mutant that appears at a given position x_{M} in the system. As mutants generally appear upon division, the appearance of new mutants is proportional to R, which thus also matters for the overall likelihood that a mutant appears and fixes. Mutants are introduced at numerical integration time $t = 500\,\mathrm{h}$, when steady state is reached. **B**: Product of the ratio M/B and the number R of reproductions per unit volume and unit time versus x_{M}. This quantity yields the fixation probability of a mutant that appears proportionally to reproduction rate. Parameter values are the same as in Figure 1B, and F and B are initially at steady state as in Figure 1B, while mutants are introduced locally by using the initial condition in Eqs. 3, with a total number $N_M = 3.33 \times 10^{-11}$ of mutants introduced in the system.

In our deterministic continuous description, bacterial species or strains coexist forever (except in the washout case where they are all wiped away), reflected by the fact that M/B is nonzero at steady state. However, the fate of individual mutants is in fact affected by demographic fluctuations known as genetic drift [42], so that in a finite system, mutants eventually either take over the population or disappear. Here, on a short time scale, mutants either reach deterministic steady state in coexistence with the wild-type, or they get extinct stochastically. If they reach steady state, then, on a longer time scale, proportional to population size [42],

one of the two types takes over. What is the probability that a mutant lineage that has reached steady state then fixes in the population? In a well-mixed system, the fixation probability of a neutral mutant is given by the ratio of the number of mutants to the total number of individuals [42]. In our gut model, the steady-state ratio M/B is independent of x in the deterministic limit (note that throughout we have $M \ll B$ so that here $M/(M+B) \approx M/B$, and we only discuss M/B). Moreover, in Eqs. 1, the only non-linearity in the evolution of B and M comes from the dependence of F on B and M. Here, since we introduce a very small amount of mutants, $M_0 \ll B$, when B is at stationary state, and since the overall bacterial population is very large, F remains almost constant through the evolution of M, which entails that the equations for B and M are then approximately linear. Because in the linear case, the equations on averages across replicates of a stochastic system coincide with those of the deterministic large-size limit [51], the fixation probability of neutral mutants in the stochastic case is given by the deterministic steady-state ratio $M/(M+B) \approx M/B$. In section S12 of the Supplementary material, we provide a validation of our deterministic analysis by stochastic simulations. Fig. S16 demonstrates the good agreement between the two descriptions regarding the fate of neutral mutants appearing at various locations in the gut. In particular, it confirms that the deterministic steady-state ratio M/B yields the mutant fixation probability, for each given mutant introduction position x_{M}. Given the dependence of the ratio M/B on the initial position x_{M} of the mutants (see Figure 2A and discussion above), mutants appearing close to the entrance of the gut are much more likely to fix than those appearing close to its exit, in the regime with strong spatial dependence (see Figure 1C).

Where in the gut do the mutants that fix originate? To address this question, we need to account for the apparition of mutants as well as for their fixation. Assume that mutations occur upon division, which is the case for replication errors. Then, mutants appear at a position x_{M} proportionally to the local number

$$R(x_{\mathrm{M}}) = B(x_{\mathrm{M}})\rho(x_{\mathrm{M}}) \tag{6}$$

of reproduction events per unit volume and unit time (where the reproduction rate ρ is given by Eq. 5). This number is small close to the exit of the gut because food is exhausted, but it is also small close to its entrance because bacteria are scarce, and it features a maximum at an intermediate location (see Figure 2A). What ultimately sets the location where mutants that fix tend to originate is the product of M/B and R, whose dependence on the mutant initial position x_{M} is depicted in Figure 2B. It features a strong spatial dependence, with a maximum at an intermediate position in the gut. In the Supplementary material section S8, we study R and $R\,M/B$ for various parameter values, and show that this maximum of $R\,M/B$ at an intermediate position in the gut is obtained robustly in the regime with strong spatial dependence (see Figure S10). Furthermore, Fig. S16 demonstrates the good agreement of the results obtained in our deterministic model with those from stochastic simulations.

Spatial structure in the gut increase the fixation probability of neutral mutants

What is the overall probability $\mathcal{F}$ that neutral mutants fix in the gut, averaged over their possible positions of origin? It can be expressed as the integral over all possible initial mutant locations x_{M} of the fixation probability given x_{M}, multiplied by the probability that the mutant originates at this location x_{M}:

$$\mathcal{F} = \frac{\int_0^L R(x_{\mathrm{M}}) \frac{M(x_{\mathrm{M}})}{B(x_{\mathrm{M}})} dx_{\mathrm{M}}}{\int_0^L R(x_{\mathrm{M}}) dx_{\mathrm{M}}}. \tag{7}$$

How is the overall fixation probability $\mathcal{F}$ of a neutral mutant affected by the spatial dependence of food and bacterial concentrations in the gut? To address this question, Figure 3 depicts $\mathcal{F}$ versus total population size $N_{\mathrm{T}} = S\int_0^L B(x)dx$ for different velocities v and diffusion coefficients D. In order to include concentration profiles with different degrees of spatial dependence, quantified by $[F(0) - F(L)]/F_{\mathrm{in}}$ (see Figure 1C), several values of D were chosen, and for each of them, a range of velocities v was chosen using Figure 1C so that it includes flat profiles for small velocities, spatial profiles for intermediate velocities, and again flat profiles close to the washout limit. Throughout, the food inflow $v\,F_{\mathrm{in}}$ at the entrance of the gut was held constant to allow comparison. In a well-mixed system, we would have $\mathcal{F} = N_{\mathrm{M}}/N_{\mathrm{T}}$, where N_{M} denotes the initial number of mutants in the system and N_{T} the total number of bacteria in the system [42]. We find an excellent agreement with this expectation in the case of flat concentration profiles. This is evident for small N_{T} values, which correspond to the largest velocities considered and thus to the washout limit, when the concentration profiles are the flattest (see Figure 3). Conversely, in the strongly spatial regime (red symbols in Figure 3), the fixation probability deviates from the well mixed system expectation, becoming substantially larger than it, and almost independent of the total population. For large N_{T}, which corresponds to small velocities, and hence flat profiles again, the fixation probability slowly converges back to the well-mixed system expectation (see Figure 3).

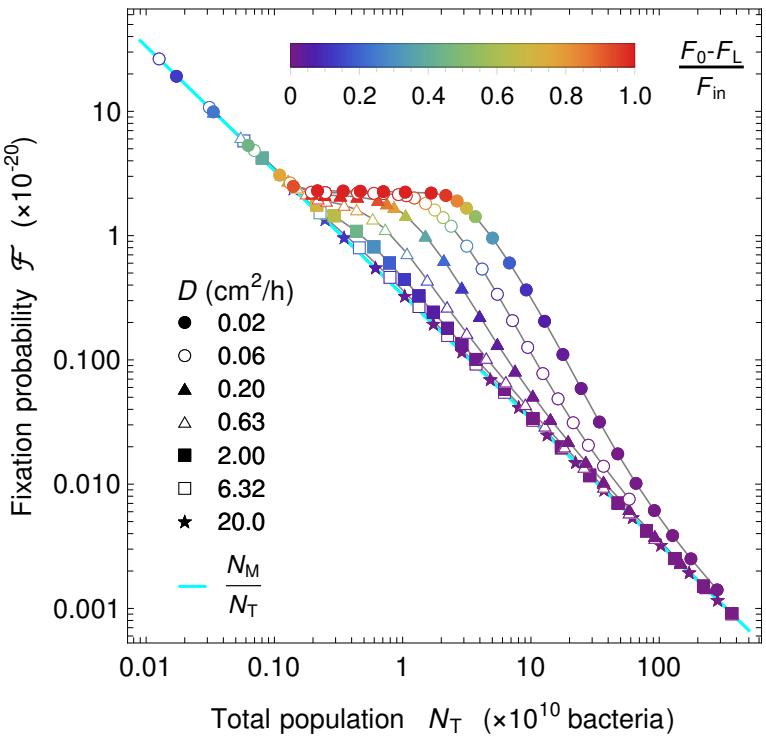

Figure 3: Fate of a neutral mutant versus population size in the gut. Fixation probability $\mathcal{F}$ of mutants appearing proportionally to reproduction rate is shown versus total population size N_T, for different diffusion coefficients D (see corresponding markers on the right hand-side of the heat map in Fig. 1C). Markers are colored by the level of spatial dependence of the concentration profiles, quantified by $[F(0) - F(L)]/F_{in}$ as in Fig. 1C. For strong spatial dependence (red), a plateau is observed, evidencing a large difference with the well-mixed expectation $\mathcal{F} = N_M/N_T$. For each value of D, v is varied while keeping $v\,F_{in}$ constant, and fixation probability is calculated from Eq. 7 and total population by integrating the sum of the mutant and the wild type in the total space (volume). Parameter values: $D \in [0.02, 20.0]$ cm^2/h, $v \in [0.001, 2.4]$ cm/h, $k = 0.1$ mM, $r = 0.42\,\mathrm{h}^{-1}$, $v\,F_{in} = 1$ mM cm/h, and $\alpha = 6.13 \times 10^8$ bacteria/mM and initial conditions as in Figure 2.

The fixation probability of neutral mutants results from an active population

Why is the fixation probability of neutral mutants larger in the gut in the presence of strong spatial dependence than in a well-mixed population with the same size? An important difference is that not all bacteria are actively reproducing in the gut, while they all have the same replication rate in a well-mixed population. More precisely, in the regime with strong spatial dependence, most replications occur in the region such that the local number of reproduction events $R(x)$ per unit volume and unit time (see Eqs. 6 and 5) is substantial, i.e. visually, under the local replication rate curve, which coincides with the zone where bacterial concentration increases (see Figure 4A). Quantitatively, we define the "active population", i.e. the region with active reproduction, by comparing the replication rate to its maximum possible value, see Figure 4A and the Supplementary material Section S9 for details.

Can the active population, smaller than the total population, and comprising the reproducing bacteria, quantitatively explain the fixation probability observed in the gut in the presence of strong spatial dependence? In order to assess this, we set out to significantly change active population size, and thus the total number of reproduction events, by varying the food inflow $v\,F_{in}$ at the entrance of the gut, while holding the diffusion coefficient constant at $D = 0.02$ cm/h. We took several velocity values, but only retained those such that concentration profiles were strongly spatially dependent. Figure 4B shows the fixation probability $\mathcal{F}$ versus the size N_A of the active population in this spatial regime. Our results agree very well with the relation

$$\mathcal{F} = \frac{N_M}{N_A}, \tag{8}$$

where N_M is the initial number of mutant bacteria. This corresponds to the well-mixed expectation for the fixation probability of N_M mutants in a population of N_A bacteria, which confirms that the active population is the one that matters for the process of mutant fixation. This explains why the fixation probability is higher in the spatial system than in the well-mixed one, as well as the shape of the curves in Fig. 3. Indeed, in each of these curves, as v is decreased at a given D, N_T increases. Fig. 1C shows that the system then goes from quasi washout where N_T is small and $N_A \approx N_T$ to the strongly spatial regime where $N_A \ll N_T$, and finally toward

less strongly spatial regimes where replication is slow in the whole system. When $N_A \approx N_T$, Eq. 8 reduces to the well-mixed expectation $\mathcal{F} = N_M/N_T$, while it strongly deviates from it when $N_A \ll N_T$. Moreover, in the strongly spatial regime, almost all food is consumed. As the food inflow $v\,F_{in}$ is constant in Fig. 3, the overall production rate of new bacteria, which is approximately rN_A as most reproductions occur in the active population, is then constant too, yielding a constant N_A. Accordingly, Fig. S1A shows that when v is decreased within the strongly spatial regime, the transition from low to high bacterial concentration gradually shifts upstream in the gut while retaining the same shape, and thus N_A remains constant while N_T increases. This explains the plateau observed in Fig. 3.

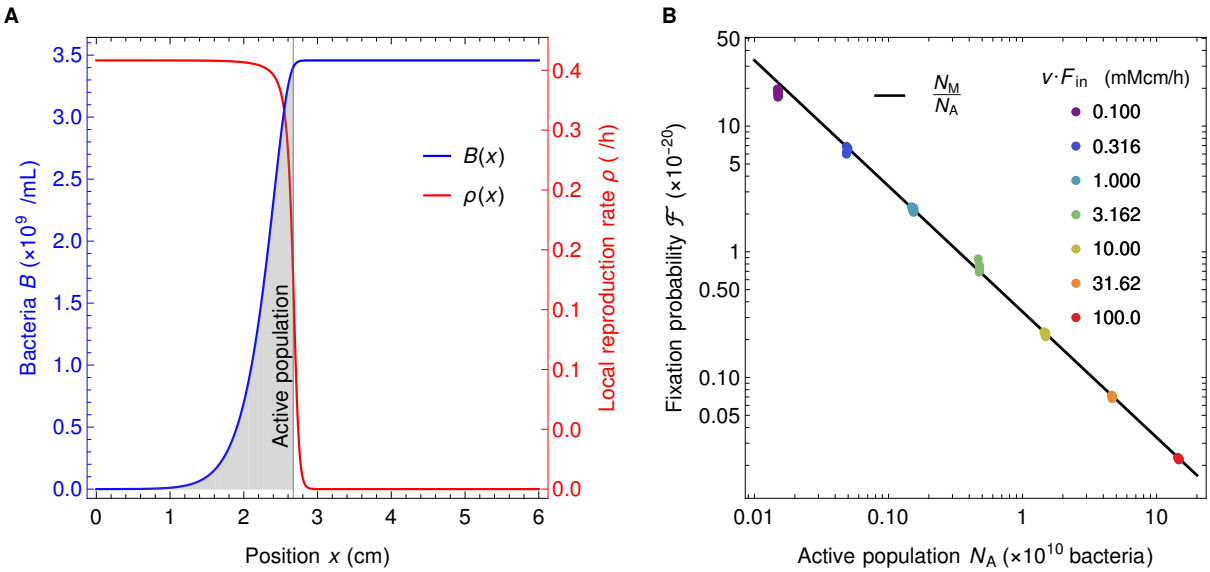

Figure 4: Active population explains the behavior of neutral mutant fixation probability in the gut. A: The active bacterial population (gray shaded area) is defined as the total number of bacteria between the points $x = 0$ and $x = x^*$, where x^* is defined as $F(x^*) = k$, so that $B(x^*) = \alpha F_{\rm in}(1 - k/F_{\rm in})$. This corresponds to the region where bacteria have significant reproduction rates. Parameters are $v\,F_{\rm in} = 1\,{\rm mM\,cm/h}$, $v = 0.181\,{\rm cm/h}$, $D = 0.02\,{\rm cm^2/h}$, $k = 0.1\,{\rm mM}$, $r = 0.42\,{\rm h^{-1}}$, and $\alpha = 0.613 \times 10^9\,{\rm bacteria/mM}$. **B**: Fixation probability $\mathcal{F}$ of neutral mutants in the gut in the regime with strong spatial dependence versus active population N_A, for different values of food inflow $v\,F_{\rm in}$ (different colors). Each set of markers with a given color contains between 6 and 11 different points (often overlapping). Diffusion coefficient is the same for all points, $D = 0.02\,{\rm cm^2/h}$, velocities are $v \in [0.135, 0.171]$ cm/h, $v \in [0.14, 0.18]$ cm/h, $v \in [0.15, 0.185]$ cm/h, $v \in [0.15, 0.181]$ cm/h, $v \in [0.15, 0.186]$ cm/h, $v \in [0.155, 0.182]$ cm/h, $v \in [0.165, 0.186]$ cm/h for $v\,F_{\rm in} = 0.1$ to 100.0 mM cm/h, respectively. Other parameters and initial conditions are as in Figure 3. Only the data points satisfying $[F(0) - F(L)]/F_{\rm in} > 0.9$ are retained, ensuring that we focus on the plateau of the fixation probability with respect to the total population (see Fig. 3). The black line corresponds to $\mathcal{F} = N_M/N_A$. Panel A corresponds to one of the green dots in panel B.

In the Supplementary material, section S10, we demonstrate the generality of the conclusions obtained here by systematically investigating the three dimensionless parameters that fully describe the stationary state of the system. Eq. 8 holds in all cases considered, provided that the food concentration profile is strongly spatial (see Figures S12, S13 and S14). Furthermore, we demonstrate that the range of parameters considered in the present study matches the realistic one in the human colon. Finally, Fig. S17 in section S12 of the Supplementary material shows that our prediction in Eq. 8 is validated by stochastic simulations.

Discussion

We addressed bacterial evolution in the gut within a minimal model that incorporates flow and gradients of food and bacterial concentrations along the gut. We focused on the colon lumen, where the vast majority of our microbiota is located, and we studied parameter ranges relevant for the human colon. We considered neutral mutants appearing in the gut. Estimates of bacterial population sizes in the human colon [1, 52] and of fitness effects [53] show that a substantial fraction of spontaneous mutations occurring in gut bacteria is expected to be effectively neutral (see Supplementary material, section S11). The dynamics of bacteria and food was described using a system of partial differential equations based on Refs. [39, 40]. In the long term, in a finite-size system, mutants either disappear or take over due to stochastic fluctuations, and the stationary proportion of mutants in our continuous and deterministic description gives their fixation probability. We demonstrated that, in the regime where the profiles of food and bacterial concentrations are strongly spatial, with abundant food and few bacteria upstream, and vice-versa downstream, the stationary concentration of mutants is higher

if they start upstream. However, for mutations occurring at replication, the small upstream concentration of bacteria means that few mutants appear there. Accordingly, we found that successful mutants are more likely to originate from an intermediate position along the gut. We studied the overall long term mutant proportion for neutral mutants appearing spontaneously upon division, which also gives their fixation probability. We found that in the almost well-mixed regime, it is given by the ratio of the initial number of mutants to the total bacterial population size, consistently with the well-mixed expectation. By contrast, when the profiles of food and bacterial concentrations are strongly spatial, which is the relevant regime in the gut [39, 40], this fixation probability becomes substantially larger than the well-mixed expectation. Thus, the spatial structure of the gut favors the spread of neutral mutants and the evolution of the population composition. Furthermore, we rationalized this increase of the fixation probability by demonstrating that it stems from the fact that only a subset of the bacterial population is actively replicating. This active population is located upstream, where there is enough food to allow substantial replication. It gives an effective population size [12, 42] for the fixation of neutral mutants in the complex structured population of the gut.

