# Emergent evolutionary forces in spatial models of microbial growth in the human gut microbiota

Olivia M. Ghosh[1] and Benjamin H. Good[2]

[1] *Department of Physics, Stanford University, Stanford, CA 94305*
[2] *Department of Applied Physics, Stanford University, Stanford, CA 94305*

The genetic composition of the gut microbiota is constantly reshaped by ecological and evolutionary forces. These strain-level dynamics can be challenging to understand because they emerge from complex spatial growth processes that take place within a host. Here we introduce a general population genetic framework to predict how stochastic evolutionary forces emerge from simple models of microbial growth in spatially extended environments like the intestinal lumen. Our framework shows how fluid flow and longitudinal variation in growth rate combine to shape the frequencies of genetic variants in sequenced fecal samples, yielding analytical expressions for the effective generation times, selection coefficients, and rates of genetic drift. We find that the emergent evolutionary dynamics can often be captured by well-mixed models that lack explicit spatial structure, even when there is substantial spatial variation in species-level composition. By applying these results to the human colon, we find that continuous fluid flow and simple forms of wall growth are unlikely to create sufficient bottlenecks to allow large fluctuations in mutant frequencies within a host. We also find that the effective generation times may be significantly shorter than expected from traditional average growth rate estimates. Our results provide a starting point for quantifying genetic turnover in spatially extended settings like the gut microbiota, and may be relevant for other microbial ecosystems where unidirectional fluid flow plays an important role.

The human gut harbors a diverse microbial ecosystem that plays an important role in maintaining health and reducing susceptibility to disease (1, 2). The composition of the gut microbiota is shaped by interactions between hundreds of different species (3, 4), which must continually reproduce in order to balance their daily excretion in feces (5, 6). These high rates of cell turnover create enormous opportunities for rapid evolutionary change: the trillions of bacteria within a single human colon produce more than a billion new mutations every day (7), and tens of thousands of bacterial generations will typically elapse within a single human lifetime (1, 8). Consistent with these estimates, a growing number of empirical studies, ranging from experimental evolution in mouse models (9, 10, 11, 12) to longitudinal sequencing of human subjects (7, 13, 14, 15, 16, 17, 18, 19, 20), have shown that genetic variants can sweep through resident populations of gut bacteria on timescales of weeks and months. Understanding the causes and consequences of this genetic turnover is likely to be important for predicting the stability and resilience of this diverse microbial ecosystem (21).

A central challenge in understanding the evolution of the gut microbiota stems from the complex structure of the gut environment. While the frequencies of genetic variants are typically tracked in homogenized fecal samples, most microbial growth occurs in the spatially extended setting of the large intestine. At present, little is known about how this spatial structure influences the evolutionary dynamics that are observed in sequenced fecal samples. This makes it difficult to draw quantitative conclusions from existing genetic data: What does it mean if the frequency of a variant increases dramatically over a few days? or if it eventually declines in frequency a few weeks later? Is it reasonable to interpret these coarse-grained measurements using well-mixed models from population genetics? or are spatially explicit models necessary for extracting quantitative insights?

Previous studies of gut microbial biogeography have shown that growth conditions can vary over a huge range of length scales, ranging from the micron scales of intercellular interactions (22, 23, 24) to macroscopic compartments like the mucosa, lumen, and intestinal crypts (25). Even within the comparatively simpler environment of the central lumen — thought to be the primary site of growth for some of the most abundant species of *Bacteroides* and *Firmicutes* (26) — significant longitudinal gradients are established by the basic anatomy of the intestinal tract. Primary nutrients are supplied at one end of the colon, where microbial densities are relatively low, and are rapidly consumed by bacteria in the ascending colon (27). The constant influx of fluid, along with additional mixing from contractions of the intestinal walls (26, 28), causes both bacteria and their metabolic byproducts to drift toward lower regions of the colon, where they are eventually excreted in feces. These physical processes combine to produce large gradients in nutrient concentrations, pH, and excretion rates across the longitudinal axis of the large intestine (27, 28, 29), which can have important consequences for the relative growth rates and abundances of different species (28, 30, 31, 32, 33, 34).

The evolutionary implications of these spatial gradients are less well understood, since the details will depend

on the phenotypic differences that exist among more closely related strains. However, even in the absence of phenotypic differences, the longitudinal structure of the lumen will lead to general physical effects that can be appreciated by revisiting our back-of-the-envelope calculation for the production rate of new mutations. While billions of mutations may be produced each day, not all of these emerge on equal footing. Mutations that occur near the lower end of the colon are very likely to be excreted in feces before they can produce viable offspring. Conversely, mutations that occur near the entrance of the colon are significantly more likely to produce descendants that can persist and grow within the host over many generations.

These physical considerations suggest that only a subset of the luminal population will tend to contribute to its long-term genetic composition. Similar arguments apply for estimates of the effective generation time, which can vary substantially in different regions of the colon (27, 35, 36). These arguments suggest that the relevant evolutionary parameters must emerge from some sort of average over this existing longitudinal structure. However, the weights in this average will be determined not by the current bacterial densities and growth rates, but rather by their long-term contribution to the future population. Our limited understanding of these weights — even in the simplest models of luminal growth — leaves many important questions unresolved: Can fluid flow alone generate sufficient population bottlenecks that would allow large fluctuations in mutant frequencies within a host? How do absolute growth rate differences map on to the frequency trajectories that can be observed in vivo? And how do the answers to these questions depend on the physical properties of the gut environment, which vary dramatically across species, and can be modulated by drugs or other dietary perturbations?

In this study, we introduce a mathematical framework for predicting how stochastic evolutionary forces emerge from explicit spatial models of microbial growth in the intestinal lumen. Building on earlier work in Refs. (37, 38, 39, 40), we show that the emergent evolutionary dynamics can often be captured by spatially homogeneous models with an appropriate set of effective parameters, even when the ecological effects of spatial structure are very strong. We derive analytical expressions to show how the effective selection coefficients, generation times, and population sizes are related to the underlying growth rates and flow profiles along the large intestine. By applying this framework to parameter estimates from the human colon (28), we find that fluid flow alone is unlikely to generate large fluctuations in the frequencies of sequenced mutations, while the diffusion of slowly growing cells can still have a large impact on the long-term evolution of the population. These results provide a framework for understanding how spatial structure influences the genetic composition of the gut microbiota, and may be relevant for other microbial ecosystems (e.g. hot springs and acid mine drainage systems) where unidirectional flow is thought to play an important role (41, 42).

## ANALYSIS

We consider the dynamics of mutant lineages in a spatial model of luminal growth that incorporates continuous fluid flow, growth, and active mixing from contractions of the intestinal walls (26, 28). A key assumption of this model is that the mixing process is strong enough to destroy any cross-sectional spatial structure within the lumen (28), yielding an effective diffusion process in the longitudinal dimension (denoted by $x$) with some effective flow rate $v(x)$ and effective diffusion coefficient $D(x)$ (Fig. 1A). The local density of species $\mu$ at position $x$ and time $t$ can then be described by a reaction diffusion equation,

$$\begin{aligned}\partial_t \rho_\mu(x,t) = \partial_x^2[D(x)\rho_\mu(x,t)] - \partial_x[v(x)\rho_\mu(x,t)] \\ + \lambda_\mu[\vec{r}(x,t)]\rho_\mu(x,t)\,,\end{aligned} \tag{1}$$

where $\lambda_\mu[\vec{r}(x,t)]$ is a function that describes the net growth rate of species $\mu$. We assume that this net growth rate depends on a set of local environmental variables $\vec{r}(x,t)$, which could represent local nutrient concentrations, pH, or the abundances of other species. These environmental variables will be described by their own set of reaction diffusion equations, which are coupled to the species densities $\rho_\mu(x,t)$. In our analysis below, we will assume that the ecosystem has approached a steady-state, in which the species densities $\rho_\mu(x,t)$ and environmental variables $\vec{r}(x,t)$ can be described by time invariant profiles $\rho_\mu(x)$ and $\vec{r}(x)$. This is a crucial simplification, as it means that the growth rate $\lambda_\mu[\vec{r}(x,t)]$ can be replaced by a simple spatially varying growth function $\lambda_\mu(x) \equiv \lambda_\mu\,[\vec{r}(x)]$, which is conditionally independent of the surrounding community. The steady-state growth profiles are uniquely determined by the steady-state density profiles,

$$\lambda_\mu(x) = \frac{\partial_x^2[D(x)\rho_\mu(x)] - \partial_x[v(x)\rho_\mu(x)]}{\rho_\mu(x)}\,, \tag{2}$$

such that the realized growth rates can in principle be determined from static snapshots of a community's spatial composition.

### Dynamics of mutant lineages

We wish to understand the dynamics of genetic variants within species in the steady-state communities described above. Since these variants may initially be present in just a small number of cells, the deterministic dynamics

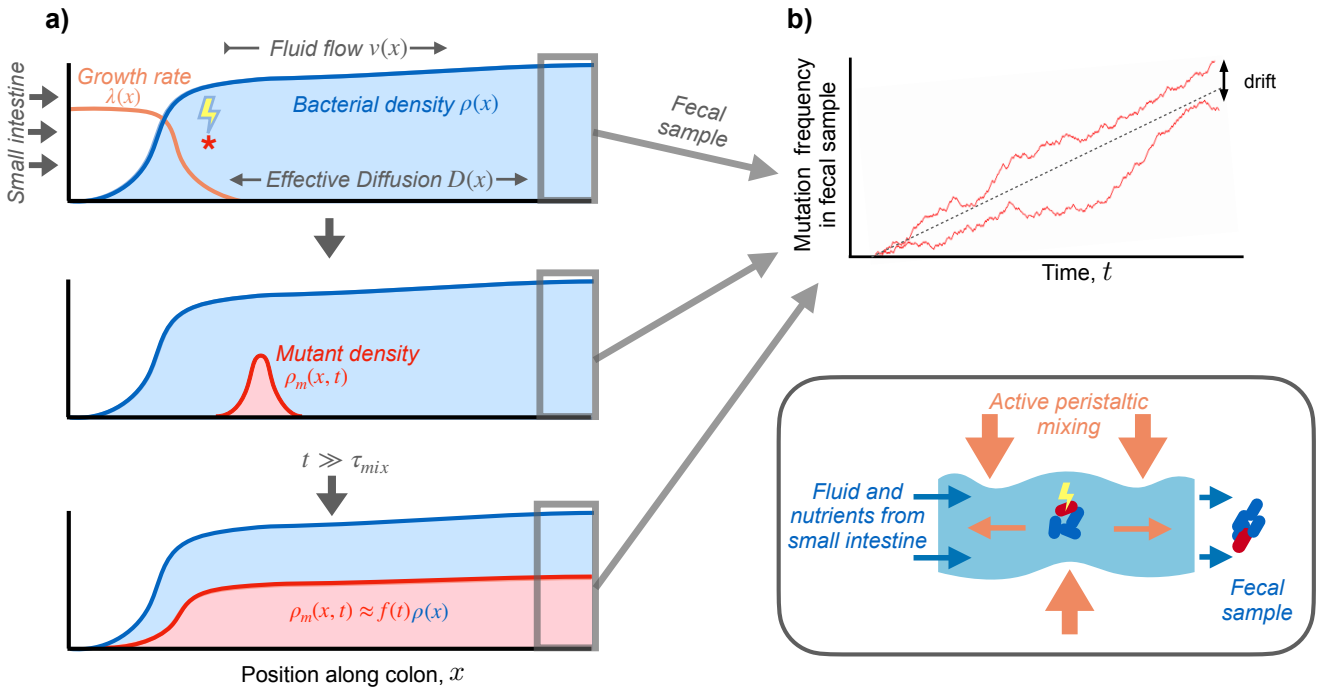


FIG. 1 **Dynamics of a genetic variant in a spatially extended population in the intestinal lumen.** (a) Schematic of spatial growth model. Nutrients enter the colon from the small intestine (left) and fecal samples are generated at the distal end (right). Intestinal wall contractions mix contents of lumen (inset), producing effective diffusion in the axial direction (26, 28). Continuous fluid flow through the colon generates longitudinal gradients in steady-state bacterial growth rates [$\lambda(x)$, orange line] and population densities [$\rho(x)$, shaded blue region]. Mutation events (red star) found new lineages that begin to grow and spread out in space [$\rho_m(x,t)$, shaded red region]. After $\tau_{\text{mix}}$ generations, the spatial distribution of the mutant lineage approaches that of the wildtype population, but with a smaller overall size. (b) Mutation frequencies observed in fecal samples over time. Two hypothetical trajectories are shown (red lines), illustrating the stochastic fluctuations that arise due to genetic drift.

in Eq. (1) must be augmented to account for stochastic fluctuations in cell numbers in both space and time. As long as the frequency of the mutant lineage is sufficiently rare, we expect it to have a negligible influence on the steady state values of $r_i(x)$ and $\rho_\mu(x)$, such that the growth function of the mutant can be described its own steady state growth profile $\lambda_m(x)$. In Supplementary Information 1, we use a microscopic birth-and-death model to show that the density profile of the mutant lineage can be described by the coarse-grained stochastic differential equation

$$\begin{aligned}\partial_t \rho_m(x,t) = \partial_x^2[D(x)\rho_m(x,t)] - \partial_x[v(x)\rho_m(x,t)] \\ + \lambda_m(x)\rho_m(x,t) \\ + \sqrt{\Upsilon_m(x)\rho_m(x,t)}\cdot\eta_\lambda(x,t) \\ + \frac{\partial}{\partial x}\left[\sqrt{D(x)\rho_m(x,t)}\cdot\eta_D(x,t)\right],\end{aligned} \tag{3}$$

where $\eta_\lambda(x,t)$ and $\eta_D(x,t)$ are independent Brownian noise terms with $\langle\eta_i(x,t)\eta_j(x',t')\rangle = \delta_{ij}\delta(x-x')\delta(t-t')$ (43), and $\Upsilon_m(x)$ is the variance in offspring number of the local birth and death process. In the absence of cell death, $\Upsilon_m(x)$ will be proportional to the local growth rate $\lambda_m(x)$; we expect this to be a reasonable approximation in the human gut, where live cells comprise a large fraction of the total excreted biomass (5, 44, 45).

Equation (3) is a stochastic variant of the familiar Fisher-KPP equation (46, 47), which is frequently used to model the spread of species or genetic variants through space (46, 47, 48, 49). An important difference in our setting is that the expansion velocity $v(x)$ is now set by the fluid flow in the environment, such that the mutant and wildtype populations are continuously "treadmilling" in place. We will show below that this typically leads to "pushed" rather than "pulled" expansions (50, 51), which will have important implications for the stochastic dynamics of genetic drift in this system (52, 53).

Equation (3) provides a quantitative model for predicting the spread of a genetic variant within a resident gut community under the combined action of fluid flow, mixing, and spatially varying growth rates. While the spatiotemporal dynamics of Eq. (3) are often quite complicated, further simplifications arise in scenarios where the emergent evolutionary forces (e.g. natural selection and genetic drift) are much slower than the "ecological" processes of growth, flow, and mixing. These ecological forces will lead to a characteristic *mixing timescale* $\tau_{\text{mix}}$ (discussed in more detail below), which represents the

time required for the descendants of a mutant lineage to disperse across space (Fig. 1A). For times much larger than $\tau_{\text{mix}}$, the shape of the mutant density profile will approach the steady-state population profile,

$$\rho_m(x,t) \approx f(t)\rho(x) \quad (t \gg \tau_{\text{mix}}), \tag{4}$$

but with a prefactor $f(t)$ that depends on the growth rate $\lambda_m(x)$, the stochastic noise terms in Eq. (3), and the initial density profile $\rho_m(x,0)$. This prefactor $f(t)$ also coincides with the relative frequency of the mutant lineage in later fecal samples (Fig. 1B). Thus, the dynamics of $f(t)$ provide a crucial connection between the spatial growth dynamics in Eq. (3) and the frequencies of mutant lineages that are measured in metagenomic sequencing data.

Much of the relevant behavior can be understood by considering a perfectly neutral lineage $[\lambda_m(x) = \lambda(x)]$ founded by a single individual a position $x$ (Fig. 1A). In Supplementary Information 2, we show that the distribution of $f(t)$ at intermediate times ($t \gg \tau_{\text{mix}}$) can be captured by a well-mixed Wright-Fisher model

$$\frac{\partial f}{\partial t} = \sqrt{\Lambda f(t)}\,\eta(t), \tag{5}$$

with an effective initial frequency

$$f(0|x) \approx u(x) \equiv \frac{D(x)\rho(x)e^{-\int_0^x \frac{v(x')}{D(x')}dx'}}{\int D(x)\rho(x)^2 e^{-\int_0^x \frac{v(x')}{D(x')}dx'}}, \tag{6}$$

and an effective rate of genetic drift

$$\Lambda = \int \left[\Upsilon(x)u(x)^2 + 2D(x)\left(\partial_x u(x)\right)^2\right]\rho(x)\,dx, \tag{7}$$

which is independent of the mutation's initial position. Equation (5) implies that this new mutation will require a time of order $\tau_{\text{drift}} \sim f/\Lambda$ to drift to frequency $\sim f$ in future fecal samples (54, 55). These dynamics will be consistent with the separation-of-timescales assumption as long as $\tau_{\text{drift}} \gg \tau_{\text{mix}}$, which requires that $f(t) \gg \Lambda\tau_{\text{mix}}$. Since $\Lambda$ is inversely proportional to the bacterial density, and $\tau_{\text{mix}}$ is independent of density (see below), we expect these results to apply for a broad range of frequencies when population sizes are sufficiently large.

At intermediate times ($\tau_{\text{mix}} \ll t \ll 1/\Lambda$), the descendants of the mutant lineage will be approximately evenly distributed across the population (Fig. 1A), so the conditional fixation probability of the entire lineage is simply $f(t)$. Averaging over the random values of $f(t)$ using Eq. (5), we find that

$$p_{\text{fix}}(x) = \int f \cdot p(f,t|x)df = u(x), \tag{8}$$

so that Eq. (6) has a simple interpretation as the fixation probability of a neutral mutation at position $x$ (37). Similarly, the product $g(x) = u(x)\rho(x)$ represents the spatial distribution of future common ancestors of the population (37) — regardless of whether they have a mutation or not — which provides a quantitative measure of the spatial locations that have an outsized influence on the genetic composition of future fecal samples. By modelling the long-term accumulation of neutral mutations within these populations, we obtain a simple prediction for the effective generation time,

$$\frac{1}{\tau_g} = \lambda_e \equiv \int [\lambda(x) + d(x)]u(x)\rho(x)\,dx, \tag{9}$$

where $d(x)$ is the local death rate (Supplementary Information 2.3). This allows us to re-express the "wall-clock" rate of genetic drift in Eq. (7) as a more traditional effective population size:

$$N_e = \frac{\int [\lambda(x) + d(x)]u(x)\rho(x)\,dx}{\int \left[\Upsilon(x)u(x)^2 + 2D(x)\left(\partial_x u(x)\right)^2\right]\rho(x)\,dx}, \tag{10}$$

which is proportional to the total bacterial density as expected. When cell death is negligible $[\lambda(x)+d(x) \approx \lambda(x)$, $\Upsilon(x) \approx \lambda(x)]$, these effective parameters are completely determined by the steady state density profile $\rho(x)$.

#### Analytical solutions for a minimal model of bacterial growth

We can understand the implications of these results by examining a minimal model of bacterial growth in the gut, which yields analytical solutions for the growth and density profiles in Eqs. (9) and (10). We assume that flow and diffusion are uniform in space, such that there is a characteristic length scale $\sim D/v$ that sets the transition between diffusive and advective motion. We further assume that growth depends on a single limiting nutrient with a step-like dose response curve (Fig. 2A). These assumptions produce a spatial growth profile with a sharp transition at a critical position $x = \ell$ (Fig. 2B), whose location depends on the relative magnitudes of $\lambda$, $v$, and $D$. In Supplementary Information 4, we obtain analytical expressions for the resulting population density profiles, which are illustrated in Fig. 2C,D.