Studies addressing the impact of spatial population structure on evolution generally consider discrete patches of population with migrations between them, and the same environment in each of them [5–21]. While complex population structures with asymmetric migrations can impact the fixation probabilities of beneficial and deleterious mutants [13, 14, 21], that of neutral mutants appearing uniformly in the population, e.g. upon division, is unaffected [14, 21]. Similarly, chaotic hydrodynamic flow has been predicted to impact non-neutral mutant fixation probabilities, but not neutral ones [54]. In the gut, the flow can be viewed as yielding asymmetric migrations. Strikingly, we found that the fixation probability of neutral mutants could strongly differ from the well-mixed case. Aside the fact that the gut is a continuous system, a crucial difference with the above-cited models of population structure is that, due to directional hydrodynamic flow, the environment varies along the gut, in particular the food and bacterial concentrations, and thus the bacterial division rate. Environmental gradients can strongly impact evolution; for instance, gradients of antibiotics can increase the speed at which antibiotic resistance emerges [32–35]. The coupling of bacterial concentration gradients due to antibiotics with convective flow also has complex implications on evolution [55]. Hydrodynamic flow itself can strongly impact fixation probabilities and fixation times, as has been shown in the case of compressible flows relevant for large-scale turbulent systems such as bacterial populations living at the surface of oceans [29–31]. In these situations, flow reduces the effective population size for fixation probability, and microorganisms born near a flow source are more likely to fix than those born in a flow sink [31]. Albeit obtained in a different hydrodynamic regime, these results share similarities with ours, and together, they demonstrate that hydrodynamic flow, and in particular convective flow, can strongly impact evolution at various scales, from the gut to the ocean.

In addition to hydrodynamic flow and gradients, the gut comprises an upstream zone with few bacteria and rapid growth. This is reminiscent of expanding fronts in populations that invade a new environment [56, 57], which feature reduced competition and reduced effective population sizes, with important consequences on evolution [58–60]. In these cases, the dynamics is different depending on whether the traveling waves characterizing expansion are driven by the leading edge (pulled, as e.g. Fisher waves [47–49]) or by the bulk of the wave (pushed), yielding different wave velocities [57, 61]. Contrary to population expansion on solid substrates [56], the gut features directional hydrodynamic flow. The associated velocity v is imposed, as opposed to a traveling wave velocity. Besides, boundary conditions put us in a different parameter regime compared to Fisher waves. One may worry that our deterministic model may not be appropriate in the upstream region. However, this concern is alleviated by the directional flow, which transports bacteria downstream. Specifically, bacteria take at least 20 minutes to replicate (here we took a typical replication time of 100 minutes), and since they are transported with the flow, the lineage of an upstream bacteria will be broadly distributed, including where there are many bacteria, before being large enough to affect F sufficiently to modify the dynamical equations via $F/(k+F)$. Furthermore, the main findings from our deterministic model are validated by stochastic simulation results (Supplementary material section S12).

Extending our study from neutral mutants to beneficial and deleterious ones, and studying fixation times and the rate of evolution in the gut, would be interesting topics for future work. Note that given the very large numbers of bacteria at play, fixation is expected to be slow. However, even before fixation, our results show that the proportion of mutants is increased by the gut structure compared to a well-mixed system. Our stochastic simulation results in Fig. S18 confirm that the timescale for the increase of the average proportion of mutants is much faster than the one for mutant fixation. Furthermore, while the minimal model used here captures some key characteristics of the gut, with a net flow, an effective mixing that is on limited length scales, and a stable bacterial population, the reality of the gut is more complex. In particular, muscle contractions in peristalsis and segmentation [62, 63] mean that the radius of the gut is variable, and yield complex mixing dynamics. Besides, several food sources and several bacterial species are present, yielding complex ecological dynamics. Bacterial populations in the colon lumen can also interact with those in the mucus and in crypts. In addition, assuming a constant food inflow is a simplification, and in real life food inflow is variable, depending e.g. on

the timing of meals, thus adding time variability to the spatial gradients we considered here. Despite all these complications, our results, which can be interpreted simply through the active population, have the potential to be general, and can be tested in more detailed models.

Note. While this manuscript was in revision, an independent and complementary study [64] was released on BioRxiv.

Acknowledgments. This work was supported by an Emergence Grant from Sorbonne Université (to A.-F. B.). A.-F. B. acknowledges funding by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 851173).

References

[1] R. Sender, S. Fuchs, and R. Milo. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biology*, 14(8):e1002533, 2016.

[2] Peter J Turnbaugh, Ruth E Ley, Micah Hamady, Claire M Fraser-Liggett, Rob Knight, and Jeffrey I Gordon. The human microbiome project. *Nature*, 449(7164):804, 2007.

[3] Timothy F Landers, Bevin Cohen, Thomas E Wittum, and Elaine L Larson. A review of antibiotic use in food animals: perspective, policy, and potential. *Public health reports*, 127(1):4–22, 2012.

[4] C Rose, Alison Parker, Bruce Jefferson, and Elise Cartmell. The characterization of feces and urine: a review of the literature to inform advanced treatment technology. *Critical reviews in environmental science and technology*, 45(17):1827–1879, 2015.

[5] S. Wright. Evolution in Mendelian populations. *Genetics*, 16(2):97–159, 1931.

[6] M. Kimura and G. H. Weiss. The Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation with Distance. *Genetics*, 49(4):561–576, Apr 1964.

[7] T. Maruyama. On the fixation probability of mutant genes in a subdivided population. *Genet. Res.*, 15:221–225, 1970.

[8] T. Maruyama. A simple proof that certain quantities are independent of the geographical structure of population. *Theor. Popul. Biol.*, 5(2):148–154, Apr 1974.

[9] T. Nagylaki. The strong-migration limit in geographically structured populations. *J Math Biol*, 9(2):101–114, Apr 1980.

[10] M. Slatkin. Fixation probabilities and fixation times in a subdivided population. *Evolution*, 35(3):477–488, 1981.

[11] N. Barton. The probability of fixation of a favoured allele in a subdivided population. *Genet. Res.*, 62:149–157, 1993.

[12] M. C. Whitlock and N. H. Barton. The effective size of a subdivided population. *Genetics*, 146(1):427–441, May 1997.

[13] M. C. Whitlock. Fixation probability and time in subdivided populations. *Genetics*, 164(2):767–779, Jun 2003.

[14] E. Lieberman, C. Hauert, and M. A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, Jan 2005.

[15] Kamran Kaveh, Natalia L. Komarova, and Mohammad Kohandel. The duality of spatial death-birth and birth-death processes and limitations of the isothermal theorem. *Royal Society Open Science*, 2(4):140465, 2015.

[16] Laura Hindersin and Arne Traulsen. Most undirected random graphs are amplifiers of selection for birth-death dynamics, but suppressors of selection for death-birth dynamics. *PLOS Computational Biology*, 11(11):1–14, 11 2015.

[17] Karan Pattni, Mark Broom, Jan Rychtář, and Lara J. Silvers. Evolutionary graph theory revisited: when is an evolutionary process equivalent to the moran process? *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2182):20150334, 2015.

[18] Bahram Houchmandzadeh and Marcel Vallade. The fixation probability of a beneficial mutation in a geographically structured population. *New Journal of Physics*, 13(7):073020, Jul 2011.

[19] B. Houchmandzadeh and M. Vallade. Exact results for fixation probability of bithermal evolutionary graphs. *Biosystems*, 112(1):49–54, Apr 2013.

[20] G. W. Constable and A. J. McKane. Population genetics on islands connected by an arbitrary network: an analytic approach. *J Theor Biol*, 358:149–165, Oct 2014.

[21] L. Marrec, I. Lamberti, and A.-F. Bitbol. Toward a universal model for spatially structured populations. *Phys Rev Lett* 127(21):218102, Nov 2021.

[22] I. Gordo and P. R. A. Campos. Adaptive evolution in a spatially structured asexual population. *Genetica*, 127:217, 05 2006.

[23] L. Perfeito, M. I. Pereira, P. R. Campos, and I. Gordo. The effect of spatial structure on adaptation in Escherichia coli. *Biol. Lett.*, 4(1):57–59, Feb 2008.

[24] S. Kryazhimskiy, D. P. Rice, and M. M. Desai. Population subdivision and adaptation in asexual populations of Saccharomyces cerevisiae. *Evolution*, 66(6):1931–1941, Jun 2012.

[25] R. Korona, C. H. Nakatsu, L. J. Forney, and R. E. Lenski. Evidence for multiple adaptive peaks from populations of bacteria evolving in a structured habitat. *Proc. Natl. Acad. Sci. USA*, 91:9037–9041, 1994.

[26] A. F. Bitbol and D. J. Schwab. Quantifying the role of population subdivision in evolution on rugged fitness landscapes. *PLoS Comput. Biol.*, 10(8):e1003778, Aug 2014.

[27] J. R. Nahum, P. Godfrey-Smith, B. N. Harding, J. H. Marcus, J. Carlson-Stevermer, and B. Kerr. A tortoise-hare pattern seen in adapting structured and unstructured populations suggests a rugged fitness landscape in bacteria. *Proc Natl Acad Sci U S A*, 112(24):7530–7535, Jun 2015.

[28] M. T. France and L. J. Forney. The Relationship between Spatial Structure and the Maintenance of Diversity in Microbial Populations. *Am Nat*, 193(4):503–513, 04 2019.

[29] P. Perlekar, R. Benzi, D. R. Nelson, and F. Toschi. Population dynamics at high Reynolds number. *Phys Rev Lett*, 105(14):144501, Oct 2010.

[30] S. Pigolotti, R. Benzi, M. H. Jensen, and D. R. Nelson. Population genetics in compressible flows. *Phys Rev Lett*, 108(12):128102, Mar 2012.

[31] A. Plummer, R. Benzi, D. R. Nelson, and F. Toschi. Fixation probabilities in weakly compressible fluid flows. *Proc Natl Acad Sci U S A*, 116(2):373–378, 01 2019.

[32] Q. Zhang, G. Lambert, D. Liao, H. Kim, K. Robin, C Tung, N. Pourmand, and R. H. Austin. Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments. *Science*, 333(6050):1764–1767, 2011.

[33] Philip Greulich, Bartłomiej Waclaw, and Rosalind J Allen. Mutational pathway determines whether drug gradients accelerate evolution of drug-resistant cells. *Physical Review Letters*, 109(8):088101, 2012.

[34] R. Hermsen, J. B. Deris, and T. Hwa. On the rapidity of antibiotic resistance evolution facilitated by a concentration gradient. *Proc. Natl. Acad. Sci. USA*, 109:10775–10780, 2012.

[35] Michael Baym, Tami D Lieberman, Eric D Kelsic, Remy Chait, Rotem Gross, Idan Yelin, and Roy Kishony. Spatiotemporal microbial evolution on antibiotic landscapes. *Science*, 353(6304):1147–1151, 2016.

[36] Gregory P Donaldson, S Melanie Lee, and Sarkis K Mazmanian. Gut biogeography of the bacterial microbiota. *Nature Reviews Microbiology*, 14(1):20, 2016.

[37] Ken Kurokawa, Takehiko Itoh, Tomomi Kuwahara, Kenshiro Oshima, Hidehiro Toh, Atsushi Toyoda, Hideto Takami, Hidetoshi Morita, Vineet K Sharma, Tulika P Srivastava, et al. Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Research*, 14(4):169–181, 2007.

[38] Peter J Turnbaugh, Micah Hamady, Tanya Yatsunenko, Brandi L Cantarel, Alexis Duncan, Ruth E Ley, Mitchell L Sogin, William J Jones, Bruce A Roe, Jason P Affourtit, et al. A core gut microbiome in obese and lean twins. *Nature*, 457(7228):480–484, 2009.

[39] Jonas Cremer, Igor Segota, Chih Yu Yang, Markus Arnoldini, John T. Sauls, Zhongge Zhang, Edgar Gutierrez, Alex Groisman, and Terence Hwa. Effect of flow and peristaltic mixing on bacterial growth in a gut-like channel. *Proc. Natl. Acad. Sci. USA*, 113(41):11414–11419, 2016.

[40] Jonas Cremer, Markus Arnoldini, and Terence Hwa. Effect of water flow and chemical environment on microbiota growth and composition in the human colon. *Proc. Natl. Acad. Sci. USA*, 114(25):6438–6443, 2017.

[41] P. A. P. Moran. Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(1):60–71, 1958.

[42] W. J. Ewens. *Mathematical Population Genetics*. Springer-Verlag, 1979.

[43] Herbert F Helander and Lars Fändriks. Surface area of the digestive tract–revisited. *Scandinavian journal of gastroenterology*, 49(6):681–689, 2014.

[44] Richard A Cone. Barrier properties of mucus. *Advanced drug delivery reviews*, 61(2):75–85, 2009.

[45] Alexander Swidsinski, Jutta Weber, Vera Loening-Baucke, Laura P Hale, and Herbert Lochs. Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *Journal of clinical microbiology*, 43(7):3380–3389, 2005.

[46] Darka Labavić, Claude Loverdo, and Anne-Florence Bitbol. Hydrodynamic flow and concentration gradients in the gut enhance neutral bacterial diversity, (Version v1.0.0), http://doi.org/10.5281/zenodo.4704653. *Zenodo*, 2021.

[47] R. A. Fisher. The wave of advance of advantageous genes. *Ann. Eugenics*, 7:353–369, 1937.

[48] A. Kolmogorov, I. Petrovsky, and N. Piscounoff. Etude de l'équation de la diffusion avec croissance de la quantite de matiere et son application à un probleme biologique. *Moscow University, Bull. Math.*, 1:1–25, 1937.

[49] J. D. Murray. *Mathematical Biology I – An Introduction*. Springer-Verlag, 1989.

[50] D. Herbert, R. Elsworth, and R. C. Telling. The continuous culture of bacteria; a theoretical and experimental study. *Microbiology*, 14(3):601–622, 1956.

[51] Leif Gustafsson and Mikael Sternad. When can a deterministic model of a population system reveal what will happen on average? *Mathematical biosciences*, 243(1):28–45, 2013.

[52] Marcus J Claesson, Orla O'Sullivan, Qiong Wang, Janne Nikkilä, Julian R Marchesi, Hauke Smidt, Willem M De Vos, R Paul Ross, and Paul W O'Toole. Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PloS one*, 4(8):e6669, 2009.

[53] Lydia Robert, Jean Ollion, Jérôme Robert, Xiaohu Song, Ivan Matic, and Marina Elez. Mutation dynamics and fitness effects followed in single cells. *Science*, 359(6381):1283–1286, 2018.

[54] F. Herrerías-Azcué, V. Pérez-Muñuzuri, and T. Galla. Stirring does not make populations well mixed. *Sci Rep*, 8(1):4068, 03 2018.

[55] M. Gralka, D. Fusco, S. Martis, and O. Hallatschek. Convection shapes the trade-off between antibiotic efficacy and the selection for resistance in spatial gradients. *Phys Biol*, 14(4):045011, 07 2017.

[56] O. Hallatschek, P. Hersen, S. Ramanathan, and D. R. Nelson. Genetic drift at expanding frontiers promotes gene segregation. *Proc. Natl. Acad. Sci. USA*, 104:19926–19930, 2007.

[57] O. Hallatschek and D. R. Nelson. Gene surfing in expanding populations. *Theor Popul Biol*, 73(1):158–170, Feb 2008.

[58] J. M. Travis, T. Munkemuller, O. J. Burton, A. Best, C. Dytham, and K. Johst. Deleterious mutations can surf to high densities on the wave front of an expanding population. *Mol. Biol. Evol.*, 24(10):2334–2343, Oct 2007.

[59] L. Bosshard, I. Dupanloup, O. Tenaillon, R. Bruggmann, M. Ackermann, S. Peischl, and L. Excoffier. Accumulation of Deleterious Mutations During Bacterial Range Expansions. *Genetics*, 207(2):669–684, 10 2017.

[60] M. Gralka, F. Stiewe, F. Farrell, W. Mobius, B. Waclaw, and O. Hallatschek. Allele surfing promotes microbial adaptation from standing variation. *Ecol. Lett.*, 19(8):889–898, 08 2016.

[61] G. Birzu, O. Hallatschek, and K. S. Korolev. Fluctuations uncover a distinct class of traveling waves. *Proc Natl Acad Sci U S A*, 115(16):E3645–E3654, 04 2018.

[62] A. C. Ailiani, T. Neuberger, J. G. Brasseur, G. Banco, Y. Wang, N. B. Smith, and A. G. Webb. Quantitative analysis of peristaltic and segmental motion in vivo in the rat small intestine using dynamic MRI. *Magn Reson Med*, 62(1):116–126, Jul 2009.

[63] J. D. Huizinga and W. J. Lammers. Gut peristalsis is governed by a multitude of cooperating mechanisms. *Am J Physiol Gastrointest Liver Physiol*, 296(1):1–8, Jan 2009.

[64] Olivia M Ghosh and Benjamin H Good. Emergent evolutionary forces in spatial models of luminal growth in the human gut microbiota. *bioRxiv*, `https://doi.org/10.1101/2021.07.15.452569`, 2021.

[65] J. Crank. *The Mathematics of Diffusion*. Oxford science publications. Clarendon Press, 1979.

[66] P.C. Fife. *Mathematical aspects of reacting and diffusing systems*. Springer-Verlag, 1979.

[67] K. S. Korolev, M. Avlund, O. Hallatschek, and D. R. Nelson. Genetic demixing and evolution in linear stepping stone models. *Rev Mod Phys*, 82(2):1691–1718, Jun 2010.

[68] B. Perthame. *Parabolic Equations in Biology – Growth, Reaction, Movement and Diffusion*. Springer, 2015.

[69] M. D. Bramson. Convergence of solutions of the Kolmogorov equation to travelling waves . *Mem. Am. Math. Soc.*, 44(285):1–190, 1983.

[70] E. Brunet. Some aspects of the Fisher-KPP equation and the branching Brownian motion , 2016.

[71] NI McNeil. The contribution of the large intestine to energy supplies in man. *The American journal of clinical nutrition*, 39(2):338–342, 1984.

[72] JC Debongnie and SF Phillips. Capacity of the human colon to absorb fluid. *Gastroenterology*, 74(4):698–703, 1978.

[73] Karin Kovárová-Kovar and Thomas Egli. Growth kinetics of suspended microbial cells: from single-substrate-controlled growth to mixed-substrate kinetics. *Microbiology and molecular biology reviews*, 62(3):646–666, 1998.

[74] Ivan Ramirez, Eveline IP Volcke, Rajagopal Rajinikanth, and Jean-Philippe Steyer. Modeling microbial diversity in anaerobic digestion through an extended adm1 model. *Water research*, 43(11):2787–2800, 2009.

[75] Alonzo W Lawrence and Perry L McCarty. Kinetics of methane fermentation in anaerobic treatment. *Journal (Water Pollution Control Federation)*, pages R1–R17, 1969.

[76] Sherwood L. Gorbach, Andrew G. Plaut, Laila Nahas, Louis Weinstein, Gunter Spanknebel, and Ruven Levitan. Studies of intestinal microflora: II. microorganisms of the small intestine and their relations to oral and fecal flora. *Gastroenterology*, 53(6):856–867, 1967.

[77] Biomechanics of the digestive system. In Takami Yamaguchi, Takuji Ishikawa, and Yohsuke Imai, editors, *Integrated Nano-Biomechanics*, Micro and Nano Technologies, pages 71–99. Elsevier, Boston, 2018.

[78] JH Cummings, DJ Jenkins, and HS Wiggins. Measurement of the mean transit time of dietary residue through the human gut. *Gut*, 17(3):210–218, 1976.

[79] M. Proano, M. Camilleri, S. F. Phillips, G. M. Thomforde, M. L. Brown, and R. L. Tucker. Unprepared human colon does not discriminate between solids and liquids. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 260(1):G13–G16, 1991.