We observe two characteristic regimes that depend on the value of the dimensionless parameter $\alpha \equiv 4D\lambda/v^2$, which compares the characteristic timescales of growth ($\lambda^{-1}$) and local dilution ($\sim D/v^2$). When $\alpha < 1$, growth is not sufficient to overcome the constant loss of individuals due to fluid flow, and the population cannot permanently establish itself in the gut — a scenario we refer to as "washout" (26). At the opposite extreme ($\alpha \gg 1$), growth is sufficiently rapid that the nutrients are depleted long before flow is relevant ($\ell \ll D/v$), and diffusive mixing is still the dominant form of motion. This diffusion-dominated growth produces a nearly uniform population density profile, and a distribution of future ancestors $[g(x) = u(x)\rho(x)]$ that is primarily composed

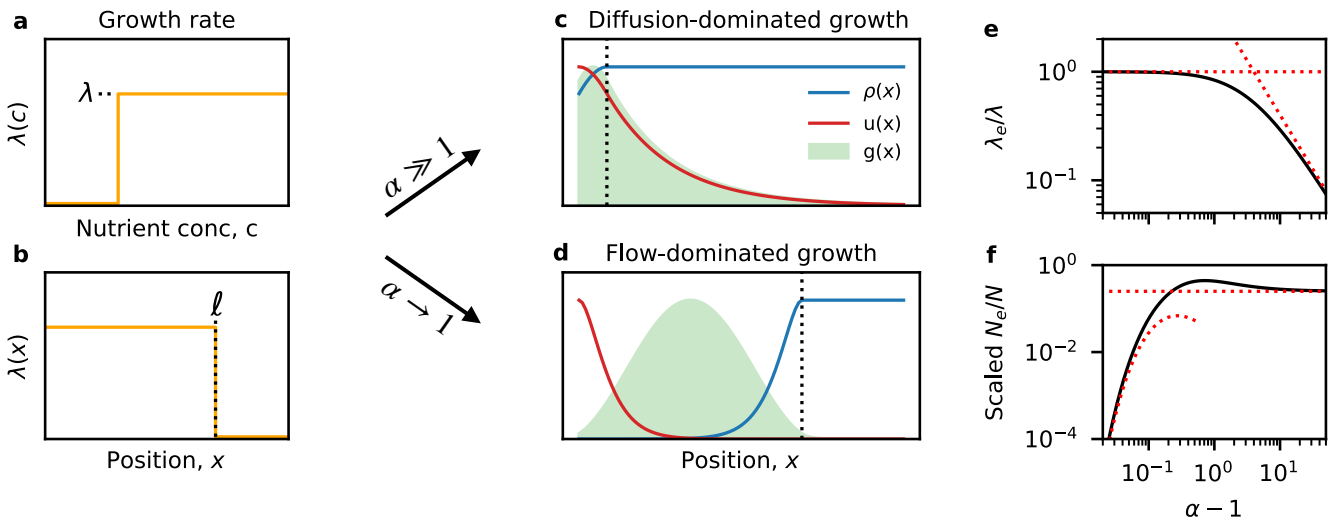


FIG. 2 **Analytical solutions for a minimal model of bacterial growth.** (a,b) Schematic of bacterial growth rates. A step-like growth response curve (a) leads to a sharp transition between high growth ($\lambda(x) \approx \lambda$) and no-growth ($\lambda(x) \approx 0$) at a critical position $\ell$ (b). (c, d) Spatial profiles of population density, $\rho(x)$ (blue line), neutral fixation probability, $u(x)$ (red line), and distribution of future common ancestors, $g(x) = u(x)\rho(x)$ (green shaded region), for examples of (c) diffusion-dominated growth ($4D\lambda/v^2 \approx 10$) and (d) flow-dominated growth ($4D\lambda/v^2 \approx 1.04$). In both cases, the dashed black lines indicate the boundary of the growth region in panel b. (e) Scaled effective growth rate, $\lambda_e/\lambda$, and (f) scaled effective population size, $(N_e/N) \cdot (vL/2D)$, as a function of the washout parameter $\alpha \equiv 4D\lambda/v^2$. Black lines denote the analytical predictions in Supplementary Information 4.2, while the red lines illustrate the asymptotic behavior in the $\alpha \gg 1$ and $\alpha - 1 \ll 1$ limits.

of slowly growing cells (Fig. 2C). This leads to a reduced effective growth rate, $\lambda_e \sim v^2/D \ll \lambda$ (Fig. 2E), and an effective population size, $N_e \sim \rho(L)D/v$, that is independent of the length of the colon (Fig. 2F).

Between these two extremes, we also observe an intermediate regime close to the washout threshold ($\alpha \to 1$), where flow dominates over diffusion throughout most of the growth region ($D/v \ll \ell$). This produces a strong spatial gradient in the density profile, and a distribution of future ancestors [$g(x) = u(x)\rho(x)$] that is primarily concentrated within the growth region (Fig. 2D). The effective growth rate therefore approaches its maximum value $\lambda_e \approx \lambda$ (Fig. 2E), while the effective population size is dramatically reduced (Fig. 2F). By examining the contributions to Eq. (10), we find that these lower effective population sizes are driven by fluctuations in a narrow region near the entrance of the colon ($x \lesssim D/v$), rather than the comparatively larger sub-population of dividing cells ($x \sim \ell$). This shows how long-term lineage dynamics can have a major impact on the emergent evolutionary parameters in our model.

While the enhanced fluctuations are most pronounced near the washout threshold ($\alpha \approx 1$), we note that washout itself does not play a major role in this behavior. In Supplementary Information 4.3, we show that the results in Fig. 2 are robust to the addition of a washed out region at the proximal end of an otherwise established population. This suggests that it is the accumulated growth over many diffusion lengths — rather than washout per se — that is the primary cause of the reduced effective population size when $\alpha \to 1$. It also suggests that our qualitative results are do not depend on the specific nature of the boundary between the large and small intestine: the small number of cells that diffuse into this low-growth region do not dramatically change the dynamics that occur downstream.

#### Dynamics of selected mutations

We can use a similar separation-of-timescales approach to analyze the fates of selected mutations as well (Supplementary Information 3). Provided that the growth rate differences are sufficiently small (see below), we find that the emergent dynamics on timescales $t \gg \tau_{\text{mix}}$ can again be mapped on to well-mixed Wright-Fisher model

$$\frac{\partial f}{\partial t} = s_e \lambda_e f(t) + \sqrt{\frac{\lambda_e f(t)}{N_e}} \cdot \eta(t)\,, \tag{11}$$

with an effective selection coefficient

$$s_e = \frac{1}{\lambda_e} \int [\lambda_m(x) - \lambda(x)] u(x)\rho(x)\, dx\,. \tag{12}$$

This equation encapsulates the relevant growth rate differences across the ancestral region $g(x) \propto u(x)\rho(x)$. For a global increase in growth rate [$\lambda_m(x) = (1+s)\lambda(x)$], the effective selection coefficient $s_e$ is proportional to $s$ as expected. However, Eq. (12) also allows us to account

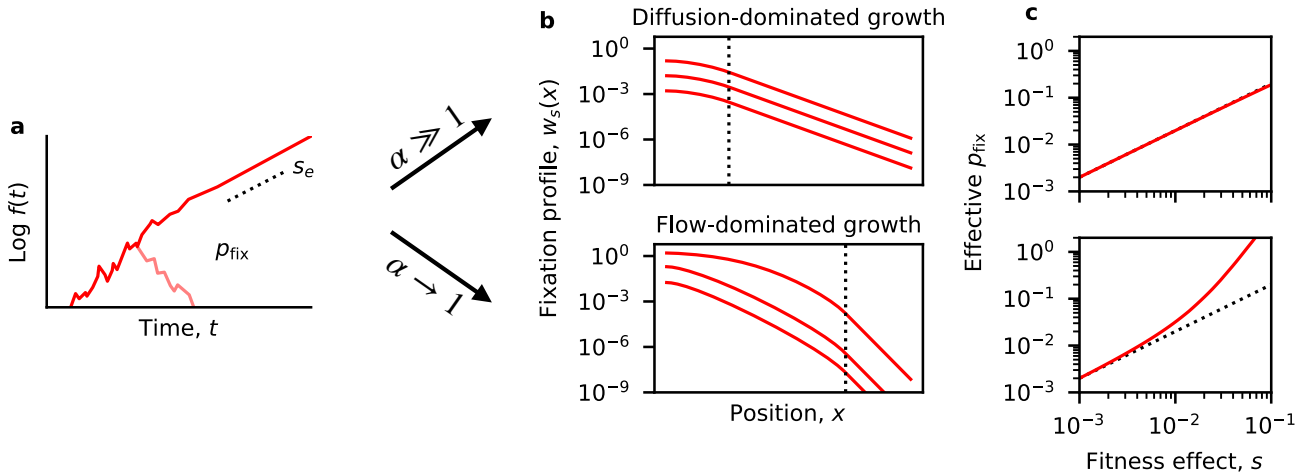


FIG. 3 **Dynamics of selected mutations.** (a) Schematic trajectory of a beneficial mutation, which either establishes and grows exponentially at rate $s_e$ (dashed line) or else goes extinct (light red line). (b) The spatially-resolved fixation probability $w_s(x)$ of a beneficial mutation that originates at position $x$. Red lines denote the numerical solutions of Eq. (13) for the minimal model in Fig. 2 with a constant growth rate advantage $\lambda_m(x) = (1+s)\lambda(x)$ (Supplementary Information 4.4). The dashed lines indicate the transition from $\lambda(x) \approx \lambda$ to $\lambda(x) \approx 0$. The top panel illustrates a population in the diffusion-dominated growth regime ($4D\lambda/v^2 \approx 1.9$) while the bottom panel illustrates a population in the flow-dominated growth regime ($4D\lambda/v^2 \approx 1.04$). In both cases, fixation profiles were calculated for $s = 10^{-3}$, $10^{-2}$, and $10^{-1}$. (c) The effective fixation probability $p_{\rm fix}$ obtained by averaging $w_s(x)$ the range of initial positions. The top and bottom panels denote the same populations as in (b). Dashed lines indicate the separation-of-timescales prediction, $p_{\rm fix} = 2s_e$. Deviations from this prediction emerge in the flow-dominated growth regime, due to the enhanced fixation probability of mutations that originate at larger $x$ values (illustrated in panel b).

for tradeoffs in growth rate that emerge at different spatial locations. The relative weights of fitness effects in these different microenvironments are determined by the product $\lambda(x)g(x)$, which biases the effective selection coefficients toward regions with higher local growth rates.

The mapping in Eq. (11) allows us to obtain a precise picture of the dynamics a beneficial mutation as observed in sequenced fecal samples. When $\tau_{\rm mix} \ll t \ll 1/s_e\lambda_e$, the dynamics of the selected lineage will be identical to the neutral mutations described above. With probability $\sim N_e s_e u(x)$, the mutant lineage will eventually establish and grow deterministically as $\sim e^{s_e\lambda_e t}$, and will fix within the population on a characteristic timescale $T_{\rm fix} \sim \frac{1}{s_e\lambda_e}\log(2N_e s_e)$ (54, 55). These dynamics will be consistent with the separation-of-timescales assumption as long as $s_e\lambda_e\tau_{\rm mix} \ll 1$.

For stronger selection strengths ($s_e\lambda_e\tau_{\rm mix} \gg 1$), the separation-of-timescales assumption will eventually break down, and growth rate differences will start to play an important role during the mixing process itself. In this case, we show in Supplementary Information 3.1 that an analogous fixation probability can still be obtained numerically by solving the ordinary differential equation

$$0 = \lambda_m(x)w(x) + D(x)\partial_x^2 w(x) + v(x)\partial_x w(x) - \frac{\Upsilon_m(x)w(x)^2}{2} - D(x)(\partial_x w(x))^2\,, \tag{13}$$

which gives the fixation probability of a beneficial mutation that arises at a given position $x$. Equation (13) is similar to existing branching process models of fixation in spatially or genetically subdivided populations (56, 57, 58), but with an additional non-linearity that accounts for the conservative noise of spatial diffusion. The total fixation probability can be obtained from Eq. (13) by averaging $w(x)$ over the local production rate of new mutations, $\mu(x)\rho(x)$, yielding

$$p_{\rm fix} \equiv \frac{1}{N_e}\frac{\int \mu(x)\rho(x)w(x)\,dx}{\int \mu(x)\rho(x)u(x)\,dx}\,. \tag{14}$$

When $s_e\lambda_e\tau_{\rm mix} \ll 1$, we recover the linear relation $p_{\rm fix} = 2s_e$ predicted by the separation-of-timescales approximation. Deviations from this behavior at larger values of $s_e$ can therefore provide insight into the mixing timescale $\tau_{\rm mix}$, which we have so far treated as an implicit parameter.

In Fig. 3, we compute these fixation probabilities numerically for the minimal model in Fig. 2, assuming a global increase in growth rate [$\lambda_m(x) = (1+s)\lambda(x)$] and a replication-dominated mutation rate [$\mu(x) \approx \mu\lambda(x)$]. We find good agreement with the separation-of-timescales prediction at sufficiently low values of $s$. This shows that $\tau_{\rm mix}$ is primarily determined by the ancestral region [$g(x) = u(x)\rho(x)$] rather than the total length of the colon ($L \to \infty$). For diffusion-dominated growth ($4D\lambda/v^2 \gtrsim 1$), this linear approximation remains accurate for growth rate advantages as large as $s{\sim}100\%$,

demonstrating the broad utility of the separation-of-timescales approach in this regime. In contrast, we find that deviations can emerge at much lower values of $s$ in the flow-dominated growth regime ($4D\lambda/v^2 - 1 \ll 1$), such that the fixation probabilities are often much larger than expected under the separation-of-timescales approximation (Fig. 3C). The spatial fixation profiles in Fig. 3B show that this increase is primarily driven by enhanced fixation probabilities in lower regions of the colon (Fig. 3C), which acts to expand the range of production of successful beneficial mutations. These deviations can be rationalized by the longer times required for lineages to migrate across the ancestral region $\sim v/\ell$, such that even small growth rate advantages can produce large numbers of additional offspring before complete mixing is achieved. However, despite these enhanced fixation probabilities, we find that the long-term growth rates of established mutations continue to match the separation-of-timescales prediction above. By solving a related equation for the average density profile (Supplementary Information 4.5), we find that the exponential growth rates eventually approach $S \sim s_e\lambda_e$ even when $s_e\lambda_e\tau_{\text{mix}} \gg 1$. This shows that finite mixing can have a differential impact depending on the timescales of the corresponding evolutionary processes (e.g. establishment vs fixation).

## APPLICATION TO THE HUMAN COLON

Our analytical results have shown that the effective evolutionary parameters can strongly depend on the relative magnitudes of growth, flow, and mixing across the large intestine. How do these results extend to bacteria growing in the human colon? While direct measurements of the growth rates and density profiles are difficult to obtain in this system, rough estimates for many of the key parameters were recently obtained by Cremer *et al* (28) using a generalization of the model in Eq. (1). By explicitly modeling the consumption and production metabolites along the length of the human colon ($L{\sim}2m$), Cremer *et al* (28) obtained predictions for the steady-state growth rates and density profiles for a pair of representative species from the *Bacteroides* and *Firmicutes* phyla (Fig. 4A).

An important feature of this model is a spatially varying flow rate, due to the absorption of water by the colonic epithelium. Cremer *et al* (28) estimated that this effect causes the flow rate to decline from $v \approx 30\mu m/s$ to $v \approx 5\mu m/s$ over the length of the proximal colon ($\approx 30cm$). Given an estimated mixing rate of $D{\approx}10^6\mu m^2/s$, the effective diffusion lengthscales range from $D/v \approx 3 - 20cm$, and the corresponding timescales range from $D/v^2{\approx}16$ minutes to 11 hours. These estimates suggest that the *in vitro* growth rates of the representative strains ($\lambda = 0.6 - 1$ hr$^{-1}$) will be close to washout at the highest flow rates, but far from washout elsewhere, making it difficult to directly apply the results from our minimal model in Figs 2 and 3.

To overcome this challenge, we developed a numerical approach for calculating the effective evolutionary parameters in Eqs. (9), (10), and (14) using the steady-state density profiles estimated by Cremer *et al* (28) (Supplementary Information 5). The results are shown in Fig. 4. In both cases, we find that the flow rate decreases sufficiently rapidly across the proximal colon that the corresponding ancestral regions [$g(x) = u(x)\rho(x)$] include a substantial contribution from slowly growing cells. For example, the *B. theta* curves in Fig. 4E predict that about a third of the future common ancestors will come from regions where growth rates are <0.5/day, while another third will come from nutrient-rich regions where growth rates are >0.5/hr. This suggests that these resident populations will lie on the boundary between diffusion-dominated and flow-dominated growth in Figs 2 and 3 above. This has several implications for the emergent evolutionary dynamics:

Figure 4F shows that the effective growth rates are about 3 times smaller than the maximum growth rates attained at the entrance of the colon. This corresponds to about 9 generations per day for the *B. theta* example and about 5 generations per day for *E. rectale*. These predictions are 5-10 times higher than the traditional ∼1/day estimates based on the daily biomass excreted in feces (59, 60), which reflects the higher-than-average growth rates experienced by evolutionarily relevant lineages. Interestingly, these higher growth rates are more consistent with direct measurements of *B. theta* generation times in germ-free mice (61), suggesting that some of the emergent evolutionary parameters may be similar in these environments, despite the vast differences in host physiology.

In addition to these shortened generation times, Fig. 4G shows that corresponding effective population sizes are about 10-20 times smaller than the total number of bacterial cells in the colon. This leads to elevated rates of genetic drift, but not at the levels necessary to drive large changes in the frequencies of genetic variants over a single host lifetime. Given the large number of bacteria in the human gut [$N \sim 10^{13}$ (62)], these estimates predict that a species with 1% relative abundance would have an absolute rate of genetic drift of about $\Lambda{\equiv}\lambda_e/N_e{\sim}10^{-9}$ per day. At this rate, genetic drift would require more than a hundred years to generate even a 1% shift in the frequency of a common genetic variant. This stands in contrast to the rapid shifts in frequency that are observed for some mutations in sequenced human fecal samples (Fig. 5). This quantitative discrepancy suggests that other factors, like natural selection or temporal bottlenecks, are required to explain these observations.

Finally, Fig. 4H shows that the mixing timescales in the human colon are sufficiently rapid that the establishment probabilities of beneficial mutations are well-

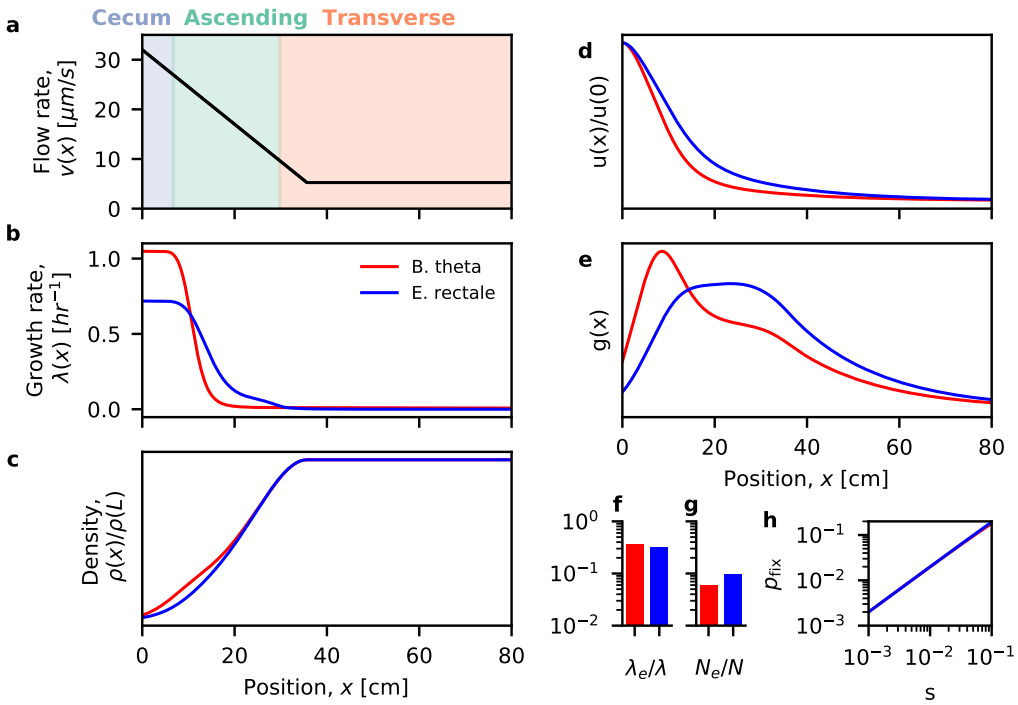


FIG. 4 **Applications to bacteria in the human colon.** (a) Estimates of the flow velocity along the first 80 cm of the human colon (28). Colors indicate the positions of different colonic segments. (b,c) Steady-state growth rate and density profiles estimated for two representative species using the results of Ref. (28). Growth rate profiles were extracted from the figures in Ref. (28), and the corresponding density profiles were calculated from Eq. (2) using the methods described in Supplementary Information 5. (d,e) Predictions for the neutral fixation profile $u(x)$ (d) and distribution of common ancestors $g(x)$ (e) obtained from the estimated parameters in (a-c) (Supplementary Information 5). In both cases, the predicted ancestral distributions contain a large contribution from regions with low growth rates. (f-g) Predictions for the effective growth rate, effective population size, and the net fixation probability of beneficial mutations (Supplementary Information 5).

approximated by the linear prediction [$p_{\text{fix}} \sim 2s_e$] across a wide-range of physiologically relevant fitness benefits. These results show that our separation-of-timescales approach can apply even for human-relevant conditions, yielding quantitative estimates for a number of key evolutionary parameters.

#### Extensions to wall growth and other spatial refugia

So far, we have assumed that the emergent evolutionary forces are dominated by the dynamics within the luminal fluid. However, our basic mathematical framework can also be extended to incorporate certain forms of wall growth (or other spatial refugia), which are known to play an important role in a variety of microbial ecosystems (23, 25, 42, 63, 64). While a complete analysis of this case is beyond the scope of the present paper, many of the key features can already be appreciated within a simple model, in which a well-mixed luminal compartment is connected to a large number of protected subpopulations on the walls of the large intestine (Fig. S1, Supplementary Information 6). By mitigating the effects of fluid flow, these wall-bound populations have an advantage in contributing genetic material to the future population. In extreme cases, these wall-bound populations can help recolonize the lumen after strong perturbations like diarrhea. However, their influence on the steady-state dynamics of genetic turnover in the gut are less well understood.

Previous work has argued that wall growth comprises a small fraction of the daily biomass production in the human gut, since the effective linear density in the wall ($\rho_w$) is much smaller than the corresponding density in the lumen ($\rho_\ell$) (26, 28). By generalizing our separation-of-timescales approach to this regime, we find that the long-term rates of genetic turnover critically depend on the *engraftment rate* $\varepsilon$, which represents the per-generation probability for a luminal cell to successfully migrate into one of the subpopulations on the wall (Supplementary Information 6). When wall growth comprises a small fraction of the total biomass production ($\rho_w\lambda_w \ll \rho_\ell\lambda_\ell$), we find even tiny rates of engraftment,

$$\varepsilon_c \approx \left(\frac{\rho_w\lambda_w}{\rho_\ell\lambda_\ell}\right)^{3/2}, \tag{15}$$

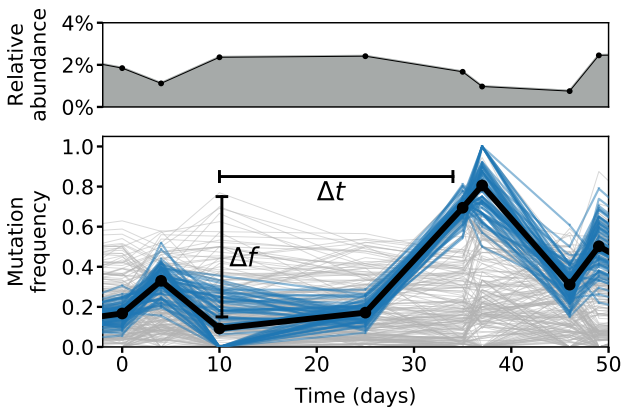


FIG. 5 **Mutation frequencies within a *Bacteroides massiliensis* population from a single human subject.** Relative abundance of *B. massiliensis* (top) and the frequencies of single nucleotide variants in this species (bottom) inferred from longitudinally sequenced fecal samples from a previous study (15) (Methods). Blue trajectories highlight a cluster of mutations that experienced a large shift in frequency during the study ($\Delta f \approx 50\%$), despite relatively small changes in the overall relative abundance of *B. massiliensis* ($\sim$2-fold; top). The timescale of this genetic shift ($\Delta t \sim 25$ days) is many orders of magnitude faster than our predicted timescale of genetic drift ($\Delta t \sim 1/\Lambda$; Fig. 4 and Eq. 16), even after accounting for the observed fluctuations in overall relative abundance.

are sufficient to ensure that genetic drift is dominated by the luminal compartment ($\Lambda \approx \Lambda_\ell$). This suggests that there is a broad parameter regime where our previous results still apply. When $\varepsilon \ll \varepsilon_c$, fluctuations within the wall population start to become more important, and the long-term rate of genetic drift is amplified by a factor of

$$\frac{\Lambda}{\Lambda_\ell} \approx \left(\frac{\varepsilon}{\varepsilon_c} + \sqrt{\frac{\rho_w}{\rho_\ell}\frac{\lambda_\ell}{\lambda_w}}\right)^{-2} \gg 1\,. \tag{16}$$

At even smaller engraftment rates, the separation-of-timescales approximation eventually breaks down, and the long-term dynamics start to be dominated by the sequential spread of strains across many isolated refugia (Supplementary Information 6). These slow spreading dynamics can lead to apparent rates of genetic drift that are even weaker than the lumen ($\Lambda \ll \Lambda_\ell$), despite the fact that the relevant population of cells is significantly smaller. This highlights how the dynamics of many distinct refugia can be qualitatively different than a single refuge on its own.

This non-monotonic behavior illustrates the importance of the engraftment rate $\varepsilon$, which is currently poorly constrained by experimental data. Previous estimates in the human gut suggest that the total wall growth is bounded by $\rho_w\lambda_w/\rho_\ell\lambda_\ell \lesssim 10^{-3}$ (26). Using these estimates as a guide, our model predicts that wall growth should have a negligible impact on genetic drift as long as $\varepsilon \gtrsim 10^{-4}$, and that engraftment rates of order $\varepsilon \sim 10^{-6}$ are required to enhance the rate of drift by a factor of $\sim 10^3$. These values are still too small to explain the rapid shifts in frequency observed in sequenced fecal samples (Fig. 5).

Our analysis also highlights the importance of the underlying timescales of these dynamics: while smaller engraftment rates can produce higher rates of genetic drift, these fluctuations still require at least $\sim \rho_\ell\lambda_\ell/\rho_w\lambda_w$ generations to propagate back into the lumen before they are eventually sampled in feces (Supplementary Information 6). On shorter time scales, we expect that wall-dominated genetic drift will generate autocorrelated fluctuations in sequenced fecal samples, similar to the background selection scenario analyzed in Ref. (40). Understanding the dynamics of these lagged fluctuations – and the signatures they leave in genetic data like Fig. 5 – could provide a promising approach for measuring the impact of wall growth from bulk metagenomic sequencing data.

## DISCUSSION

The spatial organization of microbial communities can strongly influence their ecological and evolutionary dynamics. Here we sought to quantify these effects in the human gut microbiota, whose long-term composition is shaped by lineages that manage to evade eventual dilution in feces. We have developed a mathematical framework for predicting how stochastic evolutionary forces emerge from simple models of bacterial growth in the intestinal lumen, and how they give rise to the genetic variation that can be observed in sequenced fecal samples.

While our simplified models neglect many important features of microbial and host physiology, they make several new predictions about the role of spatial structure for the within-host evolution of the gut microbiota. Chief among these is the finding that the emergent evolutionary dynamics can often be captured by effective models that lack explicit spatial structure, even in cases where longitudinal gradients have a large effect on local species composition. Similar results are implicit in earlier studies that focused on neutral fixation and heterozygosity (37, 38, 39, 65). Here we showed that this equivalence extends to wider range of evolutionary phenomena, including the establishment and spread of beneficial mutations. This suggests that well-mixed models from population genetics might have broader utility for modeling the dynamics of genetic variants in sequenced fecal samples, despite their explicit spatial structure. In these cases, our results provide a basis for connecting the inferred parameters (e.g. effective selection coefficients and population sizes) with specific weighted averages over the underlying

growth and density profiles.

Consistent with previous intuition, we found that unidirectional fluid flow can have a large impact on the emergent evolutionary parameters in the gut. Populations that are close to the washout threshold experience dramatically elevated rates of genetic drift — much larger than would be predicted by the absolute number of dividing cells (Supplementary Information 4.2). This shows how long-term lineage dynamics can play an essential role in extrapolating evolutionary parameters from absolute growth and density measurements.

By applying our results to recent estimates from the human gut, we found that continuous fluid flow (and simple forms of wall growth) are unlikely to lower effective population sizes to the point where we would expect large fluctuations in genetic composition within a host lifetime. This suggests that additional factors, such as natural selection or temporal bottlenecks, are necessary to explain the rapidly changing mutation frequencies that have been observed in sequenced human fecal samples (7, 13, 14, 15, 16, 17, 18, 19, 20). We also found that the spatial gradients generated by fluid flow can cause the effective generation time to be about 5-10 times shorter than the traditional $\sim$1/day estimates obtained from the average growth rate across the colon (59, 60). This suggests that the human gut microbiota may experience substantially more evolution than previously expected over the course of a given year. It also suggests that relatively small fitness advantages ($s \sim 1-10\%$) may be sufficient to drive large changes in the frequencies of genetic variants over daily and weekly timescales. We note that our precise estimates should be treated with a degree of caution, since they are currently derived from relatively rough estimates the underlying growth rates and density profiles in the gut (28). Our framework provides a flexible approach for refining these estimates using future measurements of spatially resolved growth rates (66) or microbial abundances (32) in human subjects. Our results also highlight the spatial locations where additional measurements would be most informative for constraining our estimates of the emergent evolutionary parameters.

Future experiments could test these predictions quantitatively using genetic barcoding techniques (67, 68, 69), both in animal models (12), and potentially even in humans (70). For example, by introducing libraries of barcoded strains and tracking their relative fluctuations in subsequent fecal samples, it should be possible to directly measure the effective rate of genetic drift in Eq. (5) to test our theoretical predictions in Fig. 4. Even stronger tests are possible if the strain labeling events are coupled to local environmental conditions (e.g. pH, mucin, etc.) (71, 72, 73). By tracking the frequencies of these labeled lineages as they spread out across the colon — and potentially into the intestinal walls (32) — one could directly measure the weighting function $u(x)$ and its associated mixing timescales. Our results provide a framework for interpreting these spatiotemporal measurements, and for using them to extract key unknown parameters like the engraftment rate $\varepsilon$ (Supplementary Information 6).

The limitations of our simple growth model create many additional directions for future work. For example, our results suggest that it will be critical to understand how temporal variation in growth conditions – either from regular circadian rhythms (74, 75) or larger perturbations like antibiotics (15) – might alter the emergent evolutionary forces we have studied here. It would also be interesting to examine how more complex forms of wall growth or other spatial refugia (61, 76, 77) might contribute to long-term genetic turnover within the gut. While the details of these models will necessarily differ from the ones studied here, our results show how similar separation-of-timescales approaches may continue to be useful for analyzing these additional scenarios as well. By "integrating out" variation over shorter length- and time-scales, these coarse-graining approaches can provide a promising route for modeling evolutionary dynamics in complex natural settings (78).

## DATA AND CODE AVAILABILITY

Source code for the data analysis and figure generation is available at Github (`https://github.com/bgoodlab/spatial_gut_evolution`).

## ACKNOWLEDGMENTS

This work was supported in part by a National Science Foundation Graduate Research Fellowship, the Alfred P. Sloan Foundation, and a Terman Fellowship from Stanford University.

## REFERENCES

[1] Cho I, Blaser MJ (2012) The human microbiome: at the interface of health and disease. *Nature Reviews Genetics* 13(4):260–270.

[2] Gilbert JA, et al. (2018) Current understanding of the human microbiome. *Nature medicine* 24(4):392–400.

[3] Eckburg PB, et al. (2005) Diversity of the human intestinal microbial flora. *science* 308(5728):1635–1638.

[4] Huttenhower C, et al. (2012) Structure, function and diversity of the healthy human microbiome. *Nature* 486(7402):207–214.

[5] Stephen AM, Cummings J (1980) The microbial contribution to human faecal mass. *Journal of Medical Microbiology* 13(1):45–56.

[6] Cummings JH, Bingham SA, Heaton KW, Eastwood MA (1992) Fecal weight, colon cancer risk, and dietary intake of nonstarch polysaccharides (dietary fiber). *Gastroenterology* 103(6):1783–1789.

[7] Zhao S, et al. (2019) Adaptive evolution within gut microbiomes of healthy people. *Cell host & microbe* 25(5):656–667.

[8] Shoemaker WR, Chen D, Garud NR (2021) Comparative Population Genetics in the Human Gut Microbiome. *Genome Biology and Evolution*. evab116.

[9] Barroso-Batista J, et al. (2014) The first steps of adaptation of escherichia coli to the gut are dominated by soft sweeps. *PLoS genetics* 10(3).

[10] Lescat M, et al. (2017) Using long-term experimental evolution to uncover the patterns and determinants of molecular evolution of an escherichia coli natural isolate in the streptomycin-treated mouse gut. *Molecular ecology* 26(7):1802–1817.

[11] Yilmaz B, et al. (2021) Long-term evolution and short-term adaptation of microbiota strains and sub-strains in mice. *Cell Host & Microbe* 29(4):650–663.

[12] Vasquez KS, et al. (2020) Quantifying the interplay between rapid bacterial evolution within the mouse intestine and transmission between hosts. *bioRxiv*.

[13] Ghalayini M, et al. (2018) Evolution of a dominant natural isolate of escherichia coli in the human gut over the course of a year suggests a neutral evolution with reduced effective population size. *Appl. Environ. Microbiol.* 84(6):e02377–17.

[14] Garud NR, Good BH, Hallatschek O, Pollard KS (2019) Evolutionary dynamics of bacteria in the gut microbiome within and across hosts. *PLoS biology* 17(1):e3000102.

[15] Roodgar M, et al. (2019) Longitudinal linked read sequencing reveals ecological and evolutionary responses of a human gut microbiome during antibiotic treatment. *bioRxiv*.

[16] Yaffe E, Relman DA (2020) Tracking microbial evolution in the human gut using hi-c reveals extensive horizontal gene transfer, persistence and adaptation. *Nature Microbiology* 5(2):343–353.

[17] Poyet M, et al. (2019) A library of human gut bacterial isolates paired with longitudinal multiomics data enables mechanistic microbiome research. *Nature medicine* 25(9):1442–1452.

[18] Zlitni S, et al. (2020) Strain-resolved microbiome sequencing reveals mobile elements that drive bacterial competition on a clinical timescale. *Genome medicine* 12:1–17.

[19] Verster AJ, et al. (2017) The landscape of type vi secretion across human gut microbiomes reveals its role in community composition. *Cell host & microbe* 22(3):411–419.

[20] Jiang X, et al. (2019) Invertible promoters mediate bacterial phase variation, antibiotic resistance, and host adaptation in the gut. *Science* 363(6423):181–187.

[21] Ferreiro A, Crook N, Gasparrini AJ, Dantas G (2018) Multiscale evolutionary dynamics of host-associated microbiomes. *Cell* 172(6):1216–1227.

[22] Earle KA, et al. (2015) Quantitative imaging of gut microbiota spatial organization. *Cell host & microbe* 18(4):478–488.

[23] Mark Welch JL, Hasegawa Y, McNulty NP, Gordon JI, Borisy GG (2017) Spatial organization of a model 15-member human gut microbiota established in gnotobiotic mice. *Proceedings of the National Academy of Sciences* 114(43):E9105–E9114.

[24] Shi H, et al. (2020) Highly multiplexed spatial mapping of microbial communities. *Nature* 588(7839):676–681.

[25] Donaldson GP, Lee SM, Mazmanian SK (2016) Gut biogeography of the bacterial microbiota. *Nature Reviews Microbiology* 14(1):20–32.

[26] Cremer J, et al. (2016) Effect of flow and peristaltic mixing on bacterial growth in a gut-like channel. *Proceedings of the National Academy of Sciences* 113(41):11414–11419.

[27] Cummings J, Macfarlane G (1991) The control and consequences of bacterial fermentation in the human colon. *Journal of Applied Bacteriology* 70(6):443–459.

[28] Cremer J, Arnoldini M, Hwa T (2017) Effect of water flow and chemical environment on microbiota growth and composition in the human colon. *Proceedings of the National Academy of Sciences* 114(25):6438–6443.

[29] Evans D, et al. (1988) Measurement of gastrointestinal ph profiles in normal ambulant human subjects. *Gut* 29(8):1035–1041.

[30] Friedman ES, et al. (2018) Microbes vs. chemistry in the origin of the anaerobic gut lumen. *Proceedings of the National Academy of Sciences* 115(16):4170–4175.

[31] Sheth RU, et al. (2019) Spatial metagenomic characterization of microbial biogeography in the gut. *Nature biotechnology* 37(8):877–883.

[32] Zmora N, et al. (2018) Personalized gut mucosal colonization resistance to empiric probiotics is associated with unique host and microbiome features. *Cell* 174(6):1388–1405.

[33] Tropini C, Earle KA, Huang KC, Sonnenburg JL (2017) The gut microbiome: connecting spatial organization to function. *Cell host & microbe* 21(4):433–442.

[34] Montassier E, et al. (2021) Probiotics impact the antibiotic resistance gene reservoir along the human gi tract in a person-specific and antibiotic-dependent manner. *Nature Microbiology*.

[35] Poulsen LK, Licht TR, Rang C, Krogfelt KA, Molin S (1995) Physiological state of escherichia coli bj4 growing in the large intestines of streptomycin-treated mice. *Journal of Bacteriology* 177(20):5840–5845.

[36] Riglar DT, et al. (2019) Bacterial variability in the mammalian gut captured by a single-cell synthetic oscillator. *Nature communications* 10(1):1–12.

[37] Hallatschek O, Nelson DR (2008) Gene surfing in expanding populations. *Theoretical population biology* 73(1):158–170.

[38] Nagylaki T (2000) Geographical invariance and the strong-migration limit in subdivided populations. *Journal of mathematical biology* 41(2):123–142.

[39] Wilkins JF (2004) A separation-of-timescales approach to the coalescent in a continuous population. *Genetics* 168(4):2227–2244.

[40] Cvijović I, Good BH, Desai MM (2018) The effect of strong purifying selection on genetic diversity. *Genetics* 209(4):1235–1278.

[41] Crump BC, Hopkinson CS, Sogin ML, Hobbie JE (2004) Microbial biogeography along an estuarine salinity gradient: combined influences of bacterial growth and residence time. *Applied and environmental microbiology* 70(3):1494–1505.

[42] Cohan FM, Ward D (2005) *Microbial diversity in hot spring cyanobacterial mats: pattern and prediction.* (Thermal Biology Institute).

[43] Gardiner C (1985) *Handbook of Stochastic Methods.* (Springer, New York).

[44] Ben-Amor K, et al. (2005) Genetic diversity of viable, injured, and dead fecal bacteria assessed by fluorescence-activated cell sorting and 16s rrna gene analysis. *Applied and environmental microbiology* 71(8):4679–4689.

[45] Maurice CF, Haiser HJ, Turnbaugh PJ (2013) Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell* 152(1-2):39–50.

[46] Fisher RA (1937) The wave of advance of advantageous genes. *Annals of eugenics* 7(4):355–369.

[47] Kolmogorov A, Petrovsky I, Piskunov N (1937) Investigation of the equation of diffusion combined with increasing of the substance and its application to a biology problem. *Bull. Moscow State Univ. Ser. A: Math. Mech* 1(6):1–25.

[48] Skellam JG (1951) Random dispersal in theoretical populations. *Biometrika* 38(1/2):196–218.

[49] Hastings A, et al. (2005) The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters* 8(1):91–101.

[50] Van Saarloos W (2003) Front propagation into unstable states. *Physics reports* 386(2-6):29–222.

[51] Wang CH, Matin S, George AB, Korolev KS (2019) Pinned, locked, pushed, and pulled traveling waves in structured environments. *Theoretical population biology* 127:102–119.

[52] Birzu G, Hallatschek O, Korolev KS (2018) Fluctuations uncover a distinct class of traveling waves. *Proceedings of the National Academy of Sciences* 115(16):E3645–E3654.

[53] Birzu G, Matin S, Hallatschek O, Korolev KS (2019) Genetic drift in range expansions is very sensitive to density dependence in dispersal and growth. *Ecology Letters* 22(11):1817–1827.

[54] Fisher DS (2007) Evolutionary dynamics in *Complex Systems*, Les Houches, eds. Jean-Philippe Bouchaud MM, Dalibard J. (Elsevier) Vol. 85, pp. 395–446.

[55] Good BH (2016) Ph.D. thesis (Harvard University, Cambridge MA).

[56] Hallatschek O (2011) The noisy edge of traveling waves. *Proceedings of the National Academy of Sciences* 108(5):1783–1787.

[57] Fisher DS (2013) Asexual evolution waves: fluctuations and universality. *Journal of Statistical Mechanics: Theory and Experiment* 2013(01):P01011.

[58] Good BH, Desai MM (2014) Deleterious passengers in adapting populations. *Genetics* 198(3):1183–1208.

[59] Savageau MA (1983) Escherichia coli habitats, cell types, and molecular mechanisms of gene control. *The american naturalist* 122(6):732–744.

[60] Gibbons R, Kapsimalis B (1967) Estimates of the overall rate of growth of the intestinal microflora of hamsters, guinea pigs, and mice. *Journal of bacteriology* 93(1):510–512.

[61] Li H, et al. (2015) The outer mucus layer hosts a distinct intestinal microbial niche. *Nature communications* 6(1):1–13.

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

[63] Mark Welch JL, Rossetti BJ, Rieken CW, Dewhirst FE, Borisy GG (2016) Biogeography of a human oral microbiome at the micron scale. *Proceedings of the National Academy of Sciences* 113(6):E791–E800.