[80] Beth Gibson, Daniel J Wilson, Edward Feil, and Adam Eyre-Walker. The distribution of bacterial doubling times in the wild. *Proceedings of the Royal Society B*, 285(1880):20180789, 2018.

[81] T. Bataillon and S. F. Bailey. Effects of new mutations on fitness: insights from models and data. *Ann N Y Acad Sci*, 1320:76–92, Jul 2014.

[82] S. F. Levy, J. R. Blundell, S. Venkataram, D. A. Petrov, D. S. Fisher, and G. Sherlock. Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature*, 519(7542):181–186, Mar 2015.

Supplementary material

Contents

S1 Numerical methods **14**
S1.1 Discretization of the partial differential equation system . . . 14
S1.2 Obtaining stationary profiles without and with mutants . . . 15
S1.3 Conversion between food and bacteria concentrations . . . 16

S2 Stationary profiles without mutants **16**
S2.1 Ordinary differential equation description . . . 16
S2.2 Dimensionless form . . . 17
S2.3 Some stationary profiles . . . 17

S3 Comparison to Fisher waves **19**

S4 Washout limits **21**

S5 Correspondence between the spatial system and the chemostat **24**
S5.1 Main matching condition . . . 24
S5.2 Constraints from each separate system . . . 24
S5.3 Additional matching conditions . . . 25
S5.4 Properties of the matching chemostats . . . 25

S6 Early dynamics of mutant bacteria concentration **28**

S7 Stationary state of mutant bacteria concentration versus the initial position x_{M} of mutants **29**

S8 Location where most mutants that fix originate **30**

S9 Calculation of the active population size **30**

S10 Impact of the dimensionless parameters **31**
S10.1 Holding the dimensionless parameters fixed . . . 31
S10.2 Varying each dimensionless parameter . . . 32
S10.3 Range of the dimensionless parameters and relevance for the human colon . . . 33

S11 Relevance of neutral mutations **34**

S12 Validation by stochastic simulations **35**
S12.1 Stochastic simulation methods . . . 35
S12.2 Stochastic simulation results . . . 37

S1 Numerical methods

S1.1 Discretization of the partial differential equation system

In order to solve the system 1 numerically, we employ an explicit method using a forward difference for the time derivative at time t and a central difference for the space derivative at position x. Explicitly, the relevant differential operators are replaced by the following expressions:

$$\frac{\partial Y}{\partial t} \to \frac{Y(t,x) - Y(t-\Delta t, x)}{\Delta t}, \tag{9a}$$

$$\frac{\partial Y}{\partial x} \to \frac{Y(t,x+\Delta x) - Y(t,x-\Delta x)}{2\Delta x}, \tag{9b}$$

$$\frac{\partial^2 Y}{\partial x^2} \to \frac{Y(t,x+\Delta x) - 2Y(t,x) + Y(t,x-\Delta x)}{(\Delta x)^2}, \tag{9c}$$

where Δt and Δx represent the discrete steps in time and space, respectively. Here $Y(t,x)$ can represent the concentration of food F, or wild type bacteria B, or mutant bacteria M, at time t and coordinate x.

Substituting the differential operators in Eqs. 1 using Eqs. 9 yields

$$F(t+\Delta t,x) = F(t,x) + D\frac{F(t,x+\Delta x) - 2F(t,x) + F(t,x-\Delta x)}{(\Delta x)^2}\Delta t - \\ - v\frac{F(t,x+\Delta x) - F(t,x-\Delta x)}{2\Delta x}\Delta t - \frac{r}{\alpha}(B(t,x)+M(t,x))\frac{F(t,x)}{k+F(t,x)}\Delta t, \tag{10a}$$

$$B(t+\Delta t,x) = B(t,x) + D\frac{B(t,x+\Delta x) - 2B(t,x) + B(t,x-\Delta x)}{(\Delta x)^2}\Delta t - \\ - v\frac{B(t,x+\Delta x) - B(t,x-\Delta x)}{2\Delta x}\Delta t + rB(t,x)\frac{F(t,x)}{k+F(t,x)}\Delta t, \tag{10b}$$

$$M(t+\Delta t,x) = M(t,x) + D\frac{M(t,x+\Delta x) - 2M(t,x) + M(t,x-\Delta x)}{(\Delta x)^2}\Delta t - \\ v\frac{M(t,x+\Delta x) - M(t,x-\Delta x)}{2\Delta x}\Delta t + rM(t,x)\frac{F(t,x)}{k+F(t,x)}\Delta t. \tag{10c}$$

The boundary conditions in $x = 0$ from Eqs. 2 become

$$F(t,0-\Delta x) = F(t,0+\Delta x) + \frac{2\Delta x}{D}v\left[F_{\text{in}} - F(t,0)\right], \tag{11a}$$

$$B(t,0-\Delta x) = B(t,0+\Delta x) - \frac{2\Delta x}{D}vB(t,0), \tag{11b}$$

$$M(t,0-\Delta x) = M(t,0+\Delta x) - \frac{2\Delta x}{D}vM(t,0), \tag{11c}$$

while the boundary conditions in $x = L$ from Eqs. 2 become

$$F(t,L+\Delta x) = F(t,L-\Delta x), \tag{12a}$$

$$B(t,L+\Delta x) = B(t,L-\Delta x), \tag{12b}$$

$$M(t,L+\Delta x) = M(t,L-\Delta x). \tag{12c}$$

The spatial discrete step Δx is in general chosen to be 0.01 cm. This value is small enough to ensure convergence for most model parameters, and to have a good spatial resolution for analysis. Note however that for some parameters, the term $[F_{\text{in}} - F(t,0)]$ in Eq. 11a can be very large which can lead to numerical instability. To compensate, the spatial step, Δx, needs to be reduced. Once Δx is chosen, the value of Δt should satisfy the stability condition $\Delta t \leq (\Delta x)^2/(2D)$ [65]. Specifically, we take $\Delta t = 0.8(\Delta x)^2/(2D)$.

S1.2 Obtaining stationary profiles without and with mutants

Numerically, we determine the unique solution for food concentration that satisfies $F(x) < F_{\text{in}}$ for all x. Thus, the steady-state profile of food concentration is independent of the initial conditions as long as they are all positive. In general, for our numerical integration, we choose initial conditions not too far to the steady state, namely $F(0,x) = F_0 \in (0, F_{\text{in}})$, $B(0,x) = \alpha[F_{\text{in}} - F(0,x)]$, and $M(0,x) \ll B(0,x)$, in order to obtain faster convergence. If there is no mutant bacteria, the steady state of the wild type bacteria concentration is also uniquely defined through the relation $B(x) = \alpha[F_{\text{in}} - F(x)]$. However, if there are both wild type and mutant bacteria in the system, then the steady state solution is uniquely defined only for the total bacterial concentration, while individual concentrations depend on the initial conditions for wild type and mutant bacteria.

In practice, the stationary state in the absence of mutant bacteria is found using Eqs. 10 coded in the Fortran90 programming language (code available at `https://doi.org/10.5281/zenodo.4704653` [46]) with homogeneous initial conditions

$$F(0,x) = 0.9F_{\text{in}}, \tag{13a}$$

$$B(0,x) = \alpha[F_{\text{in}} - F(0,x)] = 0.1\alpha F_{\text{in}}. \tag{13b}$$

The time used in all numerical integrations is $t = 500$ h, which is long enough for the system to reach the steady state for all choices of parameters considered in this paper.

To find the mutant concentration profile that is crucial to our calculation of the fixation probability, we consider the system without mutants at steady state, and we assume that mutants appear at one local position, x_{M}, in

the gut segment. We again solve Eqs. 10 for time $t = 500\,\mathrm{h}$, but now with the initial conditions

$$F(0,x) = F^*(0,x), \tag{14a}$$

$$B(0,x) = B^*(0,x), \tag{14b}$$

$$M(0,x) = \begin{cases} M_0, & |x - x_\mathrm{M}| \leq \Delta x/2, \\ 0, & |x - x_\mathrm{M}| > \Delta x/2, \end{cases} \tag{14c}$$

where F^* and B^* represent steady state concentrations without mutant bacteria, while $M_0 \ll B(x_M)$ is the initial local mutant concentration. More precisely, denoting by N_M the total number of mutant bacteria introduced in the system, the quantity M_0 is the initial concentration in the segment $[x_\mathrm{M} - \Delta x/2, x_\mathrm{M} + \Delta x/2]$, where Δx is the spatial discrete step of our numerical resolutions. Hence, M_0 satisfies $N_\mathrm{M} M_0 S \Delta x$, where N_M is the total number of mutants introduced in the system. Note that since we are using central difference discretization, we need to double the M_0 value on boundaries of the segment where the mutants are introduced, in order to have the same N_M there as in the rest of the segment. In practice, we choose the value $N_\mathrm{M} = 3.33 \times 10^{-11}$ bacteria, so that for any choice of parameters used in this paper and for any initial mutant position x_M, the relation $M_0 \ll B(x_\mathrm{M})$ is satisfied. Importantly, since the stationary concentration of mutant bacteria is proportional to M_0, all results scale with it, and we are not losing generality by fixing the value of M_0.

S1.3 Conversion between food and bacteria concentrations

The initial unit of food concentration is moles per liter, and that of bacterial concentration is optical density, OD [39], which can be converted to numbers of bacteria per volume by using the calibration curve in [39]. Specifically, the conversion factor we take is $1\,\mathrm{OD}{=}3.33 \times 10^9$ bacteria/mL. Then the parameter α allows to convert between food and bacteria concentrations.

Importantly, because α is just a scaling factor, a change in this value will modify the bacterial concentration quantitatively, but the spatial profile and all the other conclusions will remain identical.

S2 Stationary profiles without mutants

S2.1 Ordinary differential equation description

Without mutants, at stationary state, Eq. 1 yields:

$$0 = D\frac{\partial^2 F}{\partial x^2} - v\frac{\partial F}{\partial x} - \frac{r}{\alpha}\frac{FB}{k+F}\,, \tag{15}$$

$$0 = D\frac{\partial^2 B}{\partial x^2} - v\frac{\partial B}{\partial x} + r\frac{FB}{k+F}\,, \tag{16}$$

with boundary conditions in Eq. 2.

Introducing $f = \alpha F + B$, we have at stationary state

$$C = D\frac{\partial f}{\partial x} - vf\,, \tag{17}$$

where C is a constant. The solution reads

$$f(x) = -\frac{C}{v} + C' e^{vx/D}\,, \tag{18}$$

where C' is a constant. Applying the boundary conditions yields $C' = 0$ and $C = -v\alpha F_\mathrm{in}$. Thus,

$$\alpha F(x) + B(x) = f(x) = \alpha F_\mathrm{in}\,, \tag{19}$$

and this specific linear combination of B and F is independent from x: food effectively gets converted into bacteria.

Now we can inject this into the equation on F to decouple it from B, yielding:

$$0 = D\frac{\partial^2 F}{\partial x^2} - v\frac{\partial F}{\partial x} - r\frac{F(F_\mathrm{in} - F)}{k+F}\,, \tag{20}$$

which is a second order nonlinear ordinary differential equation.

S2.2 Dimensionless form

Let us make the variable change $s = xv/D$, and let us introduce the function ϕ satisfying $\phi(s = xv/D) = F(x)/F_{\text{in}}$ for all x. We obtain

$$0 = \frac{\partial^2 \phi}{\partial s^2} - \frac{\partial \phi}{\partial s} - \lambda \frac{\phi(1-\phi)}{\kappa + \phi}\,, \tag{21}$$

which involves the dimensionless numbers

$$\kappa = \frac{k}{F_{\text{in}}}\,, \tag{22}$$

and

$$\lambda = \frac{rD}{v^2}\,. \tag{23}$$

The associated boundary conditions are:

$$\phi(s=0) - \frac{\partial \phi}{\partial s}(s=0) = 1\,, \tag{24}$$

$$\frac{\partial \phi}{\partial s}(s=\sigma) = 0 \tag{25}$$

where

$$\sigma = \frac{Lv}{D} \tag{26}$$

is the third dimensionless number describing the system [39], and compares transport by convection to transport by diffusion. It has the same form as the Péclet number for mass transfer, but recall that D is an effective diffusion coefficient modeling gut contractions.

S2.3 Some stationary profiles

In practice, the partial differential equations in Eqs. 1 with boundary conditions in Eqs. 2 and initial conditions in Eqs. 3 were solved numerically as explained above. Examples of profiles obtained are given in Figure S1. Note that we checked that the long-term results from their direct resolution was consistent to those obtained by numerically solving the ordinary differential equation 20 giving the stationary state of the system.

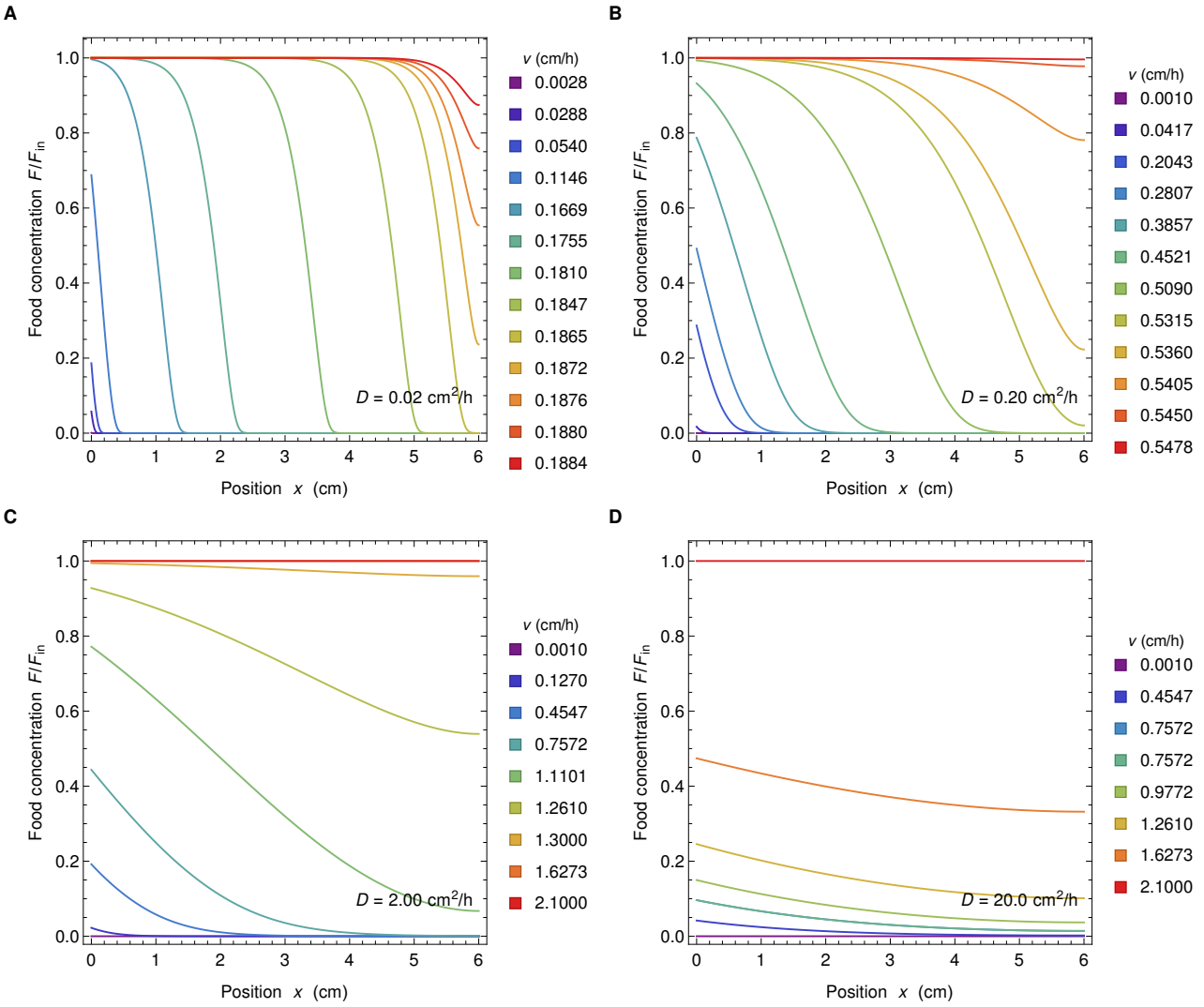

Figure S1: Food concentration profiles. Food concentration F normalized by the food concentration inflow F_{in} versus position x along the gut for four different values of the diffusion coefficient (panels A to D) and several values of velocity (different colors in each panel). Values of diffusion coefficients and velocity are indicated in each panel. Other parameter values are $k = 0.1\,\text{mM}$, $r = 0.42\,\text{h}^{-1}$, $v\,F_{\text{in}} = 1\,\text{mM}\,\text{cm/h}$, $\alpha = 6.13 \times 10^8$ bacteria/mM.

S3 Comparison to Fisher waves

In the regime with strong spatial dependence, the present model yields at steady state an upstream zone with few bacteria, high food concentration and rapid growth, and a downstream zone with many bacteria, little food and little growth. Specifically, our numerical resolutions show that $\phi = F/F_{\text{in}}$ then monotonically decays from a value close to one to a value close to zero, with a sigmoid-like shape (see Fig. 1B, as well as Fig. S1A-B for intermediate velocities). This is reminiscent of Fisher waves [47], which are (among several applications) an important model for the spread of mutant genes in populations [47], as well as for the expansion of populations that invade a new environment [49, 66], where the dynamics of mutants has been extensively discussed [56–61, 67]. Let us compare these two models.

The Fisher-Kolmogorov-Petrovsky-Piskunov (Fisher-KPP) equation [47, 48] is a reaction-diffusion equation of the form [66, 68]

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u)\,, \tag{27}$$

where f is a function satisfying $f(0) = f(1) = 0$. For $f(u) = u(1-u)$, it is known as the Fisher equation [47, 49]. Searching for a "Fisher wave" solution of Eq. 27 in the form of a traveling wave $u(x,t) = U(z = x \pm ct)$, where c is a wave velocity to be determined [66], yields:

$$\frac{d^2U}{dz^2} \pm c\frac{dU}{dz} + f(U) = 0\,. \tag{28}$$

It is known [68] that if $f'(0) > 0$, $f'(1) < 0$, and $f(x) > 0$ for all $x \in]0,1[$, then monotonic wave front solutions such that $\lim_{z\to-\infty} U(z) = 0$ or 1, and $\lim_{z\to\infty} U(z) = 1$ or 0, exist if and only if $|c| \geq c^*$, where $c^* > 0$ is a constant that can be determined by analyzing the stability of the fixed points at $(U, U') = (0,0)$ and $(U, U') = (1,0)$. In particular, $c^* = 2$ in the case of the Fisher equation [49].