[64] Dodge R, et al. (2021) A gut commensal niche regulates stable association of a multispecies microbiota. *bioRxiv*.

[65] Labavić D, Loverdo C, Bitbol AF (2021) Hydrodynamic flow and concentration gradients in the gut enhance neutral bacterial diversity. *bioRxiv*.

[66] Korem T, et al. (2015) Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples. *Science* 349(6252):1101–1106.

[67] Levy SF, et al. (2015) Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature* 519(7542):181–186.

[68] Ba ANN, et al. (2019) High-resolution lineage tracking reveals travelling wave of adaptation in laboratory yeast. *Nature* 575(7783):494–499.

[69] Jasinska W, et al. (2020) Chromosomal barcoding of e. coli populations reveals lineage diversity dynamics at high resolution. *Nature ecology & evolution* 4(3):437–452.

[70] Maldonado-Gómez MX, et al. (2016) Stable engraftment of bifidobacterium longum ah1206 in the human gut depends on individualized features of the resident microbiome. *Cell host & microbe* 20(4):515–526.

[71] Naydich AD, et al. (2019) Synthetic gene circuits enable systems-level biosensor trigger discovery at the host-microbe interface. *MSystems* 4(4):e00125–19.

[72] Merk LN, Shur A, Pandey A, Murray RM, Green LN (2020) Engineering logical inflammation sensing circuit for modulating gut conditions. *bioRxiv*.

[73] Loveless TB, et al. (2021) Molecular recording of sequential cellular events into dna. *bioRxiv*.

[74] Thaiss CA, et al. (2014) Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. *Cell* 159(3):514–529.

[75] Leone V, et al. (2015) Effects of diurnal variation of gut microbes and high-fat feeding on host circadian clock function and metabolism. *Cell host & microbe* 17(5):681–689.

[76] Lourenço M, et al. (2020) The spatial heterogeneity of the gut limits predation and fosters coexistence of bacteria and bacteriophages. *Cell Host & Microbe* 28(3):390–401.

[77] Donaldson GP, et al. (2020) Spatially distinct physiology of bacteroides fragilis within the proximal colon of gnotobiotic mice. *Nature microbiology* 5(5):746–756.

[78] Bergelson J, Kreitman M, Petrov DA, Sanchez A, Tikhonov M (2021) Functional biology in its natural context: A search for emergent simplicity. *Elife* 10:e67646.

[79] Albenberg L, et al. (2014) Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology* 147(5):1055–1063.

[80] Espey MG (2013) Role of oxygen gradients in shaping redox relationships between the human intestine and its microbiota. *Free Radical Biology and Medicine* 55:130–140.

## SUPPLEMENTAL INFORMATION (SI)

### 1. MICROSCOPIC DERIVATION OF THE REACTION-DIFFUSION MODEL

In this section, we introduce a microscopic birth-and-death model that reproduces the continuum limit described by Eq. (3). We begin by deriving a mean-field model for the total population profile in Eq. (2). We assume that space and time are divided into discrete intervals of size $\Delta x$ and $\Delta t$, respectively. In a given interval $\Delta t$, we will let $m_v(x)$ denote the probability that an individual is carried from position $x$ to position $x+\Delta x$ by advection (flow). Likewise, we will let $2m_D(x)$ denote the probability that an individual moves to either $x+\Delta x$ or $x-\Delta x$ through diffusion (mixing). Finally, we will let $m_b(x|\{\rho(x,t)\})$ and $m_d(x|\{\rho(x,t)\})$ denote the probability that an individual gives birth or dies, respectively; these probabilities will generally depend on both the position $x$ as well as the current density profile, $\rho(x,t)$. In the limit of large densities, these assumptions yield a mean-field update rule:

$$\begin{aligned}\rho(x,t+\Delta t) = \rho(x,t) + m_b(x|\{\rho(x,t)\})\rho(x,t) - m_b(x|\{\rho(x,t)\})\rho(x,t) - [2m_D(x) + m_v(x)]\,\rho(x,t) \\ + [m_D(x-\Delta x) + m_v(x-\Delta x)]\,\rho(x-\Delta x,t) + m_D(x+\Delta x)\rho(x+\Delta x,t)\,,\end{aligned} \tag{S1}$$

whose continuum limit is given by

$$\partial_t\rho(x,t) = \frac{\partial^2}{\partial x^2}\left[\frac{m_D(x)\Delta x^2}{\Delta t}\cdot\rho(x,t)\right] - \frac{\partial}{\partial x}\left[\frac{m_v(x)\Delta x}{\Delta t}\cdot\rho(x,t)\right] + \left[\frac{m_b(x|\{\rho(x)\})}{\Delta t} - \frac{m_d(x|\{\rho(x)\})}{\Delta t}\right]\cdot\rho(x,t)\,, \tag{S2}$$

This allows us to identify the macroscopic parameters

$$D(x) \equiv \frac{m_D(x)\Delta x^2}{\Delta t}\,, \quad v(x) \equiv \frac{m_v(x)\Delta x}{\Delta t}\,,\,, \tag{S3a}$$

$$b(x|\{\rho(x)\}) \equiv \frac{m_b(x|\{\rho(x)\})}{\Delta t}\,, \quad d(x|\{\rho(x)\}) \equiv \frac{m_d(x|\{\rho(x)\})}{\Delta t}\,, \tag{S3b}$$

as well as the net growth rate,

$$\lambda(x|\{\rho(x)\}) \equiv b(x|\{\rho(x)\}) - d(x|\{\rho(x)\})\,, \tag{S3c}$$

so that

$$\partial_t\rho(x,t) = \partial_x^2\left[D(x)\rho(x,t)\right] - \partial_x\left[v(x)\rho(x,t)\right] + \lambda(x|\{\rho(x)\})\rho(x,t)\,. \tag{S4}$$

At steady state, this reduces to

$$0 = \partial_x^2[D(x)\rho(x)] - \partial_x[v(x)\rho(x)] + \lambda(x)\rho(x)\,, \tag{S5}$$

which can be rearranged to obtain Eq. (2) in the main text.

The relevant boundary conditions can be derived in several different ways. In the simplest version, we assume that there is no input or output flux of bacteria from advection or diffusion through the left boundary ($x=0$). This implies that

$$m_v(x-\Delta x)\rho(x-\Delta x) + m_D(x-\Delta x)\rho(x-\Delta x) - m_D(x)\rho(x) = 0 \tag{S6}$$

which reduces to

$$v(x)\rho(x) - \partial_x[D(x)\rho(x)] = 0 \tag{S7}$$

in the continuum limit as expected. Similarly, we assume that there is no influx or outflux of bacteria from diffusion at the right boundary ($x=L$). This yields

$$m_D(x+\Delta x)\rho(x+\Delta x) - m_D(x)\rho(x) = 0 \tag{S8}$$

or

$$\partial_x[D(x)\rho(x)] = 0\,. \tag{S9}$$

A related set of boundary conditions can be derived by assuming that $\lambda(x)$ is positive only in some finite region of space, and enforcing the physical constraint that $\rho(x) \to 0$ as $x \to -\infty$ and $\rho(x) \to \text{const}$ as $x \to \infty$. We will find that this unbounded formulation can sometimes be more convenient in our later derivations, while yielding qualitatively similar results as the bounded model above.

At steady state, these boundary conditions imply that total growth rate across the colon must equal the outflow at $x = L$,

$$\int_0^L \lambda(x)\rho(x)dx = v(L)\rho(L)\,, \tag{S10}$$

such that the average per capita growth rate is given by

$$\overline{\lambda} \equiv \frac{\int_0^L \lambda(x)\rho(x)\,dx}{\int_0^L \rho(x)\,dx} = \frac{v(L)\rho(L)}{\int_0^L \rho(x)\,dx}\,. \tag{S11}$$

In the limit of a long colon, the resident population will often be dominated by regions with saturated growth $[\int_0^L \rho(x)\,dx \approx L\rho(L)]$, so that

$$\overline{\lambda} \approx v(L)/L\,. \tag{S12}$$

This allows us to recover the standard relationship between the average replication time and the total transit time across the colon (59). We note that all of these calculations assume that the daily elimination of feces does not substantially distrurb the steady-state assumption in Eq. (S5). This is partially justified by estimates that only ∼20% of the colonic volume is excreted each day (28). However, further work is required to assess the impact of these daily fluctuations on the steady-state assumptions employed here.

#### Stochastic dynamics of a rare lineage

We are now in a position to derive the stochastic dynamics of a rare mutant lineage when the larger population has reached equilibrium. As long as the mutant is sufficiently rare, we expect it to have a negligible influence on the environmental parameters, $r_i(x,t)$, so that $\lambda(x)$ will remain close to its equilibrium form. We will now consider scenarios where the mutant has a of modified birth and death rate, $b_m(x)$ and $d_m(x)$, which will also be constant in time when the mutant is sufficiently rare. Since the mutant lineage will start out with a small number of cells, its density profile $\rho_m(x,t)$ must now be described by a stochastic update rule

$$\rho_m(x, t+\Delta t) = \rho_m(x,t) + (\delta\rho_m(x))_{\text{det}} + (\delta\rho_m(x))_{\text{stoch}}\,, \tag{S13}$$

where the deterministic term is similar to the mean-field dynamics in Eq. (S4) above:

$$\begin{aligned}(\delta\rho_m(x))_{\text{det}} = {} & [m_{b,m}(x) - m_{d,m}(x)]\rho_m(x,t) - [2m_D(x) + m_v(x)]\,\rho_m(x,t) \\ & + [m_D(x-\Delta x) + m_v(x-\Delta x)]\,\rho_m(x-\Delta x,t) + m_D(x+\Delta x)\rho_m(x+\Delta x,t)\,.\end{aligned} \tag{S14}$$

The stochastic term is given by

$$\begin{aligned}(\delta\rho_m(x))_{\text{stoch}} = {} & \sqrt{m_{b,m}(x)\rho_m(x)}Z_x^b + \sqrt{m_{d,m}(x)\rho_m(x)}Z_x^d - \sqrt{m_D(x)\rho_m(x)}Z_x^{DL} - \sqrt{m_D(x)\rho_m(x)}Z_x^{DR} \\ & - \sqrt{m_v(x)\rho_m(x)}Z_x^v + \sqrt{m_D(x-\Delta x)\rho(x-\Delta x)}Z_{x-\Delta x}^{DR} \\ & + \sqrt{m_v(x-\Delta x)\rho_m(x-\Delta x)}Z_{x-\Delta x}^v + \sqrt{m_D(x+\Delta x)\rho_m(x+\Delta x)}Z_{x+\Delta x}^{DL}\,,\end{aligned} \tag{S15}$$

where the $\{Z_x^i\}$ are independent Gaussian random variables with zero mean and unit variance:

$$\langle Z_x^i Z_y^j \rangle = \delta_{i,j}\delta_{x,y}\,. \tag{S16}$$

Note that due the advection and diffusion terms, the stochastic term is not diagonal in the $\rho_m(x)$ basis, since it will generally include correlations from neighboring $x$ values. In our analysis below, we will often be interested in the

moments of quantities like $\int \phi(x)\delta\rho_m(x)\,dx$, where $\phi(x)$ is an arbitrary deterministic function. In this case, it is often helpful to diagonalize these integrals in the $\rho_m(x)$ basis by rewriting them in the form

$$\begin{aligned}\int \phi(x)[\delta\rho_m]_{\rm det}dx = \int \{&\phi(x)[m_{b,m}(x) - m_{d,m}(x)]\rho_m(x) + [\phi(x+\Delta x) - \phi(x)]m_v(x)\rho_m(x) \\ &+[\phi(x+\Delta x) - \phi(x)]m_D(x)\rho_m(x) + [\phi(x-\Delta x) - \phi(x)]m_D(x)\rho_m(x)\}\; dx\end{aligned} \tag{S17}$$

and

$$\begin{aligned}\int \phi(x)[\delta\rho_m]_{\rm stoch}dx = \int \Big\{&\phi(x)\sqrt{[m_{b,m}(x) + m_{d,m}(x)]\rho_m(x)}Z_x^{\lambda} + [\phi(x-\Delta x) - \phi(x)]\sqrt{m_D(x)\rho_m(x)}Z_x^{DL} \\ &+[\phi(x+\Delta x) - \phi(x)]\sqrt{m_D(x)\rho_m(x)}Z_x^{DR} \\ &+[\phi(x+\Delta x) - \phi(x)]\sqrt{m_v(x)\rho_m(x)}Z_x^{v}\Big\}\; dx\,,\end{aligned} \tag{S18}$$

These identities are particularly useful for calculating the generating functional of the mutant lineage, starting from an initial density profile $\rho_m(x,0)$:

$$H[\phi(x),t] = \left\langle e^{-\int \phi(x)\rho_m(x,t)}\right\rangle\,, \tag{S19}$$

Recursion relations for $H[\phi(x),t]$ can be obtained using standard approaches (57, 58). For example, by considering the change in $H[\phi(x),t]$ over a single timestep $\Delta t$, we can rewrite the generating function as

$$H[\phi(x),t+\Delta t] = \left\langle e^{-\int \phi(x)\rho_m(x)dx}\cdot e^{-\int \phi(x)[\delta\rho_m(x)]_{\rm det}dx}\cdot e^{-\int \phi(x)[\delta\rho_m(x)]_{\rm stoch}dx}\right\rangle_{Z_x^i,\rho_m(x)} \tag{S20}$$

where the angle brackets denote an average over both $Z_x^i$ and $\rho_m(x)$. By expanding the exponentials to leading order in $m_i(x)$ and performing the relevant averages, we obtain a standard system of equations

$$\begin{aligned}\frac{\partial H}{\partial t} = \Big\{&\phi(x)\frac{m_{b,m}(x) - m_{d,m}(x)}{\Delta t} + [\phi(x+\Delta x) - \phi(x)]\frac{m_v(x)}{\Delta t} + [\phi(x+\Delta x) - \phi(x)]\frac{m_D(x)}{\Delta t} \\ &+[\phi(x-\Delta x) - \phi(x)]\frac{m_D(x)}{\Delta t} - \frac{\phi(x)^2[m_{b,m}(x) + m_{d,m}(x)]}{2\Delta t} - \frac{[\phi(x+\Delta x) - \phi(x)]^2 m_v(x)}{2\Delta t} \\ &-\frac{[\phi(x+\Delta x) - \phi(x)]^2 m_D(x)}{2\Delta t} - \frac{[\phi(x+\Delta x) - \phi(x)]^2 m_D(x)}{2\Delta t}\Big\}\frac{\partial H}{\partial \phi(x)}\,.\end{aligned} \tag{S21}$$

In the continuum limit, this system reduces to the partial differential equation,

$$\frac{\partial H}{\partial t} = \left\{D(x)\partial_x^2\phi(x) + v(x)\partial_x\phi(x) + \lambda_m(x)\phi(x) - \frac{\Upsilon_m(x)\phi(x)^2}{2} - D(x)[\partial_x\phi(x)]^2\right\}\frac{\partial H}{\partial \phi(x)}\,, \tag{S22}$$

subject to the initial condition

$$H[\phi(x),0] = e^{-\int \phi(x)\rho_m(x,0)dx}\,, \tag{S23}$$

where we have defined the net growth rate,

$$\lambda_m(x) = b_m(x) - d_m(x)\,, \tag{S24}$$

and total variance in offspring number,

$$,\quad \Upsilon_m(x) = b_m(x) + d_m(x)\,. \tag{S25}$$

The generating function in Eq. (S22) is identical to the one produced by the Langevin model in Eq. (3). This shows that Eq. (3) can be viewed as the continuum limit of the microscopic model in Eqs. (S14) and (S15). Both models are similar to the linear branching processes that have been analyzed in previous work on evolutionary traveling waves (56, 57, 58). The main difference in this case is the additional nonlinear derivative of $\phi(x)$ in Eq. (S22) which arises from the conservative nature of the diffusive fluctuations in space.

The linearity of the noise term in Eqs. (3) and (S22) yields a straightforward equation for the average mutant density profile,

$$\partial_t \langle \rho_m(x,t)\rangle = \partial_x^2[D(x)\langle \rho_m(x,t)\rangle] - \partial_x[v(x)\langle \rho_m(x,t)\rangle] + \lambda_m(x)\langle \rho_m(x,t)\rangle \,, \tag{S26}$$

which is similar to the mean field dynamics in Eq. (S4). Formal solutions for the entire distribution of $\rho_m(x,t)$ can also be obtained from Eq. (S22) using the method of characteristics. If we let $\tau$ measure time before the present, then the characteristic curves $\phi(x,\tau)$ are defined by

$$\partial_\tau \phi(x,\tau) = D(x)\partial_x^2\phi(x,\tau) + v(x)\partial_x\phi(x,\tau) + \lambda_m(x)\phi(x,\tau) - \frac{\Upsilon_m(x)\phi(x)^2}{2}\phi(x,\tau)^2 - D(x)[\partial_x\phi(x,\tau)]^2 \,, \tag{S27}$$

with the initial condition $\phi(x,0) = \phi(x)$. The time-dependent generating function is then given by

$$H[\phi(x),t] = e^{-\int \phi(x,t)\rho_m(x,0)} \,. \tag{S28}$$

We will explore the behavior of this solution in several specific cases below.

## 2. DYNAMICS OF A NEUTRAL LINEAGE

We first consider the dynamics of a neutral lineage [$\lambda_m(x) = \lambda(x)$, $\Upsilon_m(x) = \Upsilon(x)$], where the characteristic function reduces to

$$\partial_\tau \phi(x,\tau) = D(x)\partial_x^2\phi(x,\tau) + v(x)\partial_x\phi(x,\tau) + \lambda(x)\phi(x,\tau) - \frac{\Upsilon(x)\phi(x)^2}{2} - D(x)[\partial_x\phi(x,\tau)]^2 \,. \tag{S29}$$

To analyze this equation, it will be useful to first consider the behavior of the linear version,

$$\partial_\tau \phi_0(x,\tau) = D(x)\partial_x^2\phi_0(x,\tau) + v(x)\partial_x\phi_0(x,\tau) + \lambda(x)\phi_0(x,\tau) \,, \tag{S30}$$

which is obtained by dropping the $\phi^2$ and $(\partial_x\phi)^2$ terms in Eq. (S29). We note that the right-hand side of Eq. (S29) is the adjoint of the deterministic dynamics for $\langle \rho_m(x,t)\rangle$ in Eq. (S26). Thus, if $\phi_0(x,\tau)$ is a solution of Eq. (S30), then

$$\langle \rho_m(x,t)\rangle \propto \cdot \frac{\phi_0(x,t)}{D(x)} e^{\int_0^x \frac{v(x')}{D(x')}dx'} \tag{S31}$$

is also a solution for $\langle \rho_m(x,t)\rangle$ for the particular initial condition

$$\langle \rho_m(x,0)\rangle \propto \cdot \frac{\phi_0(x,0)}{D(x)} e^{\int_0^x \frac{v(x')}{D(x')}dx'} \,. \tag{S32}$$

Since the average density profile of a neutral lineage must eventually approach $\langle \rho_m(x,t)\rangle \propto \rho(x)$ at long times, Eq. (S31) shows that $\phi_0(x,\tau)$ must also approach a steady state shape as $\tau \to \infty$. The linearity of Eq. (S30) implies that these steady-state solutions can be expressed in the general form

$$\phi_0^*(x|\{\phi(x',0)\}) = \left[\int C(x')\phi(x',0)dx'\right] u(x) \,, \,, \tag{S33}$$

where $C(x)$ is a weighting function that we determine below, and

$$u(x) = \frac{\rho(x)\cdot D(x) e^{-\int_0^x \frac{v(x')}{D(x')}dx'}}{\int \rho(x')^2 \cdot D(x') e^{-\int_0^{x'} \frac{v(x'')}{D(x'')}dx''}\, dx'} \tag{S34}$$

is a solution to the fixed point equation,

$$0 = D(x)\partial_x^2 u(x) + v(x)\partial_x u(x) + \lambda(x)u(x) \,, \tag{S35}$$

with an overall normalization set by $\int u(x)\rho(x)dx = 1$. We have adopted this normalization convention to be consistent with the $u(x)$ function defined in Ref. (37). In that work, the $u(x)$ function was initially derived in terms of the Greens

function of the deterministic dynamics in Eq. (S4). Here we see that it also arises as a solution to the linearized version of the neutral characteristic function in Eq. (S29). We discuss this connection in more detail below.