In our model, steady-state profiles of ϕ are solutions of Eq. 21, and thus, using $\xi = x/D$ instead of $s = xv/D$, the function $\psi = 1 - \phi$ satisfies

$$\frac{d^2\psi}{d\xi^2} - v\frac{d\psi}{d\xi} + rD\frac{\psi(1-\psi)}{\kappa + 1 - \psi} = 0\,, \tag{29}$$

which has the form of Eq. 28, with v replacing $\mp c$, and with $f : \psi \mapsto rD\,\psi(1-\psi)/(\kappa+1-\psi)$. This function satisfies $f(0) = f(1) = 0$, as well as $f'(0) > 0$, $f'(1) < 0$, and $f(x) > 0$ for all $x \in]0,1[$, since r, D and κ are positive constants. Therefore, monotonic stationary solutions such that $\lim_{z\to-\infty}\psi(z) = 0$ or 1, and $\lim_{z\to\infty}\psi(z) = 1$ or 0, exist if and only if $v \geq v^*$ (recall that in our system $v > 0$), and the same result holds for $\phi = 1 - \psi$.

In order to determine v^*, let us study the stability of the fixed points at $(\phi, \phi') = (0,0)$ and $(\phi, \phi') = (1,0)$ in Eq. 21. Introducing $\chi = \phi'$, where the prime denotes a derivative with respect to $s = xv/D$, Eq. 21 becomes:

$$\begin{cases} \phi' = \chi\,, \\ \chi' = \chi + \lambda\,\phi(1-\phi)/(\kappa+\phi)\,. \end{cases} \tag{30}$$

A linear stability analysis demonstrates that $(0,0)$ is a saddle point, while $(1,0)$ is either an unstable node if $4\lambda < \kappa + 1$ (see Fig. S2A) or an unstable spiral if $4\lambda > \kappa + 1$ (see Fig. S2B and C). Qualitative phase space analysis (similar to that presented e.g. in [49] for the Fisher equation) then shows that stationary solutions such that $\lim_{s\to-\infty}\phi(s) = 1$ and $\lim_{s\to\infty}\phi(s) = 0$ exist in both cases, but that they are monotonic if and only if $4\lambda \leq \kappa + 1$, or equivalently

$$v \geq v^* = \sqrt{\frac{4rD}{\kappa+1}}\,. \tag{31}$$

Therefore, in infinite space, there exists a stationary solution to the system of partial differential equations Eqs. 1 that is equivalent to a Fisher wave (in a moving frame), provided that Eq. 31 is satisfied. However, the velocity v is a parameter of the system, and not a wave velocity that can adjust, in contrast to the usual Fisher-KPP case.

The model studied in this paper employs the system of partial differential equations Eq. 1 in a finite segment $x \in [0, L]$ (or equivalently, $s \in [0, \sigma = Lv/D]$), with the boundary conditions Eq. 2, which become in dimensionless form:

$$\begin{cases} \chi(0) = \phi(0) - 1\,, \\ \chi(\sigma) = 0\,. \end{cases} \tag{32}$$

Is there a nontrivial solution satisfying these conditions, and if yes, how does it compare to the Fisher-wave-like monotonic solution such that $\lim_{s\to-\infty}\phi(s) = 1$ and $\lim_{s\to\infty}\phi(s) = 0$?

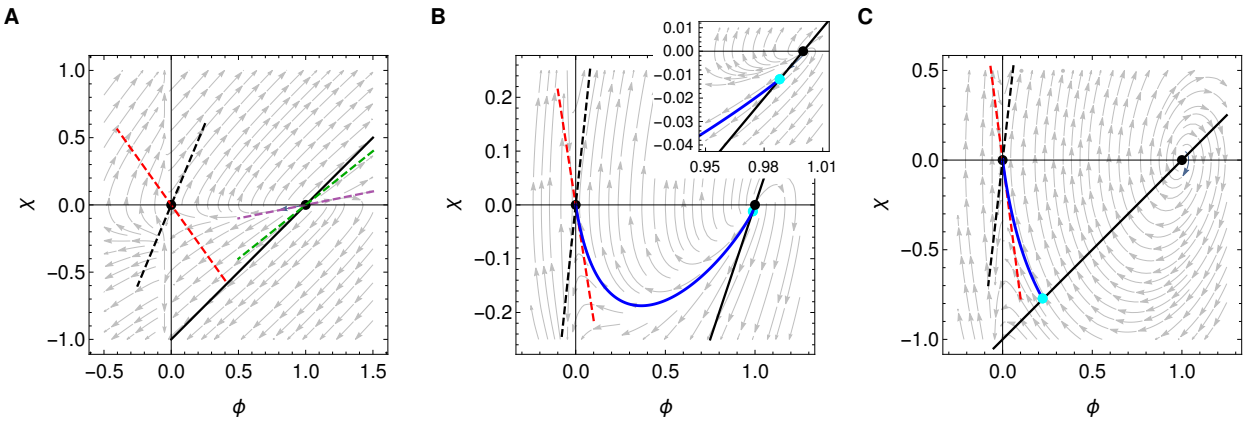

Figure S2: Phase space analysis of the system. **A**: $\lambda = 0.17$ and $\kappa = 0.05$, such that $v > v^*$ (washout). **B**: $\lambda = 0.34$ and $\kappa = 0.05$, such that $v < v^*$ (below washout, but close to it), corresponding to the parameters of Fig. 1B. **C**: $\lambda = 3.36$ and $\kappa = 0.05$, such that $v < v^*$ (substantially below washout). In all panels: Arrows: stationary solutions. Blue curve: stationary solution obtained by numerical integration of PDE Eqs. 1. Black markers: fixed points. Dashed lines: directions of the Jacobian eigenvectors at these fixed points (at $(0,0)$: stable direction in red, unstable one in black; at $(1,0)$ if it is an unstable node (i.e. in **A**): slow unstable direction in purple, fast one in green). Black line: boundary condition in $s = 0$, namely $\chi = \phi - 1$. Cyan dot: Crossing of the numerical solution (blue curve) and the boundary condition (black line). Parameters of the numerical solutions are $D = 0.1$, 0.2, 2.0 cm^2/h in panels **A** , **B**, and **C**, respectively. Other parameters are as in Fig. 1B. The inset in panel **B** shows a close-up in the vicinity of $(1,0)$.

First assume that Eq. 31 is satisfied, ensuring that the Fisher-wave-like solution exists (see Fig. S2A). Then, $(1,0)$ is an unstable node and the smallest eigenvalue of the associated Jacobian is $\ell_- = [1-\sqrt{1-4\lambda/(\kappa+1)}]/2$. This eigenvalue satisfies $0 \leq \ell_- \leq 1/2$, and is associated to the eigenvector $(1, \ell_-)$, yielding a slow direction in the phase space with slope between 0 and 1/2. Close to $(1,0)$, the Fisher-wave-like solution follows this slow direction, before curving upwards to finally approach the stable direction of the saddle point in $(0,0)$. Therefore, it only crosses the line $\chi = \phi - 1$ corresponding to the boundary condition in $s = 0$ (see Eq. 32) at the unstable node $(1,0)$. But this unstable node can only be approached at $s \to -\infty$ in nontrivial solutions, which makes it impossible to find a nontrivial solution satisfying Eq. 32. And indeed, our numerical resolutions confirm that if Eq. 31 is satisfied, bacteria are washed out (see Fig. 1C and section S4).

Let us turn to the opposite case, where $v < v^*$ (see Fig. S2B and C). Then, $(1,0)$ is an unstable spiral, and thus the stationary solution such that $\lim_{s\to-\infty} \phi(s) = 1$ and $\lim_{s\to\infty} \phi(s) = 0$ is oscillatory and comprises unphysical parts where $\phi > 1$ (implying negative numbers of bacteria). However, this spiraling behavior means that this special solution crosses the line $\chi = \phi - 1$ (see Eq. 32) in at least another point, say P, than the unstable node $(1,0)$. Thus, there may be nontrivial and physical solutions satisfying Eq. 32, among the solutions that intersect the line $\chi = \phi - 1$ between $(1,0)$ and P (examples of these intersections are shown in cyan in Figs. S2, S3 and S4). These solutions may look quite different from the Fisher-wave like solution that exists for $v \geq v^*$. In particular, they may possess a value of $\phi(0)$ substantially smaller than 1 and a value of $\phi(\sigma)$ substantially larger than 0. These differences are expected to increase as v is decreased further and further from v^* (i.e. more differences from the Fisher-wave-like solutions are expected in Fig. S2B than in Fig. S2C). Consistently, Fig. 1C demonstrates that profiles with strongest spatial dependence are obtained for v not much below the washout limit $v = v^*$, and Fig. S1 includes several numerical solutions with values substantially different from 1 and 0 at the boundaries $s = 0$ and $s = \sigma$. Fig. S3 shows the phase portraits corresponding to some of these cases, illustrating the diversity of solutions. Nevertheless, in the strongly spatial regime where v is not much below the washout limit $v = v^*$, similarities with the Fisher-wave-like solutions that exist for $v > v^*$ are expected. This case is illustrated e.g. by Fig. S3B.

Here, differences with Fisher waves are brought by the boundary conditions, and in particular by the one at $x = 0$. In addition to the different parameter regime where nontrivial solutions exist, another difference is that all horizontal translations of a traveling wave solution are also solutions in the Fisher wave case [49], while here, they would not satisfy the boundary conditions. The boundary condition at $x = 0$ is quite important as it corresponds to the transition between the small intestine and the colon. Besides, the boundary condition in $x = L$ also affects the solution, in particular by inducing a second washout limit, see Fig. 1C and section S4. However, for sufficiently large $\sigma = Lv/D$, the value of σ has little impact on the solution (see Fig. S14). Fig. S4 shows phase portraits with different values of L, illustrating its impact for small L, and its lack of impact for $L \geq 4$ cm.

Note that traveling waves arising in Fisher-KPP equations Eq. 27 (including those characterizing expanding

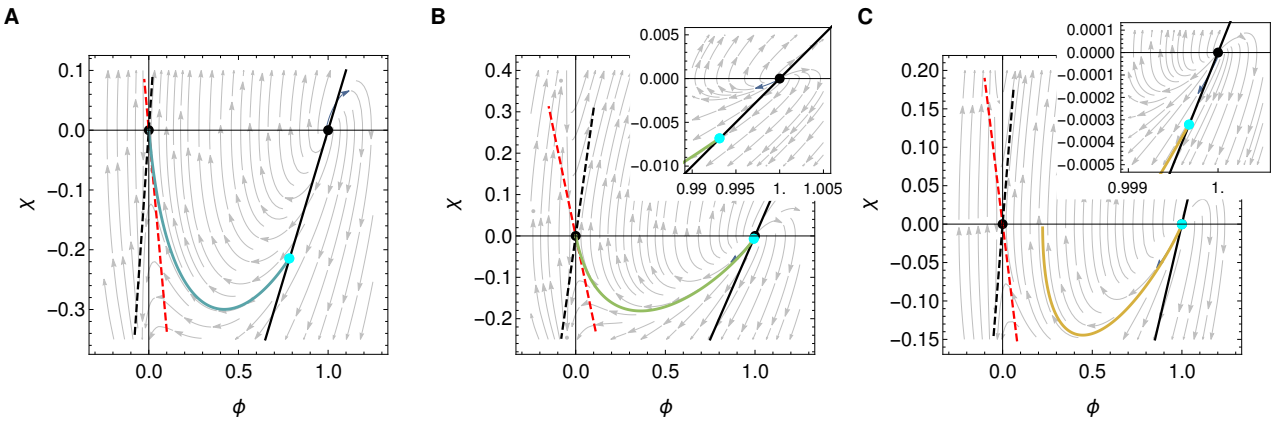

Figure S3: Phase portraits for different types of spatial profiles from Fig. S1B. A: $\lambda = 0.56$ and $\kappa = 0.039$. **B:** $\lambda = 0.32$ and $\kappa = 0.051$. **C:** $\lambda = 0.29$ and $\kappa = 0.054$. Parameters of the numerical solutions are $v = 0.386, 0.509, 0.536$ cm/h in panels **A** , **B**, and **C**, respectively. Other parameters are as in Fig. 1B. Symbols are the same as in Fig. S2 and colors of curves representing numerical solutions match those of Fig. S1B. Insets in panels **B** and **C** are close-ups in the vicinity of $(1, 0)$.

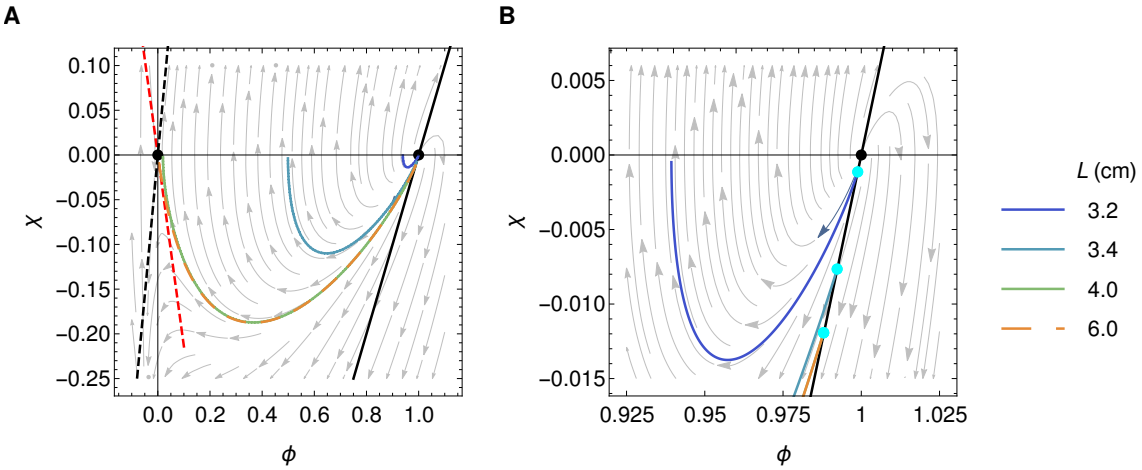

Figure S4: Phase portraits for different system lengths L. All other parameters are as in Fig. 1B. Symbols are the same as in Fig. S2 and colors of curves representing numerical solutions match those of Fig. S11B. Panel **B** is a close-up of panel **A** in the vicinity of $(1, 0)$.

populations) can be either driven by the leading edge (pulled) or by the bulk of the wave (pushed) [57, 61]. Pulled waves have a velocity that depends only on $f(0)$, and are obtained when $f'(u) \leq f'(0)$ for all $u \in [0, 1]$ and $f'(u) = f'(0) + O(h^p)$ for some $p > 0$ when $u \to 0$, in addition to the above conditions $f'(0) > 0$, $f'(1) < 0$, and $f(x) > 0$ for all $x \in]0, 1[$ [69, 70]. These conditions are all satisfied by f in Eq. 29, and thus the associated Fisher-KPP equation would lead to pulled waves. Pushed waves have a velocity that depends on the full nonlinearity of f, and can be obtained when growth rates are nonmonotonic with population density, which can occur with cooperativity (Allee effect) [57, 61]. However, here, the velocity v is an imposed parameter, in contrast to a traveling wave velocity, and the boundary conditions yield further differences, as discussed above.

Note also that it would be interesting to further investigate the stability of solutions to time-dependent perturbations, as was done e.g. for the Fisher wave in the Fisher equation [49]. Here, the analysis would be different as one would need to start from the system of partial differential equations Eqs. 1.

S4 Washout limits

The washout limits are the limits where all bacteria get washed out of the system. Mathematically, they correspond to a bifurcation point in the parameter space where the trivial steady state solution $F(x) = F_{\text{in}}$ (and $B(x) = 0$) becomes stable.

In the chemostat, this limit is easy to determine. There are two steady state solutions, $F^{\text{c}} = kc/(r - c)$ and $F^* = F^{\text{c}}_{\text{in}}$. Eigenvalues of the Jacobian associated to Eqs. 41 and 42 for both steady states change their sign at

same point in the parameter space, namely at

$$c = \frac{r}{k/F_{\text{in}}^{\text{c}} + 1} \equiv c_{\text{wo}}. \tag{33}$$

The bifurcation scenario is such that at c_{wo} the two steady state solutions collide, change stability, and one (nontrivial) solution disappears, making it a transcritical bifurcation.

In the spatial system, a complete analytical treatment is more difficult, even though one of the washout limits was found by our phase portrait analysis above (see Eq. 31). Numerically, we find one positive nontrivial solution for $v < v_{\text{wo}}$, where v_{wo} is the bifurcation point which depends on the rest of the system parameters, and only the trivial solution for $v > v_{\text{wo}}$, implying a change of the stability at $v = v_{\text{wo}}$ when the two steady states collide.

It is possible to find analytical estimates for the washout limits in the spatial system by comparing key length scales and time scales, as discussed in Ref. [39]. Let us first consider the case of large diffusion coefficients, when concentration profiles are flat regardless of the value of v. We compare the time τ_{flow} needed for a bacterium to travel through the system from the entrance to the exit to the minimal time τ_{repl} taken by a bacterium to replicate in this system (for $F = F_{in}$). If $\tau_{\text{flow}} < \tau_{\text{repl}}$, i.e.

$$\frac{L}{v} < \frac{\kappa + 1}{r}, \tag{34}$$

bacteria get washed out, so that an estimate of the washout velocity is

$$v_{\text{wo}} = \frac{rL}{\kappa + 1}, \tag{35}$$

which is in good agreement with the numerical results, as shown by Figure S5, and Figure 1C in the main text.

Next, let us compare the diffusion and flow characteristic lengths at the time of replication, τ_{repl}. Washout occurs if $L_{\text{diff}} < L_{\text{flow}}$, i.e.

$$\sqrt{2D\tau_{\text{repl}}} < v\tau_{\text{repl}}, \tag{36}$$

$$\sqrt{2\frac{D(\kappa + 1)}{r}} < \frac{v(\kappa + 1)}{r}, \tag{37}$$

which gives

$$v_{\text{wo}} = \sqrt{\frac{2rD}{\kappa + 1}}. \tag{38}$$

This second washout limit is in agreement with Eq. 31, as well as with the numerical results depicted in Figures S5 and 1, up to a factor 2. Indeed, numerically, we find a good agreement with

$$v_{\text{wo}} = v^* = \sqrt{\frac{4rD}{\kappa + 1}}. \tag{39}$$

The results in Figure S5 are consistent with those of Figure 1C where $(\kappa + 1) \leq 1.25$, and demonstrate the robustness of these results across a wide range of κ values.

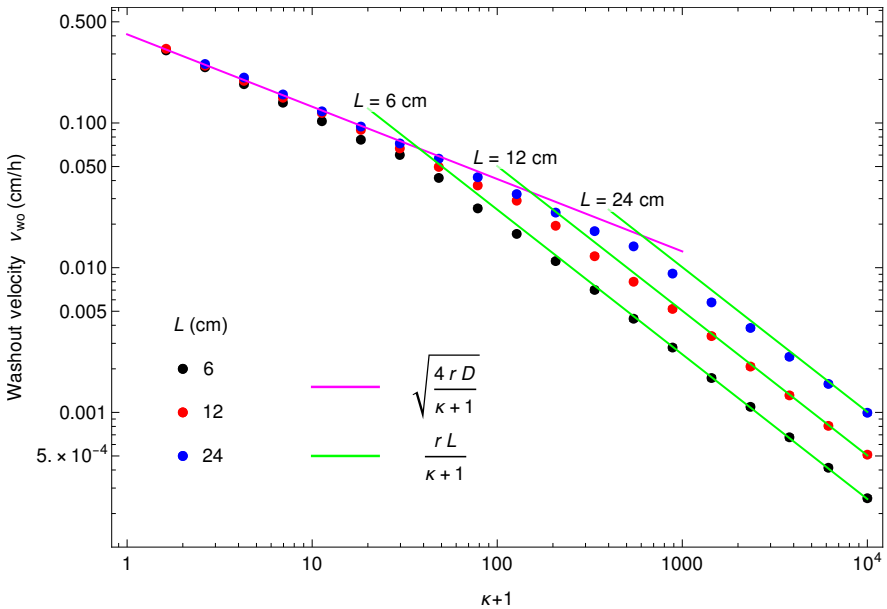

Figure S5: Washout limits in the spatial system as a function of $\kappa + 1$. The two washout limits are obtained by fixing the diffusion coefficient at $0.1\,\mathrm{cm}^2/\mathrm{h}$, and varying the velocity for 20 different values of $\kappa + 1$. The first velocity for which $F(L) > 1. - 10^{-10}$ is recorded as the washout velocity v_{wo}. The process is repeated for three different values of length $L = 6, 12, 24$ cm. Other parameters are $v\,F_{\mathrm{in}} = 1\,\mathrm{mM\,cm/h}$, $k = \kappa F_{\mathrm{in}}$ mM, $r = 0.42\,\mathrm{h}^{-1}$, $\alpha = 6.13 \times 10^8$ bacteria/mM. The green and magenta lines correspond to the two washout limits described by Eqs. 35 and 39, respectively.

S5 Correspondence between the spatial system and the chemostat

S5.1 Main matching condition

We wish to match the total number of divisions occurring in the spatial system and in the chemostat. At stationary state we thus aim to match the amount of food entering and exiting these different systems, as well as the amount of bacteria exiting them.

- Spatial system: amount of food entering: $dN_{\mathrm{F,in}}/dt = F_{\mathrm{in}} v S$; amount of food exiting: $dN_{\mathrm{F,out}}/dt = F(L) v S$; amount of bacteria exiting: $dN_{\mathrm{B,out}}/dt = B(L) v S$;
- Chemostat: amount of food entering: $dN_{\mathrm{F,in}}/dt = F^{\mathrm{c}}_{\mathrm{in}} c V$; amount of food exiting: $dN_{\mathrm{F,out}}/dt = F^{\mathrm{c}} c V$; amount of bacteria exiting: $dN_{\mathrm{B,out}}/dt = B^{\mathrm{c}} c V$.

Here V is the volume of the chemostat, c is the dilution rate i.e. the outflow rate per unit volume of the chemostat. Concentrations in the chemostat are indicated by a superscript c. Meanwhile, v denotes the velocity in the spatial system, S the section and L the length of the spatial system.

Hence, our matching condition reads:

$$\frac{F_{\mathrm{in}}}{F^{\mathrm{c}}_{\mathrm{in}}} = \frac{F(L)}{F^{\mathrm{c}}} = \frac{B(L)}{B^{\mathrm{c}}} = \frac{cV}{vS} \,. \tag{40}$$

S5.2 Constraints from each separate system

In the spatial system, we have $F(L) = F_{\mathrm{in}} \phi(\kappa, \lambda, \sigma)$ and $B(L) = \alpha \left[F_{\mathrm{in}} - F(L)\right]$, i.e. $B(L) = \alpha F_{\mathrm{in}} \left[1 - \phi(\kappa, \lambda, \sigma)\right]$, where ϕ is the dimensionless function introduced above.

In the chemostat, the following equations are satisfied:

$$\frac{dF^{\mathrm{c}}}{dt} = -\frac{r}{\alpha} \frac{F^{\mathrm{c}} B^{\mathrm{c}}}{k + F^{\mathrm{c}}} + c F^{\mathrm{c}}_{\mathrm{in}} - c F^{\mathrm{c}} \,, \tag{41}$$

$$\frac{dB^{\mathrm{c}}}{dt} = r \frac{F^{\mathrm{c}} B^{\mathrm{c}}}{k + F^{\mathrm{c}}} - c B^{\mathrm{c}} \,. \tag{42}$$

At stationary state

$$0 = -\frac{r}{\alpha} \frac{F^{\mathrm{c}} B^{\mathrm{c}}}{k + F^{\mathrm{c}}} + c F^{\mathrm{c}}_{\mathrm{in}} - c F^{\mathrm{c}} \,, \tag{43}$$

$$0 = r \frac{F^{\mathrm{c}} B^{\mathrm{c}}}{k + F^{\mathrm{c}}} - c B^{\mathrm{c}} \,. \tag{44}$$

If $B^{\mathrm{c}} \neq 0$, this yields $r F^{\mathrm{c}}/(k + F^{\mathrm{c}}) = c$, which means that the dilution rate c of the chemostat sets the effective division rate of the bacteria, a fundamental chemostat property. And then (with $c \neq 0$)

$$F^{\mathrm{c}} = \frac{kc}{r - c} \,, \tag{45}$$

$$B^{\mathrm{c}} = \alpha \left(F^{\mathrm{c}}_{\mathrm{in}} - F^{\mathrm{c}}\right) = \alpha F^{\mathrm{c}}_{\mathrm{in}} \left(1 - \frac{k}{F^{\mathrm{c}}_{\mathrm{in}}} \frac{c}{r - c}\right) \,. \tag{46}$$

Hence, our matching condition Eq. 40 reads:

$$\frac{F_{\mathrm{in}}}{F^{\mathrm{c}}_{\mathrm{in}}} = \frac{F_{\mathrm{in}} \phi(\kappa, \lambda, \sigma)}{kc/(r - c)} = \frac{\alpha F_{\mathrm{in}} \left[1 - \phi(\kappa, \lambda, \sigma)\right]}{\alpha F^{\mathrm{c}}_{\mathrm{in}} \left[1 - k/F^{\mathrm{c}}_{\mathrm{in}}\, c/(r - c)\right]} = \frac{cV}{vS} \,, \tag{47}$$

which reduces to two equations relating the spatial system (left hand-side) to the chemostat (right hand-side):

$$\phi(\kappa, \lambda, \sigma) = \frac{k}{F^{\mathrm{c}}_{\mathrm{in}}} \frac{c}{r - c} \,, \tag{48}$$

$$vS F_{\mathrm{in}} = cV F^{\mathrm{c}}_{\mathrm{in}} \,. \tag{49}$$

We assume that the parameters of the spatial system are given. Then we need to choose those of the chemostat in order to have a good matching. The parameters specific to the chemostat are $c, F^{\mathrm{c}}_{\mathrm{in}}, V$. Note that k and r are assumed to be the same in both systems. In principle Eqs. 48 and 49 allow us to fix 2 out of these 3 free chemostat parameters.

S5.3 Additional matching conditions

We may want to impose additional matching conditions between the chemostat and the spatial system:

1. Same total volume: $V = SL$. This implies $\frac{cV}{vS} = \frac{cL}{v}$, and Eq. 40 would be modified accordingly.
2. Same volume exiting per unit time: $cV = vS$. This implies $\frac{cV}{vS} = 1$, and Eq. 40 would be modified accordingly.
3. Same outflow rate relative to the total volume: $c = \frac{vS}{LS} = \frac{v}{L}$. This implies $\frac{cV}{vS} = \frac{V}{LS}$, and Eq. 40 would be modified accordingly.

We note that if we impose two of these three conditions simultaneously, then the third one is also satisfied automatically, but Eqs. 48 and 49 and all three conditions above can be satisfied simultaneously only when $F(L) = vk/(rL - v)$.

S5.4 Properties of the matching chemostats

In general, we can impose only 3 independent conditions, setting the values of $c, F^{\mathrm{c}}_{\mathrm{in}}, V$. Specifically, we have to take Eqs. 48 and 49, plus one of the three conditions numbered 1, 2 and 3 above. These three possibilities are discussed in Table S1 and illustrated in Figure S6.

Matching condition		c	$F^{\mathrm{c}}_{\mathrm{in}}$	V
1. Same total volume	$\frac{cV}{vS} = \frac{cL}{v}$	$\frac{F(L)v}{2Lk}\left(\sqrt{\frac{4Lrk}{F(L)v}+1}-1\right)$	$\frac{F_{\mathrm{in}}v}{2Lr}\left(1+\sqrt{\frac{4Lrk}{F(L)v}+1}\right)$	$L\,S$
2. Same volume exiting per unit time	$\frac{cV}{vS} = 1$	$\frac{F(L)r}{k+F(L)}$	F_{in}	$\frac{v\,S(F(L)+k)}{F(L)r}$
3. Same outflow rate relative to the total volume	$\frac{cV}{vS} = \frac{V}{SL}$	$\frac{v}{L}$	$\frac{F_{\mathrm{in}}kv}{F(L)(Lr-v)}$	$\frac{F(L)\,L\,S(Lr-v)}{kv}$

Table S1: Chemostat parameters $c, F^{\mathrm{c}}_{\mathrm{in}}, V$ as a function of the parameters of the spatial system for three different matching conditions.

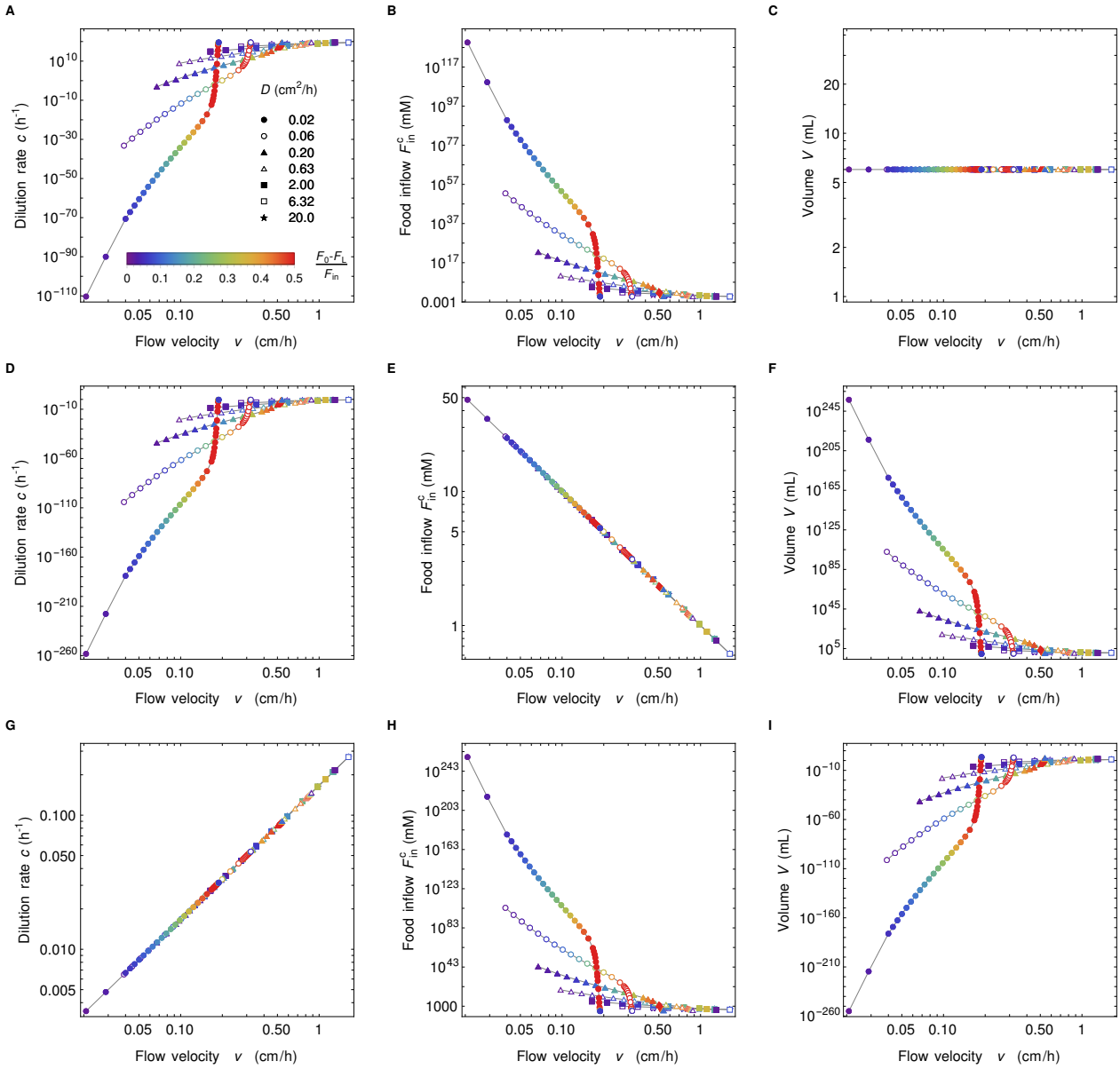

Figure S6: Chemostat parameters. Parameters of the chemostat matching the spatial system in the conditions of Figure 3. Each row in the figure (top to bottom) represents matching condition 1 to 3, while each column in the figure represents a given parameter: dilution rate c (left), food inflow F_{in}^{c} (middle), and volume V (right). Due to the very small values of food concentration exiting the gut, parameters of the chemostat system can have very large or very small values. This is particularly true for smaller diffusion constants and/or small velocities.

Finally, table S2 gives the expression of various useful quantities for the spatial system and for the chemostat.

	Spatial system	Chemostat
Reproduction rate	$\rho(x) = r\dfrac{F(x)}{k+F(x)}$	$\rho^{\mathrm{c}} = r\dfrac{F^{\mathrm{c}}}{k+F^{\mathrm{c}}}$
Reproductions per unit volume and unit time	$R(x) = B(x)\rho(x)$ $= r\alpha\dfrac{F(x)\left[F_{\mathrm{in}} - F(x)\right]}{k+F(x)}$	$R^{\mathrm{c}} = B^{\mathrm{c}}\rho^{\mathrm{c}}$ $= r\alpha\dfrac{F^{\mathrm{c}}\left(F^{\mathrm{c}}_{\mathrm{in}} - F^{\mathrm{c}}\right)}{k+F^{\mathrm{c}}}$
Total reproduction rate	$N_{\mathrm{R}} = r\alpha\, S\displaystyle\int_0^L \dfrac{F(x)\left[F_{\mathrm{in}} - F(x)\right]}{k+F(x)}dx$ $= \alpha v S\left[F_{\mathrm{in}} - F(L)\right]$	$N^{\mathrm{c}}_{\mathrm{R}} = r\alpha V\dfrac{F^{\mathrm{c}}\left(F^{\mathrm{c}}_{\mathrm{in}} - F^{\mathrm{c}}\right)}{k+F^{\mathrm{c}}}$ $= \alpha c V\left(F^{\mathrm{c}}_{\mathrm{in}} - F^{\mathrm{c}}\right)$
Total population	$N_{\mathrm{T}} = \alpha S\displaystyle\int_0^L \left[F_{\mathrm{in}} - F(x)\right]dx$	$N^{\mathrm{c}}_{\mathrm{T}} = \alpha V\left(F^{\mathrm{c}}_{\mathrm{in}} - F^{*}\right)$
Active population	$N_{\mathrm{A}} = \alpha S\displaystyle\int_0^{x^*} \left[F_{\mathrm{in}} - F(x)\right]dx, \quad x^* : F(x^*) = k$	$N^{\mathrm{c}}_{\mathrm{A}} = N^{\mathrm{c}}_{\mathrm{T}}$
Fixation probability	$\mathcal{F} = \dfrac{\int_0^L R(x_{\mathrm{M}})\frac{M(x_{\mathrm{M}})}{B(x_{\mathrm{M}})}dx_{\mathrm{M}}}{\int_0^L R(x_{\mathrm{M}})dx_{\mathrm{M}}}$	$\mathcal{F}^{\mathrm{c}} = \dfrac{N_{\mathrm{M}}}{N^{\mathrm{c}}_{\mathrm{T}}}$
Washout limit	$v_{\mathrm{wo}} = \begin{cases} \frac{rL}{k/F_{\mathrm{in}}+1} \\ \sqrt{\frac{4rD}{k/F_{\mathrm{in}}+1}} \end{cases}$	$c = \dfrac{r}{k/F^{\mathrm{c}}_{\mathrm{in}}+1}$

Table S2: Comparison of the main derived quantities in the spatial system and the chemostat.

S6 Early dynamics of mutant bacteria concentration

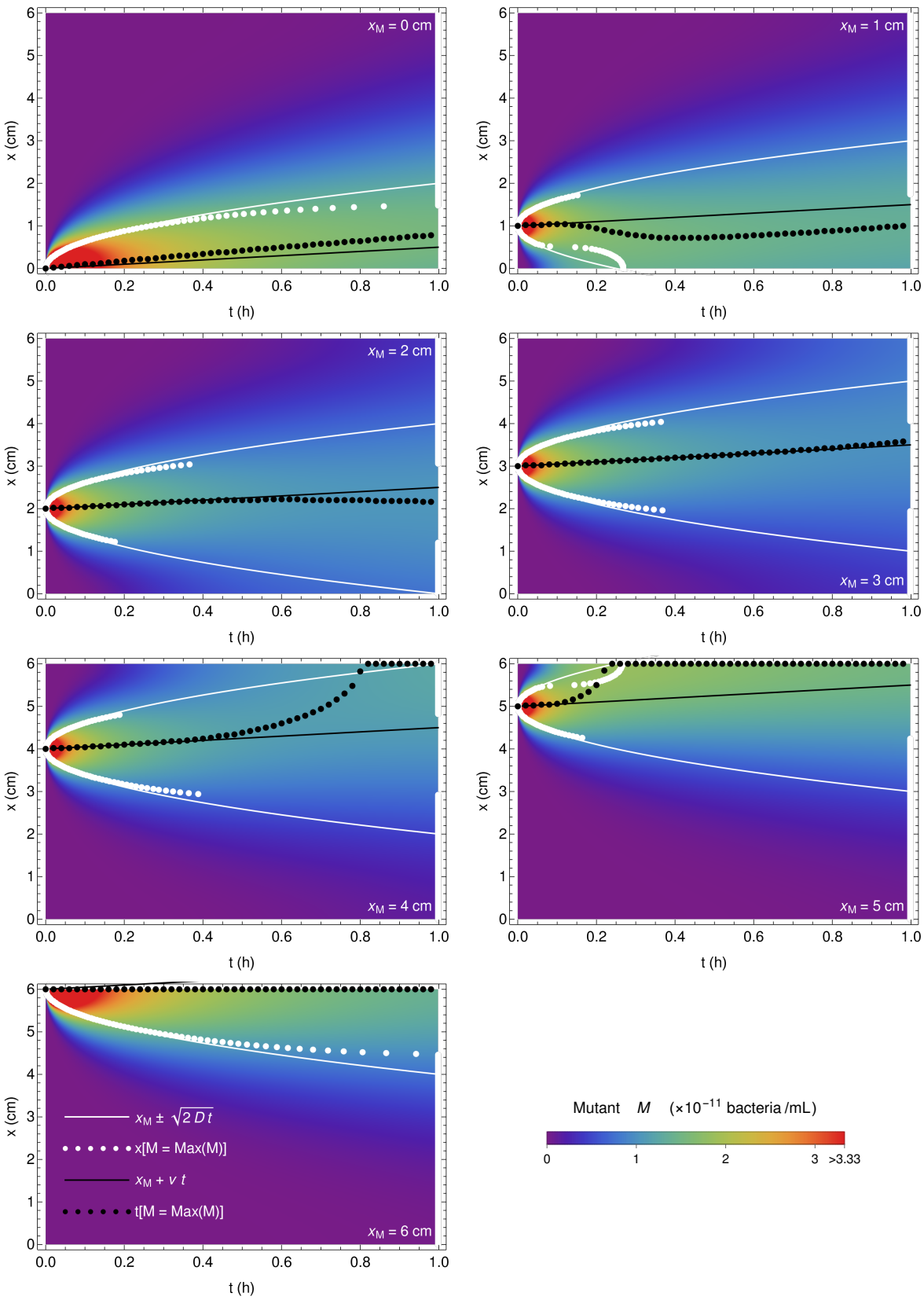

Figure S7: Early dynamics of mutant concentration. Spatio-temporal evolution of mutant concentration for $x_M = 0, 1, 2, 3, 4, 5, 6$ cm in the first hour after mutant introduction. Black points correspond to the maximum mutant concentration at time t, white points correspond to the maximum mutant concentration at position x, black line is $x_M + vt$, and white curve is $x_M \pm \sqrt{2Dt}$. The parameter values are $D = 0.4\,\mathrm{cm}^2/\mathrm{h}$, $v = 0.5\,\mathrm{cm/h}$, $k = 0.1\,\mathrm{mM}$, $r = 0.42\,\mathrm{h}^{-1}$, $v\,F_{\mathrm{in}} = 1\,\mathrm{mM\,cm/h}$, $\alpha = 6.13 \times 10^8$ bacteria/mM and $M_0 = 3.33 \times 10^{-9}$bacteria/mL.