The linearized solutions in Eq. (S33) will be important because they capture the leading order behavior of the characteristic curves $\phi(x,\tau)$ in the limit of small $\phi(x,0)$. For example, by substituting Eq. (S33) into Eq. (S28), we find that the leading order contribution to the generating function is given by

$$-\log H[\{\phi(x,0)\},t] = \int \phi(x,\tau)\rho_m(x,0)dx \to \left[\int C(x)\phi(x,0)\,dx\right]\left[\int u(x')\rho_m(x',0)dx\right] + \mathcal{O}(\phi(x,0)^2)\,. \tag{S36}$$

The average mutant density profile, $\langle\rho_m(x,t)\rangle$, then follows as

$$\langle\rho_m(x,t)\rangle \equiv -\frac{\partial \log H[\{\phi(x,0)\},t]}{\partial\phi(x,0)} \to C(x)\int u(x')\rho_m(x',0)\,dx'\,. \tag{S37}$$

In particular, this shows that if $\rho_m(x,0)$ is proportional to the steady state density profile, $\rho_m(x,0) = f\rho(x)$, then $\langle\rho_m(x,t)\rangle$ must be constant over time:

$$f\rho(x) = C(x)\int u(x')\cdot f\rho(x',0)dx' = C(x)f\,. \tag{S38}$$

This implies that $C(x) = \rho(x)$, and hence

$$\phi^*(x|\{\phi(x',0)\}) = \left[\int \rho(x')\phi(x',0)dx'\right]u(x)\,. \tag{S39}$$

The functional form of Eq. (S39) implies that $\rho(x')u(x)$ is the Greens function for the linearized dynamics in Eq. (S30) in the limit that $\tau\to\infty$, just as $\rho(x)u(x')$ is the Greens function for the deterministic dynamics of $\langle\rho_m(x,t)\rangle$ (37).

We will let $\tau_{\text{mix}}$ denote the "mixing timescale" required for $\phi_0(x,\tau)$ to approach its fixed point $\phi^*(x)$. This timescale can have a complicated dependence on $\lambda(x)$, $v(x)$, and $D(x)$, but due to the linearity of Eq. (S30), it must be independent of the overall scale of $\phi(x,0)$ — or equivalently, of $\rho_m(x,t)$. This suggests that in sufficiently large populations, we will be able to exploit a separation of timescales between the linearized growth dynamics ($\tau_{\text{mix}}$) and the slower evolutionary process of genetic drift [$\propto 1/\int\rho_m(x,t)dx$].

### 1. Variance accumulated due to genetic drift

As an example of this separation-of-timescales approach, we note that the variance of $\rho_m(x,t)$ can be obtained from the second order corrections to $\phi(x,\tau)$ in the limit that $\phi(x,0)$ is small:

$$-\log H[\{\phi(x,0)\},t] \sim \int \phi(x,0)\langle\rho_m(x,t)\rangle - \frac{1}{2}\int \phi(x,0)\phi(x'0)\text{Cov}(\rho_m(x,t),\rho_m(x',t)) + \mathcal{O}(\phi(x,0)^3)\,. \tag{S40}$$

We can calculate these contributions by computing the next order terms in the perturbation expansion for $\phi(x,\tau)$ above. By treating the nonlinear terms as a small correction, a basic iterative solution shows that the $\tau \gg \tau_{\text{mix}}$ behavior is given by

$$\begin{aligned}\phi(x,\tau) \sim &\left[\int \phi(x',0)\rho(x')dx'\right]u(x)\\ &-\frac{1}{2}\left[\int \phi(x',0)\rho(x')dx'\right]^2 \underbrace{\left[\int \left(\Upsilon(x')u(x')^2 + 2D(x)(\partial_x u(x'))^2\right)\rho(x')dx'\right]}_{\Lambda}\cdot u(x)\cdot\tau\,,\end{aligned} \tag{S41}$$

where $\Lambda$ coincides with the effective rate of genetic drift in Eq. (7) in the main text. This iterative solution in Eq. (S41) is self-consistently valid when

$$\tau_{\text{mix}} \ll \tau \ll \left[\int \phi(x',0)\rho(x')dx'\right]^{-1}\Lambda^{-1}\,, \tag{S42}$$

which can be satisfied for any fixed value of $\tau \gg \tau_{\text{mix}}$ by choosing sufficiently small values of $\phi(x,0)$. Note, however, that the converse is not true: for any fixed value of $\phi(x,0) \ll 1/\Lambda\tau_{\text{mix}}$, the perturbative solution in Eq. (S41) will eventually break down for sufficiently large $\tau$.

Fortunately, the variance of $\rho_m(x,t)$ can be extracted from the generating function $H[\{\phi(x)\},t]$ by choosing a sufficiently low value of $\phi(x,0) = \phi(x)$ so that Eq. (S41) remains valid for each desired value of $t$. By substituting this solution into Eq. (S28), we see that the second order contributions to the generating function are given by

$$\begin{aligned}\log H[\{\phi(x,0)\},t] \approx &-\left[\int \phi(x',0)\rho(x')dx'\right]\left[\int u(x')\rho_m(x',0)dx'\right] \\ &+\frac{1}{2}\left[\int \phi(x,0)\rho(x)dx\right]^2\left[\int u(x')\rho_m(x',0)dx'\right]\Lambda t\,,\end{aligned} \tag{S43}$$

so that

$$\text{Var}(\rho_m(x,t)) = \rho(x)^2\left[\int u(x')\rho_m(x',0)dx'\right]\Lambda t\,. \tag{S44}$$

This solution is self-consistently valid as long as $t \gg \tau_{\text{mix}}$. By expressing this result as a coefficient of variation:

$$c_V^2(\rho_m(x,t)) \equiv \frac{\text{Var}(\rho_m(x,t))}{\langle\rho_m(x,t)\rangle^2} = \frac{\Lambda t}{\left[\int u(x')\rho_m(x',0)dx'\right]}\,, \tag{S45}$$

we obtain a natural definition of the drift timescale,

$$\tau_{\text{drift}} \equiv \Lambda^{-1}\int u(x)\rho_m(x,0)dx\,, \tag{S46}$$

based on the time that the coefficient of variation first exceeds one.

### 2. Survival probability at long times

It will also be useful to calculate $\phi(x,\tau)$ on timescales much longer than $\tau_{\text{drift}}$. This is particularly important for calculating the survival probability of a lineage at long times. This survival probability can be calculated from an asymptotic expansion of the moment generating function in the limit that $t \gg \tau_{\text{drift}} \gg \tau_{\text{mix}}$:

$$H[\{\phi(x,0\},t] \approx [1-p_s(t)]e^{-0} + p_s(t)e^{-\mathcal{O}(t)}\,. \tag{S47}$$

Plugging in our method-of-characteristics solution for $H[\phi(x),t]$ above, we find that

$$p_s(t) \sim \int \phi(x,t)\rho_m(x,0) \qquad (t \gg \tau_{\text{drift}} \gg \tau_{\text{mix}}) \tag{S48}$$

In this case, we can no longer calculate $\phi(x,\tau)$ by appealing to the simple perturbative expansion above, since the nonlinear terms (which represent the effects of fluctuations) will eventually become important for sufficiently large $\tau$. Nevertheless, due to the separation of timescales, we expect that $\phi(x,t)$ will remain close to $\phi^*(x)$ in a sense that we will make more precise below. At long times, we know that $\phi(x,t)$ must eventually vanish (since neutral mutations must always go extinct in an infinitely large population). Motivated by the behavior in the single locus case (54, 55), we define a new function $\psi(x,\tau)$ such that

$$\phi(x,\tau) = \frac{\phi^*(x)}{\tau}\cdot\psi(x,\tau)\,, \tag{S49}$$

anticipating that $\psi(x,\tau)$ approach some time-independent shape in the limit of large $\tau$. In fact, it will be useful to split up $\psi(x,\tau)$ into three independent components:

$$\psi(x,\tau) = \underbrace{[\psi(0,\infty)]}_{B}\cdot\underbrace{\left[\frac{\psi(0,\tau)}{\psi(0,\infty)}\right]}_{f(\tau)}\cdot\underbrace{\left[\frac{\psi(x,\tau)}{\psi(0,\tau)}\right]}_{g(x,\tau)}\,, \tag{S50}$$

where $B$ is an arbitrary constant and $f(\tau)$ and $g(x,\tau)$ are unknown functions that satisfy the boundary conditions

$$f(\infty) = 1\,, \qquad g(0,\tau) = 1\,. \tag{S51}$$

The spatial boundary conditions on $\phi(x,\tau)$ provide an additional set of boundary conditions for the function $g(x,\tau)$. At $x = 0$, we have

$$0 = \partial_x \phi(0,\tau) = \frac{B}{\tau} \cdot f(\tau) g(0,\tau) \cdot \partial_x \phi^*(0) + \frac{B\phi^*(x)}{\tau} \cdot f(\tau) \partial_x g(0,\tau)\,, \tag{S52}$$

which implies that

$$\partial_x g(0,\tau) = 0\,. \tag{S53}$$

Likewise, at $x \to \infty$ we must have $\phi(x,\tau) \to 0$, which implies that

$$\lim_{x\to\infty} \phi^*(x) g(x,\tau) \to 0\,. \tag{S54}$$

Substituting these definitions into Eq. (S29), we obtain a corresponding characteristic curve for $g(x,\tau)$:

$$\begin{aligned} 0 = \frac{g(x,\tau)}{\tau} - g(x,\tau)\partial_\tau \log f(\tau) - \partial_\tau g(x,\tau) + D(x)\partial_x^2 g(x,\tau) + \left[v(x) + 2D(x)\partial_x \log \phi^*(x)\right] \partial_x g(x,\tau) \\ - \frac{B f(\tau)\phi^*(x) g(x,\tau)}{2\tau} \left(\Upsilon(x) + 2D(x)[\partial_x \log \phi^*(x) + \partial_x \log g(x,\tau)]^2\right)\,, \end{aligned} \tag{S55}$$

subject to the boundary conditions

$$g(0,\tau) = 1\,, \quad \partial_x g(0,\tau) = 0\,, \quad \lim_{x\to\infty} \phi^*(x) g(x,\tau) = 0 \tag{S56}$$

We now seek an asymptotic solution of Eq. (S55) in the limit of large $\tau$. We assume that $f(\tau)$ and $g(\tau)$ can be written as a power series in $1/\tau$:

$$f(\tau) \sim f_0 + \frac{f_1}{\tau} + \ldots\,, \quad g(x,\tau) \sim g_0(x) + \frac{g_1(x)}{\tau} + \ldots \tag{S57}$$

The boundary conditions on $f(\tau)$ and $g(x,\tau)$ imply that

$$f_0 = 1\,, \quad g_0(0) = 1\,, \quad g_1(0) = 0\,, \quad \partial_x g_i(x,0) = 0\,, \; \lim_{x\to\infty} \phi^*(x) g_i(x) = 0\,. \tag{S58}$$

After substituting these expansions into Eq. (S55) and collecting like powers of $1/\tau$, we find that the zeroth order contribution yields:

$$0 = D(x)\partial_x^2 g_0(x) + \left[v(x) + 2D(x)\partial_x \log \phi^*(x)\right] \partial_x g_0(x,\tau)\,. \tag{S59}$$

The only solution that satisfies the boundary conditions at $x = 0$ and $x = \infty$ is

$$g_0(x) = 1\,. \tag{S60}$$

The next order contribution from Eq. (S55) then yields

$$\begin{aligned} 0 = 1 + D(x)\partial_x^2 g_1(x) + \left[v(x) + 2D(x)\partial_x \log \phi^*(x)\right] \partial_x g_1(x,\tau) \\ - \frac{B\phi^*}{2} \left[\Upsilon(x) + 2D(x)(\partial_x \log \phi^*)^2\right]\,, \end{aligned} \tag{S61}$$

or

$$\partial_x^2 g_1(x) + \left[\frac{v(x)}{D(x)} + 2\partial_x \log \phi^*(x)\right] \partial_x g_1(x) = \frac{B\phi^*(x)}{2D(x)} \left[\Upsilon(x) + 2D(x)(\partial_x \log \phi^*(x))^2\right] - \frac{1}{D(x)}\,. \tag{S62}$$

This is a first order equation for $\partial_x g_1(x,)$, which has the solution

$$g_1(x) = \int_0^x \frac{e^{-\int_0^{x'} \frac{v(x'')dx''}{D(x'')}}}{\phi^*(x')^2} \int_0^{x'} e^{\int_0^{x''} \frac{v(x''')dx'''}{D(x''')}} \left[\frac{B\phi^*(x'')^3}{2D(x'')} \left[\Upsilon(x'') + 2D(x'')(\partial_x \log \phi^*)^2\right] - \frac{\phi^*(x'')^2}{D(x'')}\right] dx''\,. \tag{S63}$$

The limits of integration in Eq. (S63) were chosen to satisfy the boundary conditions $g_1(0) = 0$ and $\partial_x g_1(0) = 0$.

To apply the remaining boundary condition at $x = \infty$, it is helpful to impose the mild assumption that $\lambda(x)$ eventually vanishes for sufficiently large $x$. This implies that the solutions for $\phi^*(x)$ must scale as

$$\phi^*(x) \sim Ce^{-\int_0^x \frac{v(x')dx''}{D(x')}} \tag{S64}$$

for sufficiently large $x$. This means that the prefactor inside of the first integral must correspondingly grow as $\sim e^{+\int_0^x \frac{v(x')dx''}{D(x')}}$. Similarly, at large $x$, the inner integral must eventually approach a constant value at large $x$. The only way that the combined expression can grow more slowly than $1/\phi(x)$ is if the inner integral vanishes at $x \to \infty$:

$$\int_0^\infty e^{\int_0^x \frac{v(x')dx'}{D(x')}} \left[\frac{B\phi^*(x)^3}{2D(x)}\left[\Upsilon(x) + 2D(x)(\partial_x \log \phi^*(x))^2\right] - \frac{\phi^*(x)^2}{D(x)}\right] dx = 0\,, \tag{S65}$$

which requires that

$$B = \frac{2\int \phi^*(x)\frac{\phi^*(x)}{D(x)}e^{\int_0^x \frac{v(x')dx'}{D(x')}}dx}{\int \left[\lambda(x)\phi^*(x)^2 + 2D(x)(\partial_x\phi^*(x))^2\right]\frac{\phi^*(x)}{D(x)}e^{\int_0^x \frac{v(x')dx'}{D(x')}}dx}\,. \tag{S66}$$

Plugging this back into our expression for $\phi(x,\tau)$, and recalling that

$$\phi^*(x) = \left[\int \phi(x,0)\rho_m(x,0)\right]u(x) \tag{S67}$$

and

$$\rho(x) = \frac{u(x)}{D(x)}e^{\int_0^x \frac{v(x')dx'}{D(x')}}\,, \tag{S68}$$

we find that the characteristic curve $\phi(x,\tau)$ reduces to

$$\phi(x,\tau) \sim \frac{2u(x)}{\Lambda\cdot\tau}\,, \tag{S69}$$

where $\Lambda$ is the effective rate of genetic drift defined in Eq. (7). The survival probability is therefore given by

$$p_s(t) \sim \frac{2\int u(x)\rho_m(x,0)dx}{\Lambda\cdot t} \tag{S70}$$

We can combine this result with the average lineage size above to obtain the average size conditioned on survival:

$$\langle\rho_m(x,t)\rangle = \langle\rho_m(x,t)|{>}0\rangle\cdot p_s(t) + 0\cdot[1-p_s(t)]\,, \tag{S71}$$

which yields

$$\langle\rho_m(x,t)|{>}0\rangle = \rho(x)\cdot\frac{\Lambda t}{2}\,, \tag{S72}$$

or

$$\langle f|f>0\rangle = \frac{\Lambda t}{2}\,. \tag{S73}$$

Note that the dependence on the initial density profile $\rho_m(x,0)$ cancels out, so that the conditional mean is independent of the initial conditions for times $t \gg \tau_{\text{drift}} \gg \tau_{\text{mix}}$. Similarly, the conditional variance is given by

$$\text{Var}(f|f{>}0) = \left(\frac{\Lambda t}{2}\right)^2\,. \tag{S74}$$

Motivated by these results, we can also solve for the full distribution of $f(t)$ conditioned on survival. The key insight will be to restrict our evaluation of the generating function regimes where

$$\phi(x,0) \sim \langle\rho_m(x,t)|>0\rangle^{-1} \ll (\Lambda\tau_{\text{mix}})^{-1}\,. \tag{S75}$$

Equations (S73) and (S74) show that this will be sufficient to capture the typical values of $\rho_m(x,t)$, though it will break down for smaller sizes where the lineage is close to extinction. By restricting our attention to this regime, we can again employ the short-time expansion that we used to derive the mean and variance in above. At intermediate times $\tau_{\text{mix}} \ll \tau \ll \Lambda^{-1}$, the generating function will quickly reduce to

$$\phi(x,\tau) \approx \left[\int \phi(x',0)\rho(x')dx'\right] u(x)\,. \tag{S76}$$

Thus, it will be sufficient to consider effective initial conditions of the form $\phi(x,0^+) = z_{\text{eff}}u(x)$. The next order correction yields

$$\phi(x,\tau) = z_{\text{eff}}u(x) - z_{\text{eff}}^2 u(x)\cdot\frac{\Lambda\tau}{2} + \mathcal{O}(\Lambda\tau)^2 = \phi(x,0^+) - \phi(x,0^+)^2\frac{\Lambda\tau}{2u(x)}\,, \tag{S77}$$

which can be recast as a coarse-grained differential equation for $\phi(x,t)$:

$$\frac{\partial\phi(x,\tau)}{\partial\tau} = -\frac{\Lambda}{2u(x)}\phi(x,\tau)^2\,, \quad (\tau \gg \tau_{\text{mix}}) \tag{S78}$$

After solving this equation, and applying the initial condition $\phi(x,0^+) = \int \phi(x',0)\rho(x')dx'$, we find that

$$\phi(x,\tau) = \frac{u(x)\int \phi(x',0)\rho(x')dx'}{1+\frac{\Lambda\tau}{2}\cdot\int \phi(x',0)\rho(x')dx}\,. \tag{S79}$$

The generating function for the total mutant frequency can be obtained by setting $\phi(x,0) = z/N$. This yields

$$H_f(z,t) \approx 1-\phi(x,t) = 1 - \frac{u(x)z}{1+\frac{\Lambda t z}{2}}\,, \tag{S80}$$

and a corresponding survival probability

$$p_s(t) \approx \frac{2u(x)}{\Lambda t}\,. \tag{S81}$$

The conditional generating function is therefore given by

$$H_f(z,t|f>0) = \frac{H_f(z,t)-(1-p_s(t))}{p_s(t)} = \frac{1}{1+\frac{\Lambda t}{2}\cdot z}\,, \tag{S82}$$

which is an exponential distribution with mean $\Lambda t/2$. These expressions are equivalent to a well-mixed Wright-Fisher model,

$$\frac{\partial f}{\partial t} = \sqrt{\Lambda f}\cdot\eta(t)\,, \tag{S83}$$

with an effective initial frequency $f(0) = u(x)$ and an effective rate of genetic drift $\Lambda$.

### 3. Effective generation time

We can use this result to calculate the effective generation time by focusing on the long-term substitution rate of neutral mutations. In principle, mutations can arise through a mixture of replication errors during cell division and DNA damage between divisions. We can encapsulate both of these processes within our model using a spatially varying mutation rate,

$$\mu(x) = \mu_b\cdot b(x) + \mu_0(x) \tag{S84}$$

where $b(x)$ is the local birth rate, $\mu_b$ is the mutation rate from cell division, and $\mu_0(x)$ is the background mutation rate from DNA damage. Once a mutation occurs, our previous analysis shows that it will survive to times $t \gg \tau_{\text{mix}}$ with probability

$$p_s(x,t) \sim \frac{u(x)}{\left[\int\left(\frac{\lambda(x)u(x)^2}{2} + D(x)(\partial_x u(x))^2\right)\rho(x)\,dx\right]\cdot t}\,. \tag{S85}$$

At this point, the mutant and wildtype populations will have equivalent spatial distributions (since $t \gg \tau_{\text{mix}}$). Symmetry considerations then imply the ultimate fixation probability of the neutral mutant must be proportional to its relative size, so that

$$p_{\text{fix}}(x) = p_s(x,t) \cdot \frac{\int \langle \rho_m(x',t)|{>}0\rangle dx'}{\int \rho(x')dx'} = u(x)\,. \tag{S86}$$

This allows us to identify $u(x)$ with the fixation probability of a neutral lineage, as previously noted by Ref. (37). The long-term substitution rate of these neutral mutations is therefore given by

$$R = \int \mu(x)u(x)\rho(x)\,dx\,. \tag{S87}$$

We can connect this to the effective generation time by focusing on the subset of division-marking mutations, for which $\mu(x) = \mu_b \cdot b(x)$, and

$$R_b = \mu_b \int b(x)u(x)\rho(x)\,dx\,. \tag{S88}$$

By definition, $R_b$ must also be equal to $\mu \cdot t/\tau_g$, where $\tau_g$ is the effective generation time. This yields an expression for $\tau_g$ in terms of an average over $u(x)$ and $\rho(x)$,

$$\frac{1}{\tau_g} = \lambda_e \equiv \int b(x)u(x)\rho(x)dx\,. \tag{S89}$$

Equation (9) in the main text follows from this expression by replacing the local birth rate with its equivalent definition, $b(x) = \lambda(x) + d(x)$. We can use this result to re-express the rates of genetic drift above in more traditional units of generations:

$$\text{Var}(f(t)) = \left[\frac{\int \left(\lambda(x')u(x')^2 + 2D(\partial_{x'}u)^2\right)\rho(x')dx'}{\int \lambda(x')u(x')\rho(x')dx'}\right] \cdot f \cdot \lambda_e t \tag{S90}$$

The quantity in braces can then be identified as an "effective population size"

$$N_e^{-1} = \frac{\int \left(\lambda(x')u(x')^2 + 2D(\partial_{x'}u)^2\right)\rho(x')dx'}{\int \lambda(x')u(x')\rho(x')dx'}\,, \tag{S91}$$

which is equivalent to Eq. (10) in the main text. These definitions allow us to rewrite our earlier results in a particularly compact form:

$$\text{Var}(f(t)) \approx \frac{f \cdot \lambda_e t}{N_e}\,, \tag{S92}$$

$$p_s(x,t) = \frac{N_e u(x)}{\lambda_e t}\,, \tag{S93}$$

$$\langle \rho_m(x,t)|{>}0\rangle = \rho(x)\left(\frac{\lambda_e t}{N_e}\right)\,. \tag{S94}$$

Together, these results show that on timescales much longer than $\tau_{\text{mix}}$, the dynamics of neutral lineages resemble a Wright-Fisher model with effective generation time $\lambda_e^{-1}$ and effective population size $N_e$.