S7 Stationary state of mutant bacteria concentration versus the initial position x_{M} of mutants

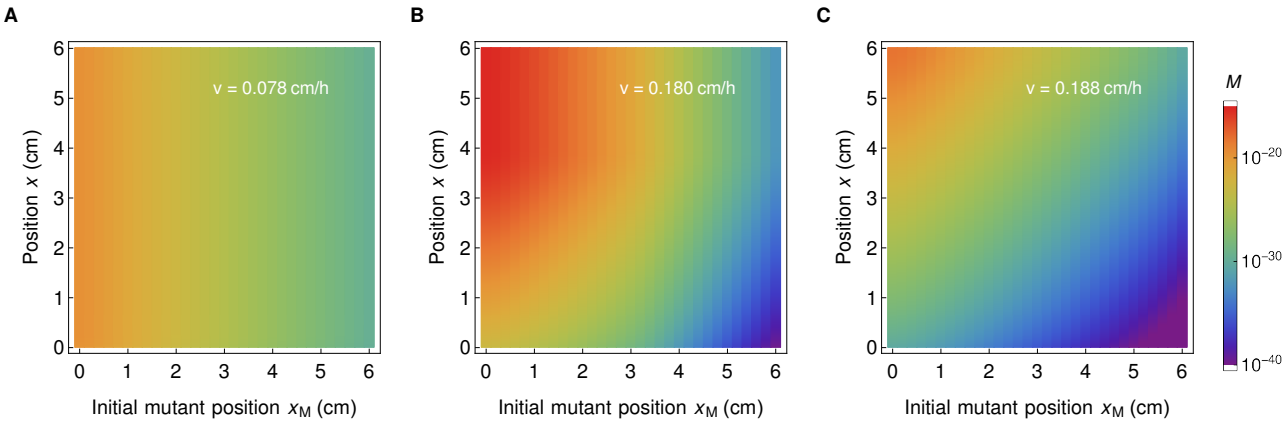

Figure S8: Stationary distribution of mutant bacteria in the gut. Concentration of mutant bacteria as a function of position x, and initial mutant position x_{M}, for $D = 0.02\,\mathrm{cm}^2/\mathrm{h}$, and for three different values of flow velocity v, corresponding to the flat concentration profile in the well mixed regime (**A**), spatial concentration profile (**B**), and close to the washout limit (**C**). In all three cases, we observe that the final concentration of the mutant bacteria is smaller if the initial position x_{M} is further along the gut. This figure corresponds to three points of the top curve in Fig. 3.

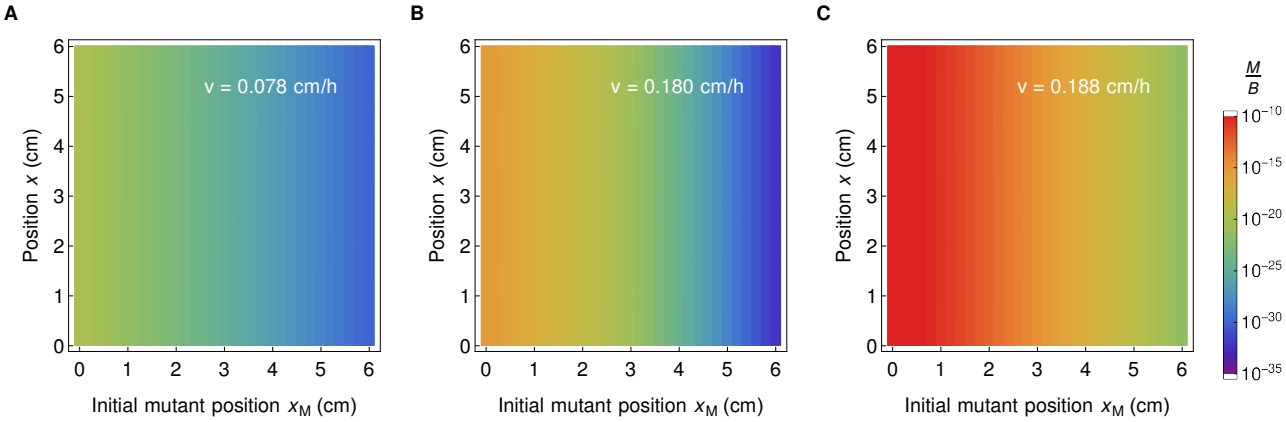

Figure S9: Stationary ratio of mutant to wild type bacteria concentrations. The ratio M/B is shown in all three cases corresponding to Fig. S8. Consistently with our analytical predictions discussed in the main text, the M/B ratio is constant along the gut. In addition, we observe that it monotonically decreases as a function of initial mutant position x_{M}, and the decrease is the most pronounced in the case of the spatial profile (panel **B**) where the ratio at the beginning and at the end of the gut are several orders of magnitude apart.

S8 Location where most mutants that fix originate

While the position of the maximum of the ratio M/B of mutant to wild type bacteria concentrations is always at the entrance of the gut (see main text and Figure 2), the position of the maximum of the number R of reproduction events per unit volume and unit time depends on parameter values. We find that for flat concentration profiles, it is located either close to the entrance of the gut (for small velocities yielding an almost well-mixed system) or at the exit of the gut (close to the washout limit). Conversely, for spatial concentration profiles, its location is intermediate (see Figure S10A). Because of this, in the regime with strong spatial dependence, we find that the position of the maximum of the product $R\,M/B$ of these two quantities ranges between 0 and $L/2$ (see Figure S10B and C). The position of the latter maximum corresponds to the location where most mutants that fix tend to originate.

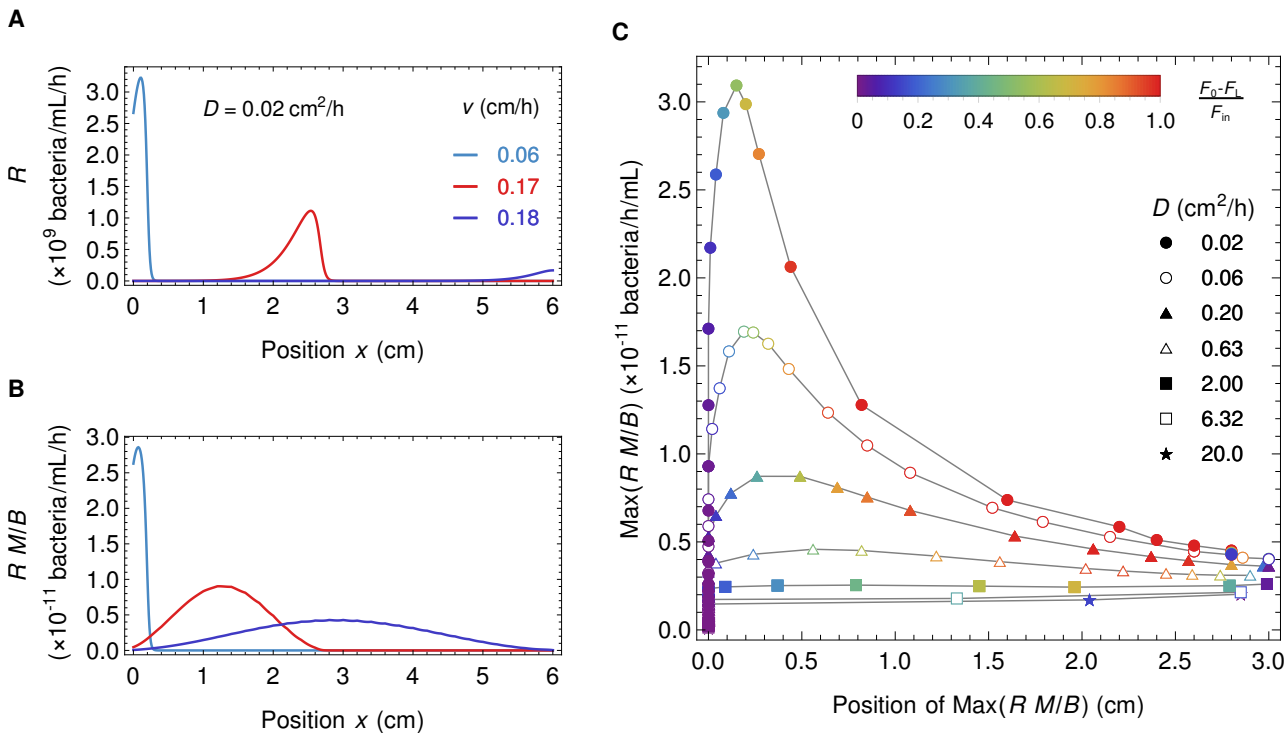

Figure S10: Maximum of R and $R\,M/B$. A: Number R of reproduction events per unit volume and unit time as a function of position along the gut for three different types of concentration profiles, almost well-mixed (light blue), spatial (red), and close to the washout limit (purple). **B:** Maximum of the product $R\,M/B$ of the reproduction events per unit volume and unit time and of the ratio of mutant and wild type bacteria as a function of space for the same three concentration profiles as in A (same colors). **C:** Maximum value of $R\,M/B$ as a function of its position for the data set in Figure 3 of the main text.

S9 Calculation of the active population size

The active population corresponds to the bacteria in the zone where reproduction rate is significant. Concretely, it is defined as

$$N_{\mathrm{A}} = S \int_0^{x^*} B(x)dx\,, \tag{50}$$

where x^* corresponds to the point in the gut segment where the food concentration equals the Monod constant k, i.e. $F(x^*) = k$, implying that the reproduction rate $\rho(x^*)$ is equal to half of the maximal possible reproduction rate, which is obtained if $F(x) \gg k$. Thus,

$$B(x^*) = \alpha[F_{\mathrm{in}} - F(x^*)] = \alpha[F_{\mathrm{in}} - k] = \alpha F_{\mathrm{in}}(1-\kappa)\,. \tag{51}$$

In the active population thus defined, the reproduction rate of bacteria is at least equal to half of its maximal possible value.

S10 Impact of the dimensionless parameters

S10.1 Holding the dimensionless parameters fixed

To illustrate the relevance of the dimensionless parameters introduced in section S2.2 to describe the stationary profiles, we vary system parameters so that we hold the dimensionless ones fixed. The reference for fixing them is Figure 2:

$$\kappa = \frac{k}{F_{\text{in}}} = \frac{0.1}{2.0} = 0.050, \tag{52}$$

$$\lambda = \frac{rD}{v^2} = \frac{0.42 \times 0.2}{0.5^2} = 0.34, \tag{53}$$

$$\sigma = \frac{Lv}{D} = \frac{6 \times 0.5}{0.2} = 15. \tag{54}$$

We first vary the gut length L in order to ease the discretization of the space. Once the L is chosen, the dimensionless parameters are fixed by adjusting v, and then F_{in}, to keep the product $v\,F_{\text{in}}$ constant. Other parameters are kept fixed. Figure S11A shows the corresponding concentration profiles. They all have the same shape, as evidenced by rescaling the food concentration by F_{in} and the spatial coordinate by L (see inset of the Figure S11A). Figure S11B shows that the fixation probability for these profiles scales with active population, $\mathcal{F} = N_{\text{M}}/N_{\text{A}}$, consistently with our expectations, since the concentration profiles are strongly spatially dependent.

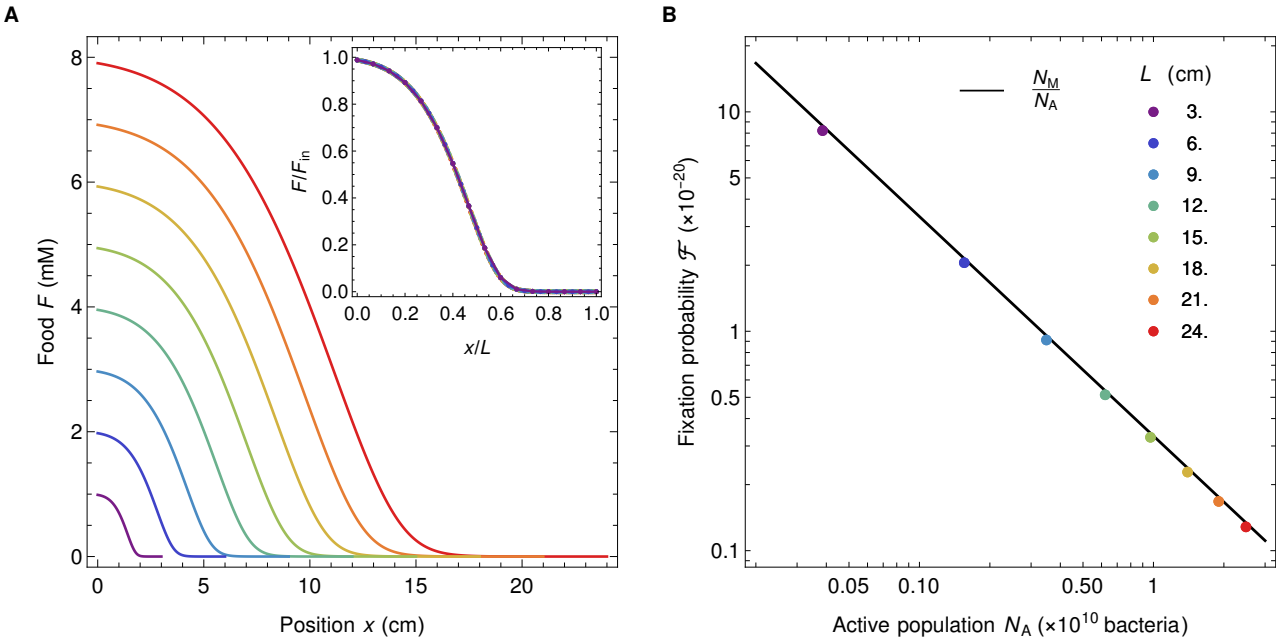

Figure S11: **A** Food concentration profiles for different system sizes and keeping nondimensional parameters, λ, κ, σ, fixed. The profiles can be rescaled by dividing food concentration by F_{in} and space by L (see inset of A) showing all the profiles are identical up to the scaling factors. **B** Fixation probability for the profiles in panel **A**. The parameters are $D = 0.2\,\text{cm}^2/\text{h}$, $v = 3./L$ cm/h, $F_{\text{in}} = 1./v$ mM, $k = 0.05/v$ mM, $r = 1.68 \cdot v^2$ h^{-1}, $\alpha = 6.13 \times 10^8$ bacteria/mM, $N_{\text{M}} = 3.33 \times 10^{-11}$ bacteria and the L values are listed in the panel **B**.

S10.2 Varying each dimensionless parameter

Finally, we systematically investigate the impact of each dimensionless parameter by varying one of them while holding the other two fixed. We keep parameters v, D and F_{in} fixed throughout, and vary k, r, and L one at a time in order to change κ, λ, and σ, respectively. Results in Figures S13, S12 and S14 show that the fixation probability is well described by $\mathcal{F} = N_{\text{M}}/N_{\text{A}}$ except in cases where the food concentration profile is less spatial because food is substantially depleted even at the entrance of the gut, which occurs for the four largest values of r in Figure S13.

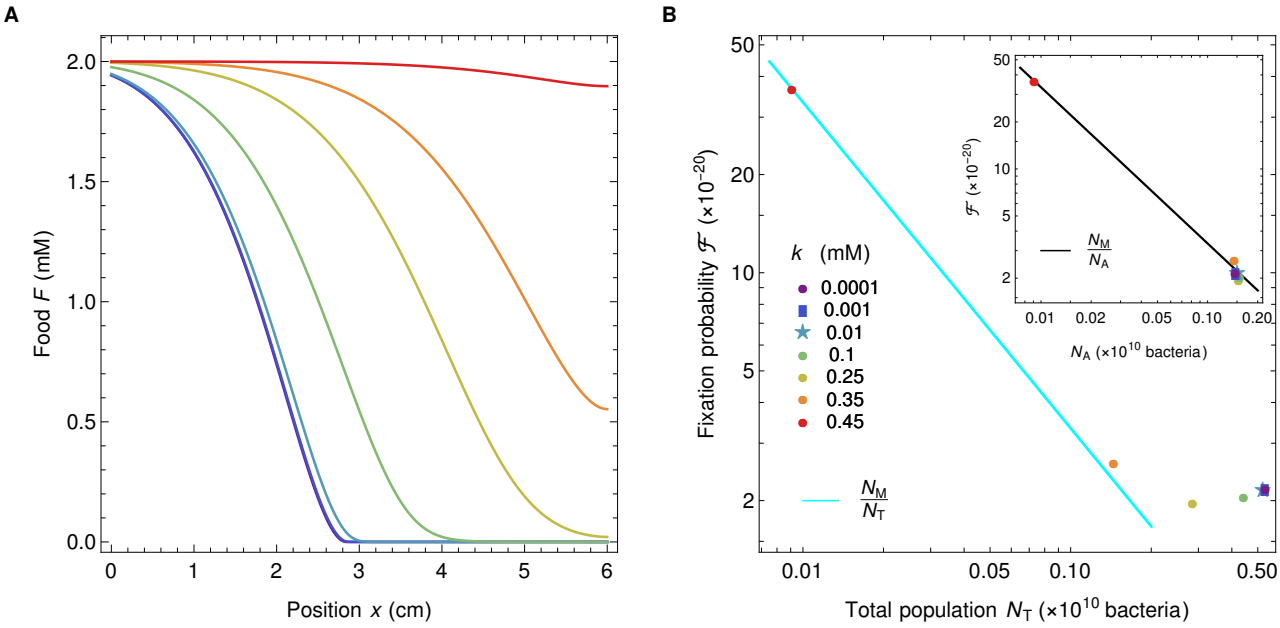

Figure S12: Varying κ. **A** Food profiles for different values of κ and fixed $\lambda = 0.336$ and $\sigma = 15$. **B** Fixation probability as a function of total population N_{T} and active population N_{A} (in the inset). Different colors correspond to different values of κ (k) and different symbols are used for overlapping data points. Parameters are $v = 0.5\,\text{cm/h}$, $D = 0.2\,\text{cm}^2/\text{h}$, $L = 6.0\,\text{cm}$, $r = 0.42\,\text{h}^{-1}$, $\alpha = 6.13 \times 10^8$ bacteria/mM, and $N_{\text{M}} = 3.33 \times 10^{-11}$ bacteria. k values are listed in the panel **B**.