## 3. DYNAMICS OF A SELECTED LINEAGE

We can use a similar set of approaches to characterize the dynamics of selected mutations. In this case, the characteristic curve in the generating function is now given by

$$\partial_\tau \phi(x,\tau) = D(x)\partial_x^2\phi(x,\tau) + v(x)\partial_x\phi(x,\tau) + \lambda_m(x)\phi(x,\tau) - \frac{\Upsilon_m(x)\phi(x)^2}{2}\phi(x,\tau)^2 - D(x)[\partial_x\phi(x,\tau)]^2\,, \tag{S95}$$

where $\lambda_m(x)$ is the position-dependent growth rate of the mutant strain. If $\lambda_m(x)$ is sufficiently close to the wildtype growth rate $\lambda(x)$, then the dynamics of $\phi(x,t)$ on timescales $\tau_{\text{mix}} \ll \tau \ll \mathcal{O}(1/s_e)$ can be derived using the same short-time perturbation approach as above. The first order contribution is given by

$$\phi(x,\tau) = \left[\int \phi(x',0)\rho(x')dx'\right] u(x) + \left[\int \phi(x',0)\rho(x')dx'\right] \left[\int [\lambda_m(x') - \lambda(x')]u(x')\rho(x')dx'\right] u(x) \cdot \tau \,. \tag{S96}$$

The average mutant density profile is therefore given by

$$\langle \rho_m(x,t)\rangle = \left[\int u(x')\rho_m(x',0)dx'\right] \rho(x)\,(1 + s_e\lambda_e t)\ , \tag{S97}$$

where

$$s_e \equiv \frac{1}{\lambda_e}\int [\lambda_m(x) - \lambda(x)]\, u(x)\rho(x)dx \tag{S98}$$

is the effective selection coefficient defined in Eq. (12) in the main text. Iterating over multiple short time windows then yields

$$\langle \rho_m(x,t)\rangle = \left[\int u(x')\rho_m(x',0)dx'\right] \rho(x)e^{s_e\lambda_e\cdot t}\,. \tag{S99}$$

For a constant growth rate advantage, $\lambda_m(x) = (1+s)\lambda(x)$, we find that

$$s_e = s \cdot \frac{\int \lambda(x)u(x)\rho(x)\,dx}{\int [\lambda(x) + d(x)]u(x)\rho(x)\,dx}\,. \tag{S100}$$

This shows that the effective selection strength is proportional to $s$ as expected. When cell death is negligible in the regions that dominate $\lambda(x)u(x)\rho(x)$, this further reduces to $s_e = s$.

### 1. Establishment probability

In contrast to the neutral scenario above, the characteristic curve for a beneficial mutation must eventually saturate at a nonzero fixed point:

$$\phi(x,\tau) \to w(x)\,, \tag{S101}$$

described by the ordinary differential equation in Eq. (13). For small growth rate differences, we can solve for the fixed point $w(x)$ using a perturbation expansion that is structurally similar to the one we used in the neutral case above. We first introduce a perturbation scale $\epsilon$ such that

$$\begin{aligned} \lambda_m(x) &= \lambda(x) + \epsilon\xi(x)\,, \\ \Upsilon_m(x) &= \Upsilon(x) + \epsilon\zeta(x)\,, \end{aligned} \tag{S102}$$

with $\epsilon \ll 1$. We then rewrite $w(x)$ as

$$w(x) = \epsilon B u(x) f(\epsilon) g(x,\epsilon)\,, \tag{S103}$$

where $u(x)$ is the linearized neutral fixed point above, and $f(\epsilon)$ and $g(x,\epsilon)$ are unknown functions that satisfy

$$f(0) = 1\,, \quad g(0,\epsilon) = 1\,, \quad \partial_x g(0,\epsilon) = 0\,, \quad \lim_{x\to\infty} g(x,\epsilon) = 0\,. \tag{S104}$$

Substituting these expressions into Eq. (13) in the main text, we obtain an analogous differential equation for $g(x,\epsilon)$:

$$\begin{aligned} 0 = {} & \epsilon\xi(x)g(x,\epsilon) + D(x)\partial_x^2 g(x,\epsilon) + [v(x) + 2D(x)\partial_x \log u]\,\partial_x g(x,\epsilon) \\ & - \frac{\epsilon B f(\epsilon)u(x)g(x,\epsilon)}{2}\left(\Upsilon(x) + \epsilon\zeta(x) + 2D(x)[\partial_x \log u(x) + \partial_x \log g(x,\epsilon)]^2\right)\ , \end{aligned} \tag{S105}$$

subject to the boundary conditions

$$g(0,\epsilon) = 1\,, \quad \partial_x g(0,\epsilon) = 0\,, \quad \lim_{x\to\infty} u(x) g(x,\epsilon) = 0\,. \tag{S106}$$

We then assume an analogous set of asymptotic expansions:

$$\begin{aligned} f(\epsilon) &= 1 + f_1 \cdot \epsilon + \mathcal{O}(\epsilon^2)\,, \\ g(x,\epsilon) &= g_0 + g_1(x)\epsilon + \mathcal{O}(\epsilon^2)\,, \end{aligned} \tag{S107}$$

and substitute these expressions into Eq. (S105). At zeroth order in $\epsilon$, we find that

$$0 = D(x)\partial_x^2 g_0(x,\tau) + [v(x) + 2D(x)\partial_x \log u(x)]\,\partial_x u(x,\tau)\,, \tag{S108}$$

and hence $g_0 = 1$ as above. At first order in $\epsilon$, we have

$$\begin{aligned} 0 = \xi(x) + D(x)\partial_x^2 g_1(x) + [v(x) + 2D(x)\partial_x \log u(x)]\,\partial_x g_1(x) \\ - \frac{B}{2u(x)}\left(\Upsilon(x)u(x)^2 + 2D(x)[\partial_x u(x)]^2\right)\,. \end{aligned} \tag{S109}$$

This is a similar first order equation as in the neutral case, and the solution is given by

$$g_1(x) = \int_0^x \frac{e^{-\int_0^{x'} \frac{v(x'')dx''}{D(x'')}}}{\phi^*(x')^2} \int_0^{x'} e^{\int_0^{x''} \frac{v(x''')dx'''}{D(x''')}} \left[\frac{B\phi^*(x'')^3}{2D(x'')}\left[\lambda(x'') + 2D(x'')(\partial_x \log\phi^*)^2\right] - \frac{\phi^*(x'')^2}{D(x'')}\xi(x'')\right] dx''\,. \tag{S110}$$

As above, the finite boundary condition at $x = \infty$ requires that $B$ is chosen to satisfy

$$B = \frac{2\int \xi(x)u(x)\rho(x)}{\int[\Upsilon(x)u(x)^2 + 2D(x)(\partial_x u)^2]\rho(x)\,dx}\,. \tag{S111}$$

This implies that

$$w(x) = 2N_e s_e u(x) + \epsilon^2\,, \tag{S112}$$

as expected from Eq. (11).

We can test the limits of this perturbation expansion by turning to a numerical solution of Eq. (13). In this case, it is helpful to work in logarithmic space by defining

$$\gamma(x) \equiv \log\left[\frac{w(x)}{w(0)}\right]\,. \tag{S113}$$

This function satisfies the related differential equation

$$0 = D(x)\partial_x^2\gamma(x) + D(x)(\partial_x\gamma(x))^2 + v(x)\partial_x\gamma(x) + \lambda_m(x) - \frac{\Upsilon_m(x)w(0)e^{\gamma(x)}}{2} - D(x)w(0)e^{\gamma(x)}[\partial_x\gamma(x)]^2\,, \tag{S114}$$

with the initial conditions

$$\gamma(0) = 0 \quad \gamma'(0) = 0\,. \tag{S115}$$

We will often be interested in scenarios where there is a critical position $x = x_0$ above which $\lambda_m(x)$ and $\Upsilon_m(x)$ vanish, and where $v(x)$ and $D(x)$ eventually attain constant values, $v(x) = v$ and $D(x) = D$. When this occurs, the solution to Eq. (S114) for $x > \ell$ is given by

$$w(x) = -\log\left[1 - \frac{D\partial_x w(x_0)}{v}\left(1 - e^{-v(x-x_0)/D}\right)\right] + w(x_0)\,, \tag{S116}$$

where the initial conditions $w(x_0)$ and $\partial_x w(x_0)$ must be chosen to ensure that $w(x) \to 0$ as $x \to \infty$. By examining the form of Eq. (S116), we see that the only way that this can happen is if $w(x_0)$ and $\partial_x w(x_0)$ are related by

$$w(x_0) = \log\left[1 - \frac{D\partial_x w(x_0)}{v}\right]\,, \tag{S117}$$

which implies that

$$w(x) = -\log\left[1 - \left(1 - e^{-w(x_0)}\right) e^{-v(x-x_0)/D}\right] . \tag{S118}$$

In terms of the logarithm $\gamma(x)$, these expressions become

$$\partial_x \gamma(x_0) + \frac{v}{D}\left[\frac{e^{w(0)e^{\gamma(x_0)}} - 1}{w(0)e^{\gamma(x_0)}}\right] = 0 , \tag{S119}$$

and

$$\gamma(x) = \log\left[-\log\left[1 - \left(1 - e^{-w(x_0)}\right) e^{-v(x-x_0)/D}\right] w(x_0)^{-1}\right] , \tag{S120}$$

respectively. These expressions allow us to obtain numerical solutions for the fixation profile by first solving for $w(x)$ over the finite region $0 \leq x \leq x_0$ and then enforcing continuity with this analytical solution for the large-$x$ limit at $x = x_0$. We used this approach to calculate the large-$s_e$ behavior of the fixation probability in Figs 3 and 4, which is described in more detail in Supplementary Information 4.4 and 5 below.

### 2. Long-term growth rate of strongly beneficial mutations

We can use a related approach to investigate the large-$s_e$ behavior of the average mutant density profile in Eq. (S26). We assume that the long-time behavior can be written in the form

$$\langle \rho_m(x,t) = e^{St}\psi(x) , \tag{S121}$$

where $S$ is an effective exponential growth rate and $\psi(x)$ is an arbitrary spatial profile. Substituting this ansatz into Eq. (S26), we find that $\psi(x)$ must satisfy the related differential equation

$$0 = \partial_x^2[D(x)\psi(x)] - \partial_x[v(x)\psi(x)] + [\lambda_m(x) - S]\psi , \tag{S122}$$

and is subject to the same boundary conditions as $\langle \rho_m(x,t)\rangle$:

$$\partial_x \psi(0) = \frac{v}{D}\psi(0) , \qquad \partial_x \psi(L) = 0 . \tag{S123}$$

When $\lambda_m(x)$ is close to $\lambda(x)$, we can use a perturbative solution of Eq. (S122) to obtain an alternate derivation of the effective selection coefficient in Eq. (12) in the main text.

To do so, we first note that the duality between $\rho(x)$ and $u(x)$ holds for $\psi(x)$ as well. This motivates us to define a function

$$\overline{\phi}(x) = D(x)\psi(x)e^{-\int_0^x \frac{v(x')}{D(x')}dx'} , \tag{S124}$$

which satisfies the dual equation

$$0 = D(x)\partial_x^2\overline{\phi}(x) + v(x)\partial_x\overline{\phi}(x) + [\lambda_m(x) - S]\,\overline{\phi}(x) . \tag{S125}$$

This is similar to Eq. (13) in the main text, except with $S\overline{\phi}(x)$ replacing the $\lambda_m(x)w(x)^2/2 + \partial_x w(x)$ term. Thus, by repeating our perturbative calculation for $w(x)$ above, we conclude that $S$ must be chosen to satisfy

$$S = \frac{\int[\lambda_m(x) - \lambda(x)]u(x)\rho(x)\,dx}{\int u(x)\rho(x)\,dx} = s_e\lambda_e . \tag{S126}$$

in agreement with the separation-of-timescales result in Eq. (12).

The advantage of Eq. (S122) is that it can still be used to solve for the exponential growth rate even when $S\tau_{\text{mix}} \gg 1$. We carry out this calculation for the minimal model in Fig. 3 in Supplementary Information 4.5 below.

## 4. ANALYTICAL SOLUTIONS FOR A MINIMAL MODEL OF BACTERIAL GROWTH

We can gain additional intuition for our results by considering a concrete model of microbial growth in the gut, where many of the important quantities can be determined analytically. In this simplified model, we assume that the flow velocity and mixing rates are spatially uniform, so that

$$v(x) \equiv v\,, \quad D(x) \equiv D\,. \tag{S127}$$

We also assume that the gut is long enough that the total length $L$ can be regarded as effectively infinite ($vL/D \to \infty$). We neglect cell death, and assume that microbial growth depends on a single nutrient that is supplied at the entrance of the colon, which is consumed by cells during division. We assume that the growth rate has a step-like dependence on the local nutrient concentration,

$$\lambda(c) = \begin{cases} \lambda & \text{if } c > K \\ 0 & \text{else} \end{cases} \tag{S128}$$

with a maximum growth rate $\lambda$ and saturation constant $K$. This function is qualitatively similar to the Monod-like models that have been employed in previous work (28, 65), but with a sharper transition between the no-growth and saturated-growth-rate regimes. This will turn out to be a crucial simplification that will allow us to obtain analytical solutions for the growth rate and density profiles that are attained at steady state.

The assumptions of our model imply that the local nutrient concentration must be a monotonically decreasing function of $x$, since the nutrients are progressively consumed as they flow down the colon. The step-like growth function then implies that the local growth profile $\lambda(x)$ must also develop a step-like form,

$$\lambda(x) = \begin{cases} \lambda & \text{if } x < \ell, \\ 0 & \text{else,} \end{cases} \tag{S129}$$

which transitions from high-growth to no-growth at a critical position $x = \ell$. The location of this transition is determined by the point where the local nutrient concentration first drops below $c(x) = K$. This depends in turn on the local density profile $\rho(x)$, which dictates the total consumption of resources between $x = 0$ and $x = \ell$.

An important feature of our minimal model is that it allows us to side-step this coupled calculation of $n(x)$ and $\rho(x)$. To see this, we note that the solutions for the steady-state density profile in Eq. (S5) will have two degrees of freedom, one of which is fixed by the overall normalization condition,

$$N = \int \rho(x)\, dx\,. \tag{S130}$$

Since there are two independent boundary conditions for $\rho(x)$ — and only one remaining degree of freedom — we see that $\ell$ can be uniquely determined by the requirement that Eq. (S5) admits a non-trivial solution. Furthermore, the solution to this "eigenvalue problem" also determines the shape of the density profile $\rho(x)$ up to an overall normalization factor, which is fixed by Eq. (S130). This shows that the explicit dependence on the nutrient profile is completely encapsulated by the population size $N$, which will vary as a function of the nutrient input flux and the saturation constant $K$. However, for a fixed value of $N$, the steady-state growth and density profiles are completely determined by Eq. (S5). We adopt this perspective in our calculations below, treating $N$ as an independent parameter that must be determined experimentally.

We note that this same argument applies for any model that generates a single parameter family of growth profiles like Eq. (S129). This suggests that Eq. (S129) may be considerably more general than the specific growth model proposed above, and can also be viewed as a purely phenomenological approximation to the empirical growth profiles in Fig. 4.

### 1. Steady state density profile

To solve for the steady-state density profile, it will be convenient to switch to diffusion units by defining the rescaled variables:

$$z \equiv \frac{vx}{2D}\,, \quad \tilde{\ell} \equiv \frac{v\ell}{2D}\,, \quad \tilde{L} \equiv= \frac{vL}{2D}\,. \tag{S131}$$

Equation (S5) can then be written in the dimensionless form

$$0 = \partial_z^2 \rho(z) - 2\partial_z \rho(z) + \alpha\theta(\tilde{\ell} - z)\rho(z) \tag{S132}$$

with $\alpha$ defined by

$$\alpha \equiv \frac{4D\lambda}{v^2}\,. \tag{S133}$$

Equation (S132) admits a simple piecewise solution for $z < \tilde{\ell}$ and $z > \tilde{\ell}$, respectively. In the bulk of the colon ($z > \ell$), the density profile satisfies

$$\partial_z^2 \rho - 2\partial_z \rho = 0\,, \tag{S134}$$

which has the general solution

$$\rho(z) = Ce^{2z} + D\,, \tag{S135}$$

where $C$ and $D$ are integration constants. The finite boundary condition at $z = \tilde{L} \to \infty$ requires that $C = 0$, so that $\rho(z)$ is simply a constant in this region.

On the other hand, when $z < \tilde{\ell}$, the density profile instead satisfies

$$\partial_z^2 \rho - 2\partial_z \rho + \alpha\rho = 0 \tag{S136}$$

which has the general solution

$$\rho(z) = Ae^z \left[\cos(kz) + b\sin(kz)\right] \tag{S137}$$

where

$$k \equiv \sqrt{\alpha - 1} \equiv \sqrt{\frac{4D\lambda}{v^2} - 1} \tag{S138}$$

and $A$ and $b$ are a pair of integration constants. Applying the boundary condition at $z = 0$ yields

$$0 = -\partial_z \rho(0) + 2\rho(0) = -\left[1 + bk\right] + 2\,, \tag{S139}$$

so that $b = 1/k$ and

$$\rho(z) = Ae^z \left[\cos(kz) + \frac{1}{k}\sin(kz)\right]\,. \tag{S140}$$

Finally, enforcing continuity at $z = \tilde{\ell}$ allows us to solve for $D$ in terms of $A$:

$$D = Ae^{\tilde{\ell}} \left[\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})\right] \tag{S141}$$

while enforcing continuity of the derivative at $\tilde{\ell}$ yields an "eigenvalue" equation for $\tilde{\ell}$:

$$\left[\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})\right] + \left[-k\sin(k\tilde{\ell}) + \cos(k\tilde{\ell})\right] = 0 \tag{S142}$$

This can be rewritten in the compact form

$$\tan\left(k\tilde{\ell}\right) = \frac{2\sqrt{\alpha - 1}}{\alpha - 2} \tag{S143}$$

whose solution is given by

$$\tilde{\ell} = \frac{1}{\sqrt{\alpha - 1}} \tan^{-1}\left(\frac{2\sqrt{\alpha - 1}}{\alpha - 2}\right) \sim \begin{cases} \frac{2}{\alpha} & \text{if } \alpha \gg 1 \\ \frac{\pi}{\sqrt{\alpha - 1}} - 2 & \text{if } \alpha - 1 \ll 1 \end{cases} \tag{S144}$$

Given this value of $\tilde{\ell}$, the full piecewise solution for the density profile can be written as

$$\rho(z) = \begin{cases} A_\rho \left[\cos(kz) + \frac{1}{k}\sin(kz)\right] e^{z} & \text{if } z \le \tilde{\ell} \\ A_\rho \left[\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})\right] e^{\tilde{\ell}} & \text{if } z > \tilde{\ell} \end{cases} \tag{S145}$$

where the normalization constant $A_\rho$ is determined by the overall population size $\int \rho(z)\,dz = N$. Equation (6) then implies that the neutral fixation probability $u(x)$ is given by

$$u(z) = \begin{cases} A_u \left[\cos(kz) + \frac{1}{k}\sin(kz)\right] e^{-z} & \text{if } z \le \tilde{\ell} \\ A_u \left[\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})\right] e^{-2(z-\tilde{\ell})-\tilde{\ell}} & \text{if } z > \tilde{\ell} \end{cases} \tag{S146}$$

where the normalization constant $A_u$ is fixed by the condition that $\int u(z)\rho(z)\,dz = 1$. The ancestral distribution $g(z) = u(z)\rho(z)$ then takes on the simple form

$$g(z) = \begin{cases} A_u A_\rho \left[\cos(kz) + \frac{1}{k}\sin(kz)\right]^2 & \text{if } z \le \tilde{\ell} \\ A_u A_\rho \left[\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})\right]^2 e^{-2(z-\tilde{\ell})} & \text{if } z > \tilde{\ell} \end{cases} \tag{S147}$$

which is independent of $N$ as expected and normalized so that $\int g(z)dz = 1$.