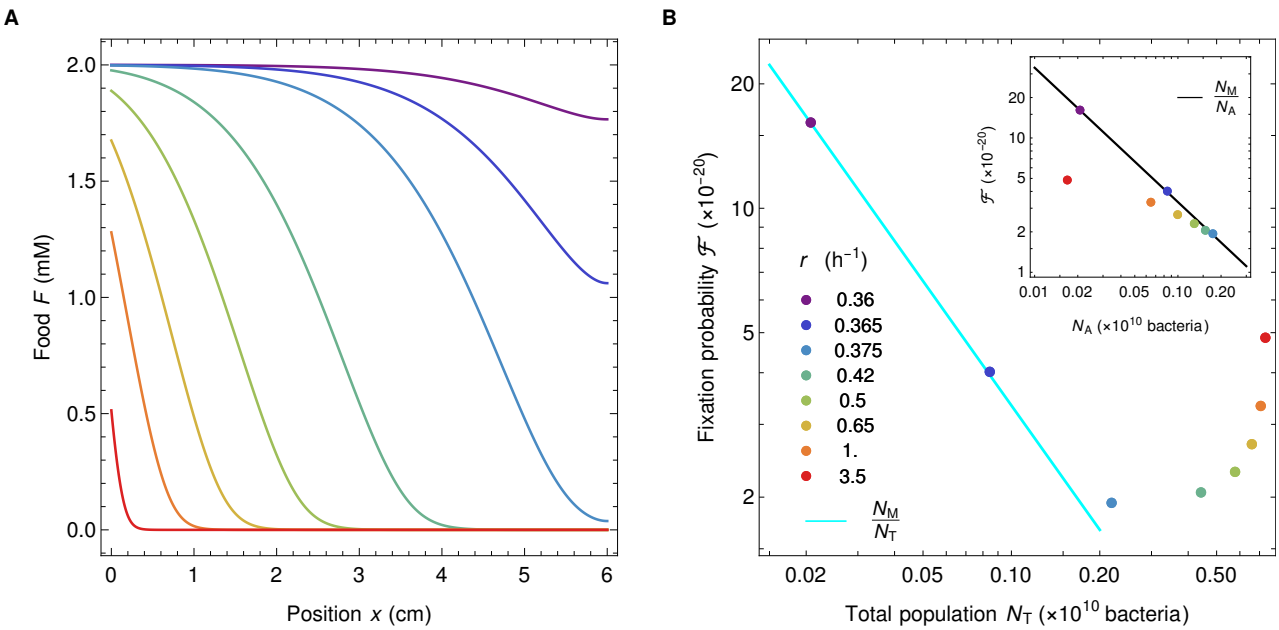

Figure S13: Varying λ. **A** Food profiles for different values of λ and fixed $\kappa = 0.05$ and $\sigma = 15$. **B** Fixation probability as a function of total population N_{T} and active population N_{A} (in the inset). Different colors correspond to different vales of λ (r). Parameters are $v = 0.5\,\text{cm/h}$, $D = 0.2\,\text{cm}^2/\text{h}$, $k = 0.1\,\text{mM}$, $k = 0.1\,\text{mM}$, $\alpha = 6.13 \times 10^8$ bacteria/mM, and $N_{\text{M}} = 3.33 \times 10^{-11}$ bacteria. r values are listed in the panel **B**.

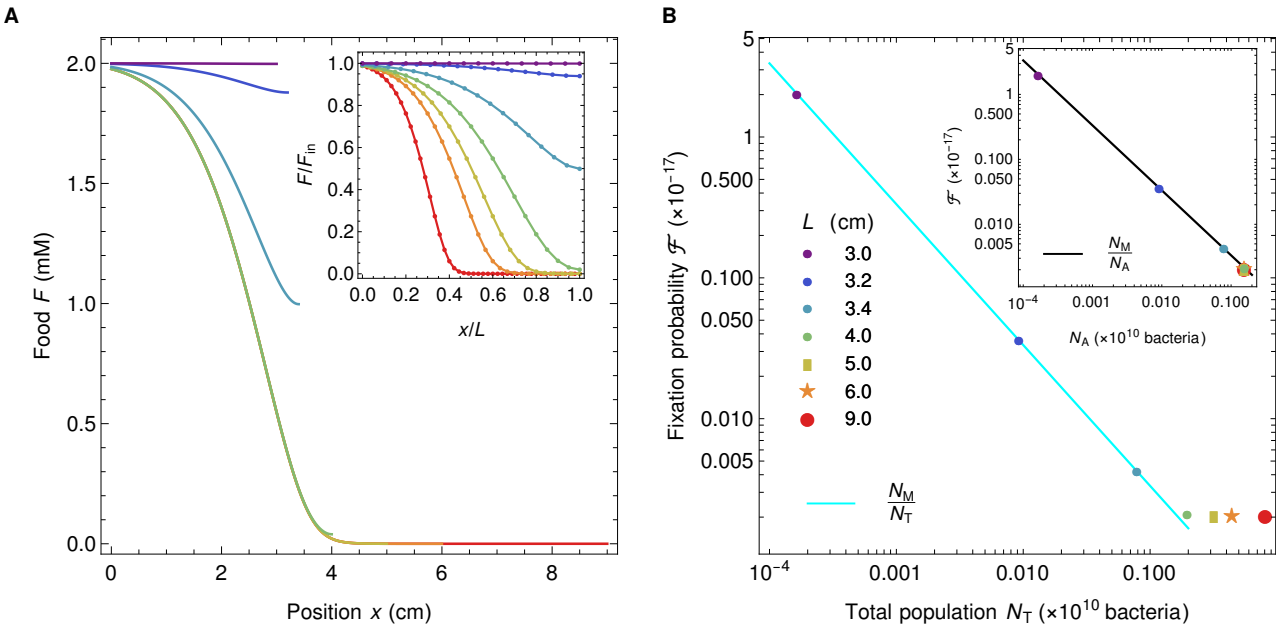

Figure S14: Varying σ. **A** Food profiles for different values of σ and fixed $\lambda = 0.336$ and $\kappa = 0.05$. The inset shows rescaled food profiles, for easier comparison of the spatial dependence. **B** Fixation probability as a function of total population N_{T} and active population N_{A} (in the inset). Different colors correspond to different values of σ (L) and different symbols are used for overlapping data points. Parameters are $v = 0.5\,\mathrm{cm/h}$, $D = 0.2\,\mathrm{cm}^2/\mathrm{h}$, $L = 6.0\,\mathrm{cm}$, $r = 0.42\,\mathrm{h}^{-1}$, $\alpha = 6.13 \times 10^8$ bacteria/mM, and $N_{\mathrm{M}} = 3.33 \times 10^{-11}$ bacteria. k values are listed in the panel **B**.

S10.3 Range of the dimensionless parameters and relevance for the human colon

In this section, we discuss the range of values of the dimensionless parameters κ, λ and σ, which are the only parameters of the model that may change the behavior of the system and affect our conclusions (recall that α is just a scaling factor that does not affect spatial profiles and conclusions, see section S1.3). We first summarize the range of values that these dimensionless parameters take in our study and in Refs. [39, 40]. Next, we discuss the realistic range of values that they can take in the human colon, and show that the range studied here is relevant for the human colon.

The parameter values employed in this paper correspond to the mini gut described in [39], where it has been proven that the mathematical model describes well the experimental setup. In [40], the same model was used to describe microbiota growth and composition in the human colon, and parameter values were thus altered to match the properties of the human gut, which is several times bigger than the mini gut. However, this change of parameter values did not significantly modify the values of the dimensionless parameters. As illustrated in Figure S11, holding the dimensionless parameters fixed fully preserves the properties of the system, including the spatial profiles of concentrations, and our conclusions on the active population remain true. Table S3 summarizes the range of dimensionless parameter values considered in Figure 3 of this paper, and compares them to those of Refs. [39, 40].

Dimensionless parameter	Figure 3	Figure 3; $\frac{F(0)-F(L)}{F_{\mathrm{in}}} > 0.9$	Ref. [39]	Ref. [40]
$\kappa = \dfrac{k}{F_{\mathrm{in}}}$	$10^{-4} - 0.21$	$0.014 - 0.083$	0.05	2.5×10^{-4}
$\lambda = \dfrac{rD}{v^2}$	$0.19 - 8.4 \times 10^6$	$0.24 - 0.42$	$0.036 - 9.3$	0.29
$\sigma = \dfrac{Lv}{D}$	$3 \times 10^{-4} - 57$	$7.59 - 56$	$1.5 - 25$	9.6

Table S3: Dimensionless parameter values used in Figure 3 of this paper and in Refs. [39] and [40].

Let us now discuss the realistic range of values of the dimensionless parameters κ, λ and σ in the human colon. Let us start by considering $\kappa = k/F_{\mathrm{in}}$, and for this, let us first estimate F_{in}. Bacteria in the large intestine consume a mix of different nutrients that have not (or not completely) been absorbed in the small intestine. A

large component of them are fibers, and a human typically ingests 25 to 100 g a day of fibers [4]. The colon input also includes a smaller or similar quantity of unabsorbed sugars and starch [71]. Given that the inflow of digesta in the colon is about 2 liters per day [72], the order of magnitude of the incoming food concentration F_{in} in the colon is in the range of 15 to 100 g/L. The Monod constant k depends on many factors, including the substrate and the bacterial strain. Even for glucose, it can range from 0.03 to 5 mg/L depending on how well adapted the bacteria are to growing on glucose [73], and it is typically higher for other substrates (20 to 300 mg/L for acetate [74]; 5 to 900 mg/L for different substrates [75]). Realistic values for $\kappa = k/F_{\text{in}}$ are thus in a wide range, but in all cases, they are much smaller than 1, at least as small as 0.1. In the next section where dimensionless parameter values are systematically varied, Figure S12 demonstrates that such values of κ give very similar outcomes, and all results collapse in the limit $\kappa \to 0$.

The two other dimensionless parameters, λ and σ, both involve the effective diffusion coefficient D, which models mixing and is thus hard to measure directly. However, it is empirically well demonstrated that there is a strong gradient of bacterial concentration along the colon [76, 77], and that most nutrients that could be used by bacteria are consumed by the end of the gut [71]. This requires that $\lambda = rD/v^2$ be larger than 1/4 (washout limit, see section S4), and less than a few units (since above, diffusion is strong enough for the system to be almost well mixed). This can be seen on Figure 1C. Accordingly, Table S3 demonstrates that despite the wide range of parameters used in our study, the range of λ is very narrow in the regime that yields strongly spatial profiles (specifically, λ is between 0.24 and 0.42 when $[F(0) - F(L)]/F_{\text{in}} > 0.9$). This is in line with the estimate from [40].

Let us finally turn to σ. We notice that $\sigma = Lv/D = Lr/(v\lambda)$. Avoiding washout requires $L/v > (\kappa+1)/r \gtrsim 1/r$ (see section S4), where the last inequality is rather tight because $\kappa \ll 1$. Since in addition $1/\lambda$ is of the order of 2 to 4 in the very spatial regime, $\sigma = Lr/(v\lambda)$ then has to be greater than a few units. Another way to estimate σ is to employ $Lr/v = \tau_{\text{dig}}/\tau_{\text{rep}}$ where $\tau_{\text{rep}} = 1/r$ is the minimal bacterial replication time (within the colon), and $\tau_{\text{dig}} = L/v$ is the typical time spent in the system. While the total transit time ranges from one to several days [78], the time spent by the digesta in the ascending colon, which is the upstream part of the colon where the strong gradients of food and bacterial concentrations are located [40], is substantially shorter, of the order of 4 hours [79], and thus $\tau_{\text{dig}} \approx 4$ h. Let us now estimate τ_{rep}. Feces weigh about 130 g/day and contain 25–50% of bacteria in mass [4]. The typical mass of a bacteria being 1 pg, this means that about 3–6$\times 10^{13}$ bacteria per day are lost in feces, to be compared with about 4×10^{13} bacteria in the colon [1], leading to about one renewal per day, which means that $\tau_{\text{rep}} \leq 24$ h. However, as it is likely that bacteria actively replicate only in the upper part of the colon, while r represents the maximal reproduction rate in the gut, the actual value of τ_{rep} is expected to be substantially smaller than this upper bond. A lower bound for τ_{rep} is given by the minimal doubling time of fast replicating bacteria such as *Escherichia coli* in good conditions, which can be as low as 20 minutes [80]. To summarize, $\sigma = \tau_{\text{dig}}/(\lambda\tau_{\text{rep}})$, with $1/\lambda$ of the order of 2–4 and $\tau_{\text{dig}}/\tau_{\text{rep}} \gtrsim 1$, and $\tau_{\text{dig}}/\tau_{\text{rep}} \approx 10$ when considering the ascending colon and the maximal replication rate. This matches well the range of values of σ considered in the present work (see Table S3).

S11 Relevance of neutral mutations

In this study, we focused on neutral mutations, as a first step towards understanding evolution in the gut. What fraction of spontaneous mutations is expected to be effectively neutral in gut bacteria? Despite the importance of beneficial mutations for adaptation, they are expected to be a small fraction of spontaneous mutations, at least in reasonably well-adapted organisms. Therefore, let us introduce the fitness cost δ of a mutation, defined by $f_M/f_W = 1 - \delta$ where f_M is mutant fitness and f_W wild-type fitness (beneficial mutations then have a negative δ). Mutations are effectively neutral if their fitness cost satisfies $N|\delta| \ll 1$ where N is the effective population size [42], which is the active population size here.

Let us estimate active population size in the human colon, where the total bacterial population is of the order of 4×10^{13} bacteria [1]. The total rate of bacterial divisions in the gut is given by the ratio of the total population size to the renewal time. This rate can also be expressed as the active population size divided by the minimal replication time τ_{rep}, because bacteria in the active population reproduce at rates close to the maximal one. Thus, the ratio of active to total population size is equal to the ratio of minimal replication time to renewal time, which, as discussed in previous section, can be as small as $20/(24 \times 60) = 1.4 \times 10^{-2}$, and is likely to be of order 0.1 (also consistent with [64]). In addition, fixation of a mutant occurs within a strain. Different strains may compete for resources, but here, for simplicity, we do not consider this aspect, which is possible if diversity is stable or if different strains employ different resources. There are at least tens of thousands of bacterial phylotypes in the gut, exact diversity characterizations being limited by sequencing depth [52]. Among all these types, some are much more abundant than others, and multiple rare types exist. Let us consider an active population size of order $N = 10^7$ bacteria.

Let us now discuss the range of δ values. Estimating the full distribution of fitness effects of mutations is difficult [81], and many studies focus on beneficial mutations [82] because of their importance in adaptation. A recent microfluidic mutation accumulation experiment [53] performed in *Escherichia coli* allowed to suppress the effect of selection and to study all spontaneous mutations. The average fitness cost $\langle\delta\rangle$ was found to be 3×10^{-3}, but it largely arises from a minority of strongly deleterious mutations. The skewness of the distribution of fitness costs was found to be positive and large, of about 17. Accurate measurements of extremely small fitness effects are highly challenging, but a Gamma distribution of the same mean and skewness would yield 10% of all mutations with a fitness cost δ smaller than 10^{-7}. Thus, a substantial fraction of all spontaneous mutations occurring in gut bacteria is expected to be effectively neutral.

S12 Validation by stochastic simulations

S12.1 Stochastic simulation methods

In our stochastic simulations, space and time are discretized. We denote the discrete steps in space and time used in these stochastic simulations by δx and δt. Recall that our model is one-dimensional and therefore the only spatial dimension to be considered is x, along the gut main axis. To simplify notations, in this section we denote by $F(x)$ the linear concentration of food and by $B(x)$ that of bacteria (which amounts to taking a gut section with unit area).

Discretizing transport. To represent transport in our discrete simulations, rates of exchange of food and bacteria between adjacent segments of lengths δx are defined.

First, to represent diffusive flow, let us introduce the rate r_d at which food moves from one segment to each one of its two closest neighboring segments (upstream or downstream). For the discrete model to converge to our continuous equations as $\delta x \to 0$, r_d has to satisfy:

$$\lim_{\delta x\to 0} r_d\,\delta x\,[F(x+\delta x)-F(x)] = D\frac{\partial F}{\partial x}\,, \tag{55}$$

where we expressed the diffusive flow per unit time from the segment at x to its immediate downstream neighbor both in the discrete and in the continuous descriptions. This leads to:

$$r_d = \frac{D}{\delta x^2}\,. \tag{56}$$

Similarly, to represent convective flow, let us define the rate r_v at which food moves from one segment to its downstream neighbor. For the discrete model to converge to our continuous equations as $\delta x \to 0$, r_v has to satisfy:

$$\lim_{\delta x\to 0} r_v\,\delta x\,F(x) = v\,F(x)\,, \tag{57}$$

where we expressed the convective flow per unit time from the segment at x to its immediate downstream neighbor both in the discrete and in the continuous descriptions. This leads to:

$$r_v = \frac{v}{\delta x}\,. \tag{58}$$

Then overall, in the discrete model, during δt, the amount of food moving from one segment to its upstream neighbor is $r_-\delta t F(x)\delta x$ where r_- is the rate of upstream transport:

$$r_- = r_d = \frac{D}{\delta x^2}\,. \tag{59}$$

Meanwhile, the amount of food moving from one segment to its downstream neighbor is $r_+\delta t F(x)\delta x$ where r_+ is the rate of upstream transport:

$$r_+ = r_d + r_v = \frac{D}{\delta x^2} + \frac{v}{\delta x}\,. \tag{60}$$

Note that the mapping with the continuous equations is obtained for $\delta x \to 0$, and thus results may deviate for finite δx.

Similarly, during δt, individual bacteria have a probability $r_-\delta t$ to jump upstream, and $r_+\delta t$ to jump downstream.