### 2. Effective generation time and population size

To calculate the effective generation time and population size, it will be helpful to first obtain explicit expressions for the normalization factors $A_\rho$ and $A_u$ above. In the limit that the length of the colon is very long ($\tilde{L} \gg \tilde{\ell}, 1$), the total population size will be dominated by slowly growing regions ($z > \tilde{\ell}$), so that

$$N = \int \rho(z)\,dz \approx A_\rho \left[\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})\right] e^{\tilde{\ell}}\tilde{L}\,, \tag{S148}$$

or

$$A_\rho \approx \frac{N}{\left[\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})\right] e^{\tilde{\ell}}\tilde{L}}\,. \tag{S149}$$

Similarly, the normalization condition $\int g(z)dz = 1$ implies that

$$(A_\rho A_u)^{-1} = \underbrace{\int_0^{\tilde{\ell}} \left[\cos(kz) + \frac{1}{k}\sin(kz)\right]^2 dz}_{I_1} + \underbrace{\int_{\tilde{\ell}}^{\infty} \left[\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})\right]^2 e^{-2(z-\tilde{\ell})}\,dz}_{I_2}\,, \tag{S150}$$

which we have split into contributions from the integrals $I_1$ and $I_2$. The latter is given by

$$I_2 = \frac{1}{2}\left[\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})\right]^2\,, \tag{S151}$$

while the former evaluates to

$$I_1 = \frac{k^2\tilde{\ell} + \tilde{\ell} + 1}{2k^2} + \frac{k^2-1}{4k^3}\sin(2k\tilde{\ell}) - \frac{1}{2k^2}\cos(2k\tilde{\ell})\,. \tag{S152}$$

Using these expressions, we can write the effective generation time in the form

$$\frac{\lambda_e}{\lambda} = \frac{1}{\lambda}\int \lambda(z)u(z)\rho(z)\,dz = \int_0^{\tilde{\ell}} u(z)\rho(z)\,dz = \frac{I_1}{I_1 + I_2}\,, \tag{S153}$$

where $I_1$ and $I_2$ are defined as above. This yields an exact expression,

$$\frac{\lambda_e}{\lambda} = \frac{4k(k^2\tilde{\ell} + \tilde{\ell} + 1) + 2(k^2-1)\sin(2k\tilde{\ell}) - 4k\cos(2k\tilde{\ell})}{[4k^3\tilde{\ell} + k^2 + 4k(\tilde{\ell}+1) + 1] + 2[k^2 + k - 1]\sin(2k\tilde{\ell}) + [k^2 - 4k - 1]\cos(2k\tilde{\ell})}\,, \tag{S154}$$

whose large and small $\alpha$ limits are given by

$$\frac{\lambda_e}{\lambda} \sim \begin{cases} \frac{4}{\alpha} & \text{if } \alpha \gg 1, \\ 1 & \text{if } \alpha - 1 \ll 1. \end{cases} \tag{S155}$$

This asymptotic behavior can also be obtained by considering the $I_1$ and $I_2$ integrals directly. When $\alpha \gg 1$ is large, the sine and cosine terms sum to $\approx 1$ throughout the entire interval $0 \leq z \leq \tilde{\ell}$. This implies that $I_1 \approx \tilde{\ell} \ll 1$ and $I_2 \approx 1/2$ and hence

$$\frac{\lambda_e}{\lambda} \approx \frac{\tilde{\ell}}{\tilde{\ell} + \frac{1}{2}} \approx \frac{4}{\alpha}\,, \qquad (\alpha \gg 1)\,. \tag{S156}$$

In the opposite extreme ($\alpha - 1 \ll 1$, $k \ll 1$, and $\tilde{\ell} \gg 1$), the sine term dominates over the cosine term through the bulk of the interval $0 \leq z \leq \tilde{\ell}$, so that

$$I_1 \approx \frac{1}{k^2} \int_0^{\tilde{\ell}} \sin(k\tilde{\ell})^2 dz = \frac{1}{k^3} \int_0^{\pi} \sin^2 u \, du = \frac{\pi}{2k^3} \tag{S157}$$

Meanwhile, Eq. (S142) shows that the sine and cosine terms sum to $\approx 1$ at $z = \tilde{\ell}$, so that $I_2 = \frac{1}{2} \ll I_1$, and

$$\frac{\lambda_e}{\lambda} \approx \frac{\frac{\pi}{2k^3}}{\frac{\pi}{2k^3} + \frac{1}{2}} \approx 1\,, \quad (\alpha - 1 \ll 1)\,, \tag{S158}$$

as desired.

We can use a similar strategy to evaluate the effective population size in Eq. (10). We begin by rewriting this expression in the form

$$\left(\frac{N_e}{N}\right)^{-1} = \frac{1}{\lambda_e} \left[\int \rho(z)\, dz\right] \int \left[\lambda(z) u(z)^2 + \frac{2\lambda}{\alpha} (\partial_z u)^2\right] \rho(z)\, dz\,, \tag{S159}$$

$$= \left(\frac{\lambda}{\lambda_e}\right) \left(\frac{\int \rho(z) dz}{A_\rho}\right) (A_\rho A_u)^2 \left[\underbrace{\int_0^{\tilde{\ell}} \frac{u(z)^2 \rho(z)\, dz}{A_\rho A_u^2}}_{I_3} + \underbrace{\int_0^{\tilde{\ell}} \frac{2(\partial_z u)^2 \rho(z)\, dz}{\alpha A_\rho A_u^2}}_{I_4} + \underbrace{\int_{\tilde{\ell}}^{\infty} \frac{2(\partial_z u)^2 \rho(z)\, dz}{\alpha A_\rho A_u^2}}_{I_5}\right]\,, \tag{S160}$$

which we have split into contributions from the integrals $I_3$, $I_4$, and $I_5$. To evaluate these expressions, it is helpful to use Eq. (S146) to show that

$$\partial_z u = \begin{cases} -A_u \left(\frac{1+k}{k}\right) \sin(kz) e^{-z} & \text{if } z \leq \tilde{\ell}, \\ -2A_u \left[\cos(k\tilde{\ell}) + \frac{1}{k} \sin(k\tilde{\ell})\right] e^{-2(z-\tilde{\ell})-\tilde{\ell}} & \text{if } z > \tilde{\ell} \end{cases} \tag{S161}$$

This implies that $I_5$ is given by

$$I_5 = \frac{2}{\alpha} e^{-\tilde{\ell}} \left[\cos(k\tilde{\ell}) + \frac{1}{k} \sin(k\tilde{\ell})\right]^3\,, \tag{S162}$$

while $I_3$ evaluates to

$$I_3 = \frac{e^{-\tilde{\ell}}}{36k^5 + 4k^3} \Big[\sin(k\tilde{\ell})[27k^4 - 24k^2 - 3] + \sin(3k\tilde{\ell})[3k^4 - 12k^2 + 1] \tag{S163}$$

$$- \cos(k\tilde{\ell})[54k^3 + 6k] - \cos(3k\tilde{\ell})[10k^3 - 6k]\Big] + \frac{16}{9k^2 + 1}\,, \tag{S164}$$

and $I_4$ is given by

$$I_4 = \frac{2}{\alpha} \left(\frac{k^2 + 1}{k}\right)^2 \Bigg[\frac{e^{-\tilde{\ell}}}{36k^5 + 40k^3 + 4k} \Big[\sin(k\tilde{\ell})[9k^4 + 26k^2 - 3] - \sin(3k\tilde{\ell})[3k^4 + 2k^2 - 1] \tag{S165}$$

$$- \cos(k\tilde{\ell})[36k^3 + 4k] + \cos(3k\tilde{\ell})[4k^3 + 4k]\Big] + \frac{8k^2}{9k^4 + 10k^2 + 1}\Bigg]\,. \tag{S166}$$

Putting everything together, we have

$$\frac{N_e}{N}\cdot\frac{vL}{2D}=\frac{e^{-\tilde{\ell}}}{\cos(k\tilde{\ell})+\frac{1}{k}\sin(k\tilde{\ell})}\cdot\frac{I_1(I_1+I_2)}{I_3+I_4+I_5}\,, \tag{S167}$$

whose large and small $\alpha$ limits are given by

$$\frac{N_e}{N}\cdot\frac{vL}{2D}\sim\begin{cases}\frac{1}{4} & \text{if } \alpha\gg 1,\\ \frac{\pi^2e^2}{2^7}(\alpha-1)^{-3}e^{\frac{-\pi}{\sqrt{\alpha-1}}} & \text{if } \alpha-1\ll 1.\end{cases} \tag{S168}$$

As above, these large and small $\alpha$ limits can also be obtained by considering the individual integrals directly. When $\alpha\gg 1$ the sine and cosine terms sum to $\approx 1$ (and $\tilde{\ell}\approx 2/\alpha\ll 1$), so

$$I_3\approx\int_0^{\tilde{\ell}}e^{-z}\,dz\approx\frac{2}{\alpha}\,, \tag{S169}$$

and

$$I_4\approx\frac{1}{\alpha}\int_0^{\tilde{\ell}}e^{-z}dz\approx\frac{2}{\alpha^2}\ll I_3\,, \tag{S170}$$

while $I_5$ evaluates to

$$I_5\approx\int_0^{\infty}\frac{8}{\alpha}e^{-4z}\,dz=\frac{2}{\alpha}\,. \tag{S171}$$

This implies that growth below $\tilde{\ell}$ (represented by $I_3$) and diffusion above $\tilde{\ell}$ (represented by $I_5$) make equal contributions to the magnitude of genetic drift. Combining this with our asymptotic expressions for $I_1$ and $I_2$ above, we find that

$$\frac{N_e}{N}\cdot\frac{vL}{2D}\approx\approx\frac{\frac{2}{\alpha}\cdot\frac{1}{2}}{\frac{2}{\alpha}+\frac{2}{\alpha}}\approx\frac{1}{4}\,,\qquad(\alpha\gg 1)\,. \tag{S172}$$

This shows that while the wall-clock rate of genetic drift ($\Lambda$) declines for large $\alpha$, the longer generation times in this regime end up compensating for this effect to yield a constant value of $N_e/N$.

In the opposite extreme ($\alpha-1\ll 1$, $k\ll 1$, and $\tilde{\ell}\gg 1$), the $I_3$ integral will be dominated by a narrow region where $z\sim\mathcal{O}(1)$, so that

$$I_3=\int_0^{\tilde{\ell}}[\cos(kz)+\frac{1}{k}\sin(kz)]^3e^{-z}\,dz\approx\int_0^{\infty}(1+z)^3e^{-z}\,dz=16\,. \tag{S173}$$

and

$$I_4=\frac{2}{\alpha}\int_0^{\tilde{\ell}}\left(\frac{1+k}{k}\sin(kz)\right)^2[\cos(kz)+\frac{1}{k}\sin(kz)]ze^{-z}\,dz\approx\int_0^{\infty}2z^2(1+z)e^{-z}\,dz=16\,. \tag{S174}$$

Finally, since $\cos(k\tilde{\ell})+\frac{1}{k}\sin(k\tilde{\ell})\approx 1$, the contribution from $I_5$ becomes

$$I_5\approx 2e^{-\tilde{\ell}}\ll I_3\,. \tag{S175}$$

Combined with our expressions for $I_1$ and $I_2$ above, this yields,

$$\frac{N_e}{N}\cdot\frac{vL}{2D}\approx e^{-\tilde{\ell}}\frac{\left(\frac{\pi}{2k^3}\right)^2}{2\cdot 16}\approx\frac{\pi^2e^2}{2^7}(\alpha-1)^{-3}e^{-\frac{\pi}{\sqrt{\alpha-1}}}\,,\qquad(\alpha-1\ll 1)\,. \tag{S176}$$

This shows that in the $\alpha-1\ll 1$ regime, genetic drift is dominated by a narrow slice of the ancestral distribution with $z\sim\mathcal{O}(1)\ll\tilde{\ell}$. Since $\rho(z)$ is exponentially increasing in this regime, the bacterial density in this narrow slice of $z$ values is also smaller than $\rho(\ell)$ by an additional factor of $\sim\exp(-\tilde{\ell})$. This leads to an effective population size that is significantly smaller than the total number of dividing cells, which scales like

$$\frac{N_d}{N}\cdot\frac{vL}{2D}=\tilde{L}\int_0^{\tilde{\ell}}A_\rho\left[\cos(kz)+\frac{1}{k}\sin(kz)\right]e^z\,dz\approx 1\,,\qquad(\alpha-1\ll 1)\,. \tag{S177}$$

### 3. Extension to multiple growth regions

To better understand the mechanisms responsible for driving the enhanced rates of genetic drift near the washout threshold ($\alpha \to 1$), we considered a generalization of our minimal model that includes a washed out region at the proximal end of the colon. We reasoned that such a model would allow us to determine whether washout itself ($\alpha < 1$) — vs an extended region of near-washout ($\alpha \to 1$) growth — is the primary cause of the behavior in Fig. 2F. For simplicity, we assumed that the washed-out region extends all the way from $x = 0$ to $x = -\infty$, so that it would have the largest possible impact on the emergent evolutionary parameters. Our generalized model assumes that the growth and flow velocity profiles can be written the piecewise form,

$$\lambda(x) = \begin{cases} \lambda_0 & \text{if } x < 0, \\ \lambda & \text{if } 0 \leq x \leq \ell, \\ 0 & \text{if } x > \ell, \end{cases} \tag{S178}$$

and

$$v(x) = \begin{cases} v_0 & \text{if } x \leq 0, \\ v & \text{if } x > 0, \end{cases} \tag{S179}$$

where $\lambda_0$ and $v_0$ represent the growth and flow parameters in the washed out region.

To solve for the steady-state population density profile, it is helpful to adopt the same units as in Eq. (S131) above, where all lengths are measured relative to the mixing lengthscale in the original the $x > 0$ region. The equation for the steady-state density profile in the new region can then be written in the dimensionless form

$$0 = \beta_0^2 \rho(z) - 2\beta_0 \partial_z \rho(z) + \alpha_0 \rho(z)\,, \tag{S180}$$

with $\alpha_0$ and $\beta_0$ defined by

$$\alpha_0 \equiv \frac{4D\lambda_0}{v_0^2}\,, \qquad \beta_0 \equiv \frac{v_0}{v}\,. \tag{S181}$$

Here, we will be primarily interested in parameter combinations that prevent the population from fully establishing in the $x < 0$ region ($\alpha_0 < 1$) and where the flow velocity is at least as large as in the $x > 0$ region ($v_0 \gtrsim v$).

The general solution to Eq. (S180) is given by

$$\rho(z) = A_0 e^{\frac{v_0 z}{v}} \left[ e^{k_0 \beta_0 z} + b_0 e^{-k_0 \beta_0 z} \right]\,, \tag{S182}$$

where $A_0$ and $b_0$ are integration constants, and

$$k_0 \equiv \sqrt{1 - \alpha_0}\,. \tag{S183}$$

The reflecting boundary condition at $z = -\infty$ requires that $b = 0$, so that

$$\rho(z) = A_0 e^{\beta_0 (1 + k_0) z}\,. \tag{S184}$$

By matching this solution to our previous solution for $\rho(z)$ in the $z > 0$ region,

$$\rho(z) = A_\rho e^z \left[ \cos(kz) + b \sin(kz) \right]\,, \tag{S185}$$

we find that

$$A_0 = A_\rho \tag{S186}$$

The derivative condition is slightly more complicated in this case, since $v(x)$ changes abruptly at $x = 0$. In this case, the conservation of particle flux across the boundary yields a generalized continuity condition,

$$D\partial_x \rho(0)^- - v_0 \rho(0)^- = D\partial_x \rho(0)^+ - v\rho(0)^+\,, \tag{S187}$$

which, in dimensionless form, becomes

$$\partial_z \rho(0)^- - 2\beta_0 \rho(0)^- = \partial_z \rho(0)^+ - 2\rho(0)^+\,. \tag{S188}$$

By applying this condition to Eqs. (S184) and (S185), we find that

$$b = \frac{1+\beta_0(k_0-1)}{k}\,. \tag{S189}$$

Finally, the continuity conditions at $z=\tilde{\ell}$ yield an analogous eigenvalue equation for $\tilde{\ell}$:

$$\cos(k\tilde{\ell}) + \left(\frac{1+\beta_0(k_0-1)}{k}\right)\sin(k\tilde{\ell}) - k\sin(k\tilde{\ell}) + (1+\beta_0(k_0-1))\cos(k\tilde{\ell}) = 0\,, \tag{S190}$$

which can be written in the compact form,

$$\tan(k\tilde{\ell}) = \frac{2\sqrt{\alpha-1}\left[1-\frac{\beta_0(1-\sqrt{1-\alpha_0})}{2}\right]}{\alpha^2-2+\beta_0(1-\sqrt{1-\alpha_0})}\,. \tag{S191}$$

The solution to this equation,

$$\tilde{\ell} = \frac{1}{\sqrt{\alpha-1}}\tan^{-1}\left(\frac{2\sqrt{\alpha-1}\left[1-\frac{\beta_0(1-\sqrt{1-\alpha_0})}{2}\right]}{\alpha^2-2+\beta_0(1-\sqrt{1-\alpha_0})}\right)\,, \tag{S192}$$

will sensitively depend on the magnitude of the compound parameter

$$\beta_0(1-\sqrt{1-\alpha_0}) = \frac{v_0}{v}\left[1-\sqrt{1-\frac{4D\lambda_0}{v_0^2}}\right]\,. \tag{S193}$$

When $\beta_0(1-\sqrt{1-\alpha_0}) \ll 1$, the solution for $\tilde{\ell}$ is only slightly perturbed from its value in our original model in Eq. (S142). Since we are primarily interested in cases where $v_0 \gtrsim v$, this condition will apply whenever

$$\alpha_0 \ll \frac{1}{\beta_0} \lesssim 1\,, \tag{S194}$$

or in terms of the original parameters,

$$\frac{4D\lambda_0}{v_0^2} \ll \frac{v}{v_0} \lesssim 1\,. \tag{S195}$$

For a fixed value of $\lambda_0 \lesssim \lambda$, this will always be satisfied for sufficiently large $v_0$.

On the other hand, when $\beta_0(1-\sqrt{1-\alpha_0}) \gtrsim 1$, the solutions for $\tilde{\ell}$ and $\rho(z)$ will be substantially different than the original model in Supplementary Information 4.1. This can occur both when $\alpha_0$ is close to the washout threshold ($1-\alpha_0 \ll 1$), but also for small values of $\alpha_0$ provided that the flow rates are sufficiently high:

$$\frac{1}{\beta_0} \lesssim \alpha_0 \ll 1\,. \tag{S196}$$

This second case corresponds to a scenario where the growth rate $\lambda_0$ is much larger than $\lambda$, but $v_0$ is still sufficiently high to ensure that the washout condition is still satisfied ($\alpha_0 \ll 1$). This represents an inversion of the usual scenario, in which the vast majority of biomass is produced in the washed-out region due to the diffusion of cells from a stable population further downstream. While this is a potentially interesting regime, we believe that it is unlikely to apply for a typical bacterial population in the human colon. We will therefore focus on the first case ($\alpha_0 \ll \beta_0^{-1} \lesssim 1$) below, while leaving a full analysis of the other regimes for future work.