Discretizing growth. Let us start from our continuous equations on concentrations of bacteria and food, and map them to a description with discrete bacteria. Recall that in this section, B and F are linear concentrations of bacteria and food. Restricting to the terms of Eq. 1 modeling growth, we have:

$$\begin{cases} \dfrac{dB}{dt} &= rB\dfrac{F}{F+k}\,, \\ \dfrac{dF}{dt} &= -\dfrac{1}{\alpha}rB\dfrac{F}{F+k}\,. \end{cases} \tag{61}$$

Denoting by $b(x) = B(x)\delta x$ (resp. $f(x) = F(x)\delta x$) the average number of bacteria (resp. the quantity of food in moles) present in the segment $[x, x+\delta x]$, these equations become:

$$\begin{cases} \dfrac{db}{dt} &= rb\dfrac{f}{f+\tilde{k}}\,, \\ \dfrac{df}{dt} &= -\dfrac{1}{\alpha}rb\dfrac{f}{f+\tilde{k}}\,, \end{cases} \tag{62}$$

$\tilde{k} = k\delta x$. Similarly, let us introduce $f_{\text{in}} = F_{\text{in}}\delta x$.

Each individual bacterium has a reproduction rate of $rf(x)/[f(x)+\tilde{k}]$, so during δt it has a probability $r\delta t\, f(x)/[f(x)+\tilde{k}]$ to reproduce. If one bacterium reproduces, then $f(x)$ is decreased by $1/\alpha$, thereby ensuring the mapping with Eqs. 62. Note that, since this quantity is fixed, for small values of δx, it may lead to transient negative local values of $f(x)$, due to discretization and to the assumption that food is only consumed in the segment where replication occurs. When this happens, we set the local reproduction rate of bacteria to zero, and thanks to diffusion, such unrealistic states are short-lived.

Simulation steps.

- Initialization: The initial discrete numbers $n_b(x)$ of bacteria in each segment $[x, x+\delta x]$ are obtained from the steady-state numerical solution $B(x)$ of the continuous equations through $n_b(x) = \lceil b(x) \rceil = \lceil B(x)\delta x \rceil$, where $\lceil . \rceil$ denotes the ceiling function.
- Equilibration without mutants:
  1. During δt, a bacterium in $[x, x+\delta x]$ has a probability $rf(x)/[f(x)+\tilde{k}]$ of reproducing. For each bacterium, a random number is drawn in a uniform distribution on the interval $[0, 1]$, and compared to the reproduction probability. This determines whether this bacterium reproduces.
  2. For each bacterium within $[x, x+\delta x]$ that reproduces, the quantity of food $f(x)$ is decreased by $1/\alpha$.
  3. After the steps above, food is updated according to the following rule: $f(x) \leftarrow f(x)\,[1-(r_+ + r_-)\delta t] + f(x+\delta x)r_-\delta t + f(x-\delta x)r_+\delta t$, where the rates r_- and r_+ are given respectively by Eq. 59 and Eq. 60. On the boundaries, $f(0) \leftarrow f(0)\,[1-(r_+ + r_-)\delta t] + f(\delta x)r_-\delta t + [v f_{in}/\delta x + r_- f(0)]\,\delta t$, and $f(L) \leftarrow f(L)\,(1 - r_+\delta t) + f(L-\delta x)r_+\delta t$.
  4. For each bacterium at each site x a random number is drawn from a uniform distribution on an interval $[0, 1]$, and compared to the probability to move to the right, $r_+\delta t$, to move to the left, $r_-\delta t$, or to stay, $1-(r_+ + r_-)\,\delta t$. Bacteria move accordingly.

  The above steps 1 to 4 are repeated until time t reaches 1000 hours. A new initial state of bacteria and food is obtained by taking the average in the time interval $t \in [100, 1000]$ h. Then, steps 1 to 4 are repeated for 100 more hours.
- Mutation: At the next replication of one bacterium, we assign to it one mutant daughter bacterium. If more than one bacterium replicates during δt, the mutant daughter bacterium is assigned to one of them, chosen uniformly at random.
- Evolution to fixation or extinction: The above steps 1 to 4 are repeated until extinction or fixation of the mutant bacteria, keeping track of what bacteria are mutant and wild-type. 141,681 different stochastic replicates were run.

Choice of the discretization parameters. Three parameters describe the discrete nature of this stochastic simulation: the discrete spatial step δx, the discrete time step δt, and the finite total number of bacteria $N_{\text{T}} = \sum_x b(x) = \int_0^L B(x)dx$. In the numerical resolutions of the main text, we focused on large bacterial population sizes. But here, as we aim to validate the results from our deterministic model using a stochastic one, we consider less realistic small population sizes, which allows for shorter computation times. However, N_{T}

should not be too small in order to avoid spontaneous extinctions of bacteria due to fluctuations within the time of the simulations. In practice, to control population size, we tune the conversion factor α between food and bacteria. In addition, taking δx or δt not small enough leads to shifts in the food profile compared to the continuous case, as demonstrated in Fig. S15.

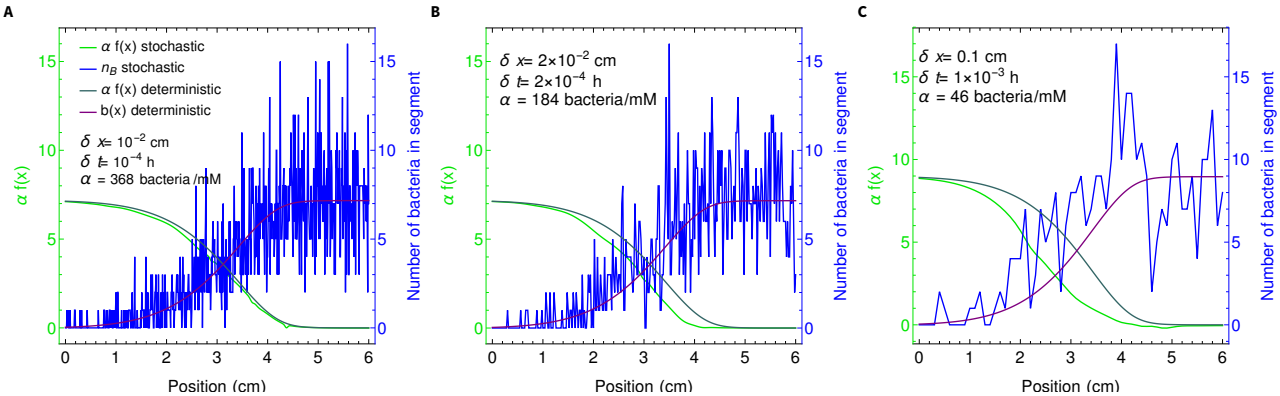

Figure S15: Dependence of the stochastic profiles on the discretization parameters, δx, δt, and N_{T} (via α). Numbers of bacteria and rescaled food quantities at time 10^3 h are shown in the stochastic and in the deterministic description versus position x in the gut. Initial conditions are the same in all panels and are taken as described in the "Initialization" item of our simulation steps. The only difference between panels is the value of the discretization parameters. Other parameters of the system are as in Fig. 1B, apart from $v = 0.514$ cm/h (and α which is specified in each panel).

In order to keep simulation times reasonably short, we chose the parameters corresponding to Fig. S15B. Note that this case exhibits a small shift in the spatial profile compared to the continuous case. However, this shift is stable, the steady state relation $B = \alpha(F_{in} - F)$ holds on average, and the source of the shift with respect to the continuous case is known to be the choice of δx and δt. The fluctuations caused by the choice of N_{T} (via α) in Fig. S15B are also stable and are not causing extinctions or changes of spatial regimes.

S12.2 Stochastic simulation results

Here, we present results obtained over 141,681 different stochastic replicates of the simulation described above, with discretization parameters corresponding to Fig. S15B.

Fig. S16 shows the fate of neutral mutants appearing at various locations in the gut. We observe a good agreement between panels A and B, obtained from our stochastic simulations, and panels C and D, obtained from our numerical resolution of the deterministic system (analogous to Fig. 2). In particular, the black curve in Fig. S16A, showing the proportion $p(x_{\mathrm{M}})$ of mutant bacteria that fix versus their introduction position x_{M}, matches the black curve in Fig. S16C, which shows the steady-state ratio M/B of mutant to wild-type bacteria concentrations. This corroborates the fact that the deterministic steady-state ratio M/B in the deterministic description behaves as the mutant fixation probability in the stochastic case, for each given mutant introduction position x_{M}.

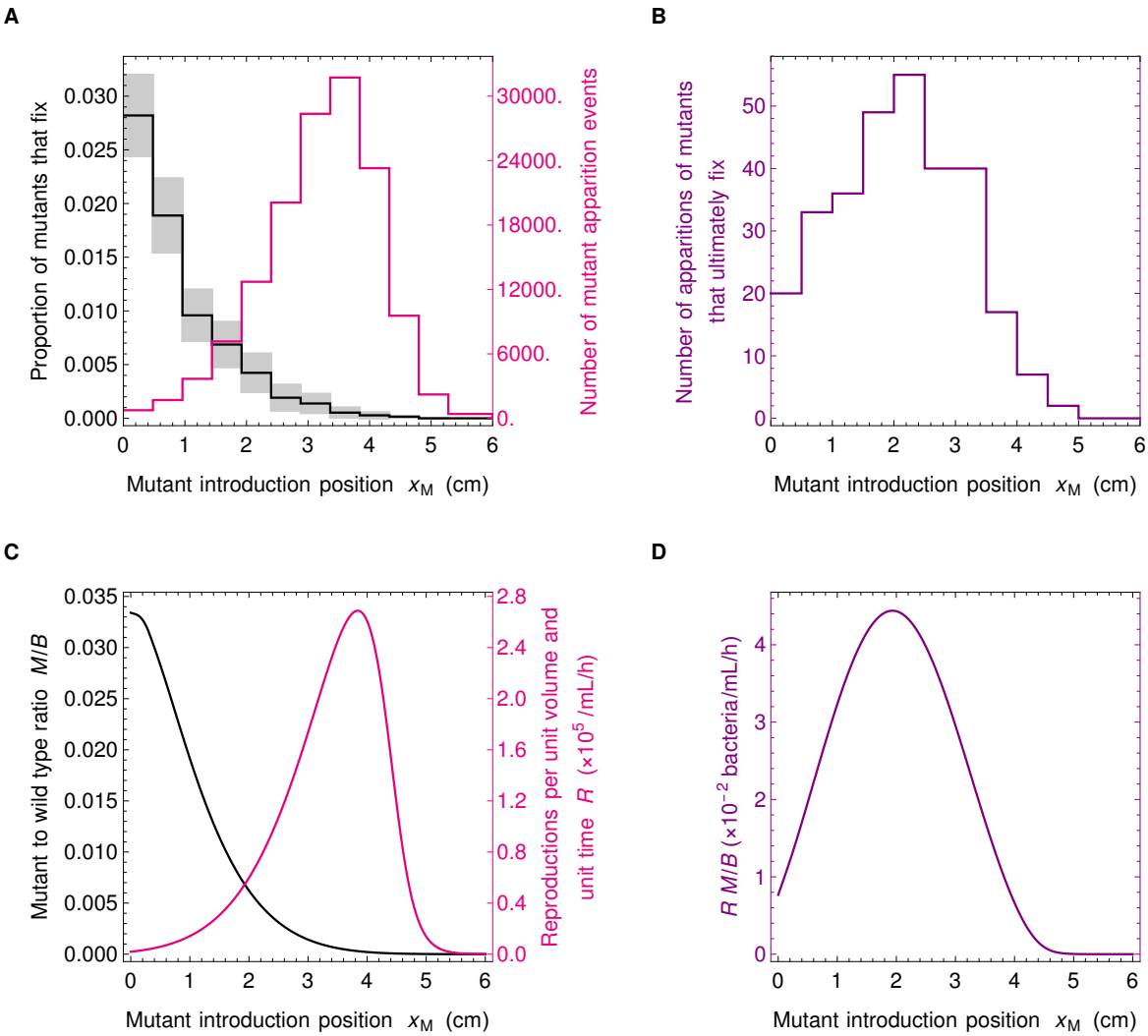

Figure S16: Fate of neutral mutants appearing at various locations in the gut: stochastic and deterministic descriptions. A: Black: proportion $p(x_{\rm M})$ of mutant bacteria that fix versus their introduction position $x_{\rm M}$, computed as $p(x_{\rm M}) = N_{MF}(x_{\rm M})/N_{MA}(x_{\rm M})$, where $N_{MF}(x_{\rm M})$ is the number of mutants that fix in the segment at $x_{\rm M}$, while $N_{MA}(x_{\rm M})$ is the number of mutants that appeared in that bin. Gray shaded boxes correspond to $p(x_{\rm M}) \pm \sqrt{N_{MF}(x_{\rm M})}/N_{MA}(x_{\rm M})$. Pink: histogram of the number of mutant apparition events at each introduction position $x_{\rm M}$. **B:** Histogram of the number of apparition events of mutant bacteria that eventually fix versus their introduction position $x_{\rm M}$. In panels A and B, parameter values are as in Fig. 1B, except $\alpha = 184$ bacteria/mM and $v = 0.514$ cm/h. Discretization parameters are $\delta x = 0.02$ cm, $\delta t = 2 \times 10^{-4}$ h. The bin size used is 0.5 cm. **C** and **D:** Corresponding numerical solutions of the deterministic system. In panels C and D, parameter values are as in Fig. 1B, except $v = 0.514$ cm/h. For direct comparison with the stochastic results shown in panels A and B, curves have been rescaled, the black one by $\alpha_d/(M_{0,d}\alpha_s)$, the pink one by α_s/α_d, and the purple one by $1/M_{0,d}$, where α_d is the value of α used in the deterministic case (see Fig. 1B), while α_s is the one employed in the stochastic case, and $M_{0,d}$ is the initial local mutant concentration at $x_{\rm M}$ in the deterministic description.

One of the main results of our study is that the fixation probability of a mutant is given by $1/N_{\rm A}$ instead of $1/N_{\rm T}$. Fig. S17 demonstrates that this result is validated by our numerical simulations.

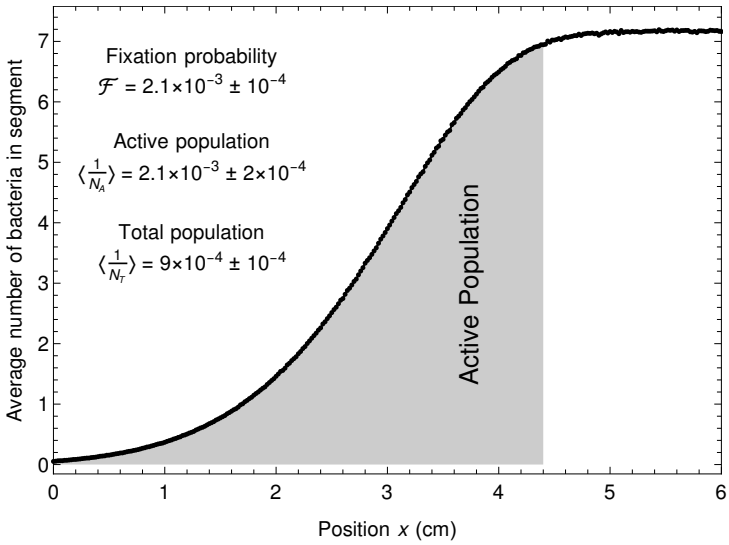

Figure S17: Fixation probability is equal to the inverse active population size. The curve shows the average profile of the bacterial population in the stochastic simulations run in panels A and B of Fig. S16. The total population size is proportional to the total area below this curve, while the active population size (defined in section S9, see also Fig. 4 of main text) is proportional to the area shaded in gray. The fixation probability $\mathcal{F}$ is computed from our stochastic simulation results as $\mathcal{F} = N_F/N_R$, where N_F is the number of fixation events and N_R the number of runs. The associated confidence interval is given by $(N_F \pm \sqrt{N_F})/N_R$. The inverses of the total and of the active population size are also given, averaged over realizations and over time. The associated confidence intervals are estimated via standard deviations.

Furthermore, Fig. S18A-B shows the distribution of mutant extinction times and mutant fixation times. Most extinction times are relatively short. The peak of the distribution of the position of mutant apparition (pink curve of panel A of Fig. S16) is about 2 cm from the downstream boundary, which takes about 4 h to cross at the velocity v employed here. Consistently, this is also the peak time for extinction events. Fixation times are much longer. Since the effective population size is the active population size, of the order of 470 and the replication rate is $r = 0.42$ h^{-1}, the fixation time is expected to be of the order of $N_\mathrm{A}/r \simeq 1000$ h, which is indeed the right order of magnitude. As bacterial populations in the gut are very large, fixation times will accordingly be very large. Fig. S18C shows the average proportion of mutants versus time. At the time of mutant apparition, it is the inverse of the total population size, but very quickly this mutant proportion becomes close to the inverse of the active population size. Therefore, the effect of active versus total population size matters much earlier than the fixation time.

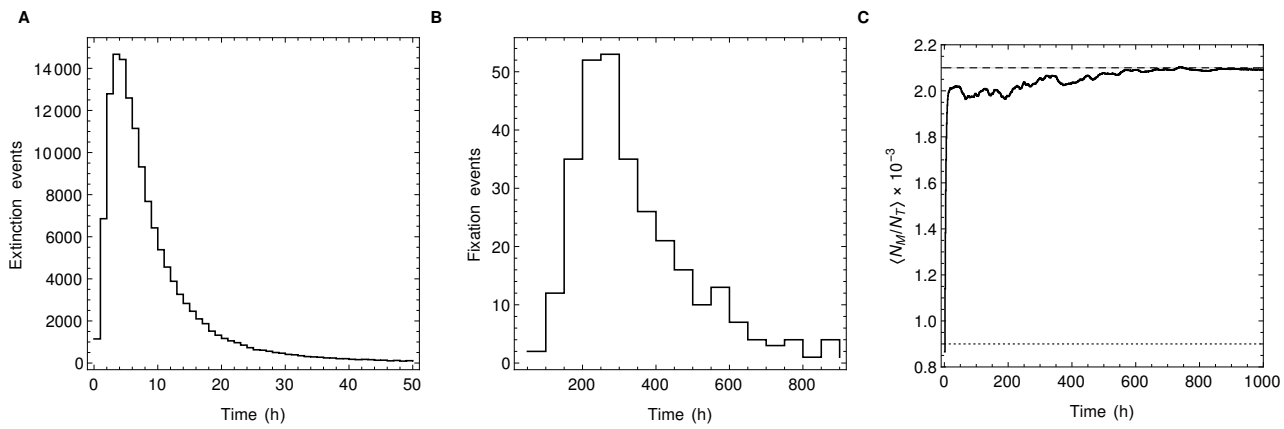

Figure S18: Time scales in stochastic simulations. A: Histogram of mutant extinction times. **B:** Histogram of mutant fixation times. **C:** Curve: average proportion of mutants versus time after mutant apparition. Dashed line: $\langle 1/N_\mathrm{A} \rangle$. Dotted line: $\langle 1/N_\mathrm{T} \rangle$. Upon mutant apparition, as there is one mutant, the mutant proportion is $1/N_\mathrm{T}$. In the long term, as the population will be either all wild-type or all mutant, the average mutant proportion tends to the fixation probability. But the time scale over which the mutant proportion becomes close to $1/N_\mathrm{A}$ is small compared to the time to fixation.