When $\alpha_0 \ll \beta_0^{-1} \lesssim 1$, the full piece-wise solution for the density profile reduces to

$$\rho(z) \approx \begin{cases} A_\rho e^{-2\beta_0|z|} & \text{if } z<0, \\ A_\rho\left[\cos(kz)+\frac{1}{k}\sin(kz)\right]e^{z} & \text{if } 0<z<\tilde{\ell}, \\ A_\rho[\cos(k\tilde{\ell})+\frac{1}{k}\sin(k\tilde{\ell})]e^{\tilde{\ell}} & \text{if } z>\tilde{\ell}. \end{cases} \tag{S197}$$

where $\tilde{\ell}$ is approximately equal to the original solution in Eq. (S142). Similarly, neutral fixation probability is given by

$$u(z) \approx \begin{cases} A_u e^{-\beta_0\alpha_0|z|/2} & \text{if } z < 0, \\ A_u \left[\cos(kz) + \frac{1}{k}\sin(kz)\right] e^{-z} & \text{if } 0 < z < \tilde{\ell}, \\ A_u[\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})]e^{\tilde{\ell}-2z} & \text{if } z > \tilde{\ell}, \end{cases} \tag{S198}$$

and the distribution of future common ancestors is

$$g(z) \approx \begin{cases} A_u A_\rho e^{-2\beta_0|z|} & \text{if } z < 0, \\ A_u A_\rho \left[\cos(kz) + \frac{1+\beta_0(k_0-1)}{k}\sin(kz)\right]^2 & \text{if } 0 < z < \tilde{\ell}, \\ A_u A_\rho[\cos(k\tilde{\ell}) + \frac{1+\beta_0(k_0-1)}{k}\sin(k\tilde{\ell})]^2 e^{2(\tilde{\ell}-z)} & \text{if } z > \tilde{\ell}. \end{cases} \tag{S199}$$

The effective evolutionary parameters can be obtained by essentially repeating the same derivation from Supplementary Information 4.2. The total population size is still dominated by the $z > \tilde{\ell}$ region, so the normalization constant $A_\rho$ is unchanged from before. The normalization condition, $\int g(z)dz = 1$, now yields a slightly different expression for the normalization constant $A_u$:

$$(A_\rho A_u)^{-1} = \underbrace{\int_{-\infty}^{0} e^{-2\beta_0|z|}\, dz}_{I_{-1}} + \underbrace{\int_0^{\tilde{\ell}} \left[\cos(kz) + \frac{1}{k}\sin(kz)\right]^2 dz}_{I_1} + \underbrace{\int_{\tilde{\ell}}^{\infty} \left[\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})\right]^2 e^{-2(z-\tilde{\ell})}\, dz}_{I_2}, \tag{S200}$$

which includes an additional contribution from the integral

$$I_{-1} \equiv \int_{-\infty}^{0} e^{-2\beta_0|z|}\, dz = \frac{1}{2\beta_0}. \tag{S201}$$

The effective growth rate can then be expressed as

$$\frac{\lambda_e}{\lambda} = \int_{-\infty}^{0} \frac{\lambda_0}{\lambda}(A_\rho A_u)e^{-2\beta_0 k_0|z|}\, dz + \int_0^{\tilde{\ell}} (A_\rho A_u)\left[\cos(kz) + \frac{1+\beta_0(k_0-1)}{k}\sin(kz)\right]^2 dz, \tag{S202}$$

$$= \frac{\frac{\alpha_0\beta_0^2}{\alpha} I_{-1} + I_1}{I_{-1} + I_1 + I_2}, \tag{S203}$$

where $I_1$ and $I_2$ are the same as in Supplementary Information 4.2 above. Similarly, the effective population size can be written as

$$\left(\frac{N_e}{N}\right)^{-1} = \frac{1}{\lambda_e}\left[\int \rho(z)\, dz\right] \int \left[\lambda(z)u(z)^2 + \frac{2\lambda}{\alpha}(\partial_z u)^2\right]\rho(z)\, dz, \tag{S204}$$

$$= \left(\frac{\lambda}{\lambda_e}\right)\left(\frac{\int \rho(z)dz}{A_\rho}\right)(A_\rho A_u)^2 \left[\underbrace{\int_{-\infty}^{0} \frac{\lambda_0 u^2 \rho(z)\, dz}{\lambda A_\rho A_u^2}}_{I_{-3}} + \underbrace{\int_{-\infty}^{0} \frac{2(\partial_z u)^2 \rho(z)\, dz}{\alpha A_\rho A_u^2}}_{I_{-4}} + \underbrace{\int_0^{\tilde{\ell}} \frac{u(z)^2\rho(z)\, dz}{A_\rho A_u^2}}_{I_3} + \underbrace{\int_0^{\tilde{\ell}} \frac{2(\partial_z u)^2\rho(z)\, dz}{\alpha A_\rho A_u^2}}_{I_4} + \underbrace{\int_{\tilde{\ell}}^{\infty} \frac{2(\partial_z u)^2\rho(z)\, dz}{\alpha A_\rho A_u^2}}_{I_5}\right], \tag{S205}$$

which includes two new contributions,

$$I_{-3} = \frac{\lambda_0}{\lambda}\int_{-\infty}^{0} e^{-2\beta_0|z|}\, dz = \frac{\alpha_0\beta_0}{2\alpha}, \tag{S206}$$

$$I_{-4} = \frac{2}{\alpha}\int_{-\infty}^{0} \left(\frac{\beta_0\alpha_0}{2}\right)^2 e^{-2\beta_0|z|}\, dz = \frac{\beta_0\alpha_0^2}{4\alpha}, \tag{S207}$$

from the $z < 0$ region. Putting everything together, we find that

$$\frac{N_e}{N} \cdot \frac{vL}{2D} = \frac{e^{-\tilde{\ell}}}{\cos(k\tilde{\ell}) + \frac{1}{k}\sin(k\tilde{\ell})} \cdot \frac{(\frac{\alpha_0\beta_0^2}{\alpha} I_{-1} + I_1)(I_{-1} + I_1 + I_2)}{I_{-3} + I_{-4} + I_3 + I_4 + I_5}\,, \tag{S208}$$

where $I_1$, $I_2$, $I_3$, $I_4$, and $I_5$ are the same as in Supplementary Information 4.2 above. Comparing these results to our previous expressions, we see that $I_{-3}$ and $I_{-4}$ are smaller than $I_3 + I_4 + I_5$ by an additional factor of $\alpha_0\beta_0 \ll 1$, so they make a negligible contribution to the denominator of Eq. (S208). This implies that $N_e$ is at least as large as in the original model in Fig. 2. Taken together, this shows that the addition of a strongly washed out region to an otherwise established population has only a small perturbative effect on the scaled effective population size (provided that the washed out growth rate does not become arbitrarily large). We can therefore conclude that it is the extra growth, rather than washout per se, that drives the enhanced rates of genetic drift in Fig. 2F.

#### 4. Establishment probability of strongly beneficial mutations

We calculated the establishment probabilities in Fig. 3B,C by numerically solving the differential equation in Eq. (S114) for the special case where $\lambda_m(x) = (1+s)\lambda(x)$. In our minimal model, this equation reduces to the simple form,

$$0 = D\gamma'' + D(\gamma')^2 + v\gamma' + (1+s)\lambda - w(0)\left[\frac{(1+s)\lambda e^{\gamma}}{2} - De^{\gamma}[\gamma'(x)]^2\right]\,, \tag{S209}$$

for $0 \le x \le \ell$, while the boundary conditions are given by Eqs. (S115) and (S119):

$$\gamma(0) = \gamma'(0) = 0\,, \tag{S210a}$$

$$\gamma'(\ell) + \frac{v}{D}\left[\frac{e^{w(0)e^{\gamma(\ell)}} - 1}{w(0)e^{\gamma(\ell)}}\right] = 0\,. \tag{S210b}$$

Numerical solutions to this boundary value problem were obtained using the shooting method. For a given combination of $s$, $\lambda$, $v$, and $D$, we scanned through a range of $w(0)$ values and computed the corresponding values of $\gamma(\ell)$ and $\gamma'(\ell)$ by integrating Eq. (S209) from the initial condition $\gamma(0) = \gamma'(0) = 0$ using the `solve_ivp` function in SciPy. We then used these $\gamma(\ell)$ and $\gamma'(\ell)$ values to determine the unique choice of $w(0)$ that satisfies the boundary condition in Eq. (S210b). The resulting profiles, $w(x) = w(0)e^{\gamma(x)}$, were used to calculate the establishment probabilities in Fig. 3.

#### 5. Long-term growth rate of strongly beneficial mutations

The establishment probabilities in Supplementary Information 4.4 can be contrasted with the long-term growth rates of beneficial mutations derived in Supplementary Information 3.2. In our minimal model, Eq. (S122) reduces to

$$0 = \partial_z^2\psi(z) - 2\partial_z\psi(z) + \alpha\left[(1+s)\theta(\tilde{\ell} - z) - sy\right]\psi(z)\,, \tag{S211}$$

where we have adopted the same diffusion units as in Eq. (S131) above, and defined the scaled exponential growth rate

$$y \equiv \frac{S}{s\lambda}\,. \tag{S212}$$

This equation admits a piecewise solution of the form,

$$\psi(x) = \begin{cases} A_1 e^{z}\left[\cos(k_1 z) + b_1\sin(k_1 z)\right] & \text{if } z < \tilde{\ell}\,, \\ A_2 e^{z}\left[e^{-k_2 z} + b_2 e^{k_2 z}\right] & \text{if } z > \ell\,, \end{cases} \tag{S213}$$

where $k_1$ and $k_2$ are given by

$$k_1 = k\sqrt{1 + \frac{s\lambda(1-y)}{k^2}}\,, \tag{S214}$$

$$k_2 = \sqrt{1 + \frac{sy}{\alpha}}\,, \tag{S215}$$

and $k = \sqrt{\alpha - 1}$ is the neutral inverse length scale defined in Eq. (S138) above. The boundary conditions at $z = 0$ and $z = \infty$ require that $b_1 = 1/k_1$ and $b_2 = 0$. The continuity conditions at $z = \tilde{\ell}$ then lead to an eigenvalue equation for $y$:

$$-k_1 \sin(k_1\tilde{\ell}) + \cos(k_1\tilde{\ell}) = -k_2 \left[\cos(k_1\tilde{\ell}) + \frac{1}{k_1}\sin(k_1\ell)\right] , \tag{S216}$$

which can be rewritten in the compact form

$$\tan\left[k\tilde{\ell}\sqrt{1 + \frac{\alpha s(1-y)}{\alpha - 1}}\right] = \left(\frac{2\sqrt{\alpha-1}}{\alpha-2}\right) \cdot \frac{\sqrt{1 + \frac{\alpha s(1-y)}{\alpha-1}}\left[1 + \frac{\sqrt{1+\alpha s y}-1}{2}\right]}{1 + \frac{\alpha s(1-y)}{\alpha-2} + \frac{1-\sqrt{1+\alpha s y}}{\alpha-2}} . \tag{S217}$$

When $s = 0$, this reduces to our earlier expression in Eq. (S143) relating $k\tilde{\ell}$ and $\alpha$. For nonzero values of $s$, it allows us to solve for $y$ as a function of $s$ and $\alpha$.

Motivated by Fig. 3, we will be particularly interested in regimes where $s \ll 1$, but not necessarily small compared to $1/\lambda_e\tau_{\text{mix}}$. In this regime, we will be able to show that the $s$-dependent terms in the square roots will always be small compared to one. Expanding Eq. (S217) to first order in $s$, we obtain

$$\frac{1 + \tan^2(k\tilde{\ell})}{\tan(k\tilde{\ell})} \cdot k\tilde{\ell} \cdot \frac{(1-y)}{2(\alpha-1)} = \frac{(1-y)}{2(\alpha-1)} + \frac{y}{4} - \frac{(1-y)}{\alpha-2} + \frac{y}{2(\alpha-2)} , \tag{S218}$$

We can evaluate this expression in the diffusion-dominated growth ($\alpha \gg 1$) and flow-dominated growth ($\alpha - 1 \ll 1$) regimes respectively. In the flow-dominated regime, we have $k\tilde{\ell} \sim \pi - 2\sqrt{\alpha-1}$, so that Eq. (S218) reduces to

$$\frac{\pi(1-y)}{4(\alpha-1)^{3/2}} = \frac{(1-y)}{2(\alpha-1)} + \frac{y}{4} + (1-y) - \frac{y}{2} , \tag{S219}$$

and hence

$$y = 1 - \frac{(\alpha-1)^{3/2}}{\pi} . \tag{S220}$$

On the other hand, in the diffusion-dominated regime, we have $k\tilde{\ell} \sim 2/\sqrt{\alpha}$, so that Eq. (S218) reduces to

$$\frac{(1-y)}{2\alpha} = \frac{(1-y)}{2\alpha} + \frac{y}{4} - \frac{(1-y)}{\alpha} + \frac{y}{2\alpha} , \tag{S221}$$

and

$$y = \frac{4}{\alpha} . \tag{S222}$$

Switching back to $S$, this implies that $S \approx s_e\lambda_e$ in both the large and small $\alpha$ limits provided that $s \lesssim 1$.

## 5. APPLICATION TO BACTERIA IN THE HUMAN COLON

We sought to apply our framework to the human gut microbiota by leveraging the extensive parameter estimates in Ref. (28). The authors of this study developed a quantitative modeling framework (similar to Eq. 1) that explicitly accounts for the interplay between fluid flow, nutrient uptake, short-chain fatty acid production, and changes in local pH, while neglecting cell death [$d(x) \ll \lambda(x)$]. The authors combined this computational model with *in vitro* growth measurements and published data on human physiology to obtain spatially resolved predictions for the steady-state growth rates and density profiles of two representative species in the human gut: *Bacteroides thetaiotaomicron* and *Eubacterium rectale*. We adapted these estimates for our evolutionary framework using the procedure described below.

We began by manually extracting the steady-state growth rate profiles from Fig. 3B in Ref. (28). Since these growth profiles represent solutions to a boundary value problem, we expect that small numerical artifacts in the extraction process will prevent these initial estimates from admitting a steady state solution in Eq. (2). To overcome this issue, we "refit" the extracted growth profiles by introducing a small horizontal shift

$$\lambda_{\text{refit}}(x) = \lambda_{\text{measured}}(x + \delta\ell) , \tag{S223}$$

and determined the unique value of $\delta\ell$ that yields a steady state solution for $u(x)$ and $\rho(x)$. We carried out this refitting step using a variant of the shooting method described in Supplementary Information 4.4, setting $s \to 0$ so that $w(x)$ becomes proportional to $u(x)$. We verified that the fitted values of $\delta\ell$ were small compared to the other length scales in the problem — e.g. the minimum diffusion length scale $D/v(0)$, or the distance required for $\lambda_{\text{measured}}(x)$ to decay to half its maximum value — which suggests that our refitting procedure should have a negligible influence on the inferred evolutionary parameters.

Once the relevant values of $\delta\ell$ were identified, we used the same numerical procedures to generate the steady-state solutions for $u(x)$, $\rho(x)$, and $w(x)$, under the assumption that $\Upsilon(x) \approx \lambda(x)$. The effective evolutionary parameters in Fig. 4F-H were obtained by numerically integrating the resulting spatial profiles over the relevant values of $x$.

**Mutation trajectory data in Fig. 5.** The example mutation trajectories for the *B. massiliensis* population in Fig. 5 were obtained from our previous study (15), in which we performed deep metagenomic sequencing on 19 fecal samples from a single individual over a six month interval. Raw sequencing data was processed exactly as described in Fig. S7F of Ref. (15), producing a list of inferred mutation trajectories (Supplemental Data 1). This particular subject received a 2-week course of oral antibiotics starting on day 57, so we only plotted data from the samples that were collected prior to the start of treatment.

## 6. EXTENSIONS TO WALL GROWTH AND OTHER SPATIAL REFUGIA

Most of our analysis has focused on a simple model where all of the dynamics take place along the longitudinal axis of the lumen. In this section, we will show how our mathematical framework can be extended to account for simple forms of cross-sectional structure as well. Our use of a one-dimensional model was motivated by previous work by Cremer *et al* (26, 28), who argued that peristaltic mixing is sufficiently rapid in the human gut to destroy any cross-sectional structure within the luminal fluid itself. In particular, given an estimated effective diffusion constant of $D \approx 10^6 \mu m^2/s$, non-motile cells will diffuse across the inner diameter of the colon ($d \sim 3mm$) in $\sim$10 minutes (26). This is significantly shorter than the rate of cell division (Fig. 4), which suggests that any cross-sectional structure will quickly average out on longer evolutionary timescales (e.g. Fig. 5).

However, this simple argument will not necessarily extend to cells in the mucosa or the intestinal crypts (25), where mixing is thought to occur more slowly. By mitigating the effects of dilution, these cells could have a natural advantage in contributing genetic material to the future population, and could therefore play an outsized role in the emergent evolutionary dynamics. While a detailed analysis of surface growth is beyond the scope of the present paper, we can gain some insight into main effects by analyzing a simpler model — designed to provide a rough approximation to wall growth in the human gut — in which a well-mixed luminal compartment is connected to $K \gg 1$ semi-isolated subpopulations on the walls of the large intestine (Fig. S1).

We assume that each of these subpopulations (or *refugia*) contains $\sim\overline{N}_r$ cells, which compete amongst themselves, but are otherwise isolated from the other subpopulations on the wall. These dynamics can be viewed as an approximation to the local competition that emerges from the diffusion of resources within a continuous surface population. Here we approximate these dynamics with a simple local growth rate for each refuge,

$$\lambda_k(t) = \lambda_w \cdot \frac{\overline{N}_r}{N_k(t)}\,, \tag{S224}$$

which is inversely proportional to the current number of cells, $N_k(t)$. We assume that the cells in the lumen compete for a separate set of resources, so that they have their own local growth rate,

$$\lambda_\ell(t) = \lambda_\ell \cdot \frac{\overline{N}_\ell}{N_\ell(t)}\,. \tag{S225}$$

We note that the steady state growth rates $\lambda_\ell$ and $\lambda_w$ will not necessarily be equal to each other, so that the typical generation times will vary between the lumen and the wall.

In addition to growth, we assume that the wall populations and the luminal compartment can exchange cells with each other. We will let $m\lambda_w$ denote the per capita rate that cells migrate from the wall into the lumen, and we will let $\varepsilon\lambda_\ell/K$ denote the corresponding rate of migration from the lumen into a particular subpopulation on the wall. Note that $m$ and $\varepsilon$ are both scaled by their local generation time, so that they can be interpreted as a traditional per generation migration probability; the parameter $\varepsilon$ coincides with the *engraftment rate* defined in the main text, and

will play a crucial role in our analysis below. For simplicity, we will assume that individual wall populations cannot exchange cells with each other directly, but only indirectly through diffusion into the lumen. We expect this to be a reasonable approximation for scenarios like the human gut, where active mixing makes it much easier to travel large distances within the lumen. Finally, we assume that the cells in the lumen are diluted out of the population at a per capita rate $\delta$. This dilution rate is the key feature that distinguishes the lumen from the subpopulations on the walls: the latter populations only dilution indirectly by exchanging cells with the lumen. This provides them with a privileged spatial location whose effects we will investigate below.

At steady state ($N_\ell \approx \overline{N}_\ell$, $N_k \approx \overline{N}_r$), the input and output from each compartment must balance, so that

$$\Delta N_\ell = \lambda_\ell \overline{N}_\ell - \delta \overline{N}_\ell + \sum_{k=1}^{K} \left[ m\lambda_w \overline{N}_r - \left(\frac{\varepsilon\lambda_\ell}{K}\right) \overline{N}_\ell \right] = 0\,, \tag{S226}$$

$$\Delta N_k = \lambda_w \overline{N}_r + \left(\frac{\varepsilon\lambda_\ell}{K}\right) \overline{N}_\ell - m\lambda_w \overline{N}_r = 0\,. \tag{S227}$$

We can rearrange these equations to write them in the non-dimensional form,

$$\left(\frac{\delta}{\lambda_\ell}\right) = 1 + \frac{KN_r\lambda_w}{N_\ell\lambda_\ell}\,, \qquad m = 1 + \frac{\varepsilon N_\ell \lambda_\ell}{KN_r\lambda_w}\,. \tag{S228}$$

Since $\overline{N}_\ell$ and $K\overline{N}_r$ are both proportional to the longitudinal density of cells, we can also write this as

$$\left(\frac{\delta}{\lambda_\ell}\right) = 1 + \frac{\rho_w\lambda_w}{\rho_\ell\lambda_\ell}\,, \qquad m = 1 + \frac{\varepsilon\rho_\ell\lambda_\ell}{\rho_w\lambda_w}\,, \tag{S229}$$

where $\rho_\ell$ and $\rho_w$ denote the longitudinal densities in the lumen and wall respectively. Similar to our luminal models above, it will be useful to view these equations as determining $\delta$ and $m$ as a function of the "input parameters" $\varepsilon$, $\rho_w$, $\rho_\ell$, $\lambda_\ell$, and $\lambda_w$. Equation (S229) shows that this relationship is mediated through two compound parameters,

$$\epsilon_w \equiv \frac{\rho_w\lambda_w}{\rho_\ell\lambda_\ell}\,, \qquad \epsilon_e \equiv \frac{\varepsilon\rho_\ell\lambda_\ell}{\rho_w\lambda_w}\,. \tag{S230}$$

The first parameter ($\epsilon_w$) represents the ratio between the total biomass produced by wall growth and the biomass produced within the lumen. Cremer *et al* (26, 28) argued that this ratio is likely to be small for the most abundant species in the human gut, largely due to the differences in the longitudinal densities (e.g. $\rho_w/\rho_\ell \lesssim 10^{-3}$). The ratio between $\lambda_w$ and $\lambda_\ell$ is less well understood *in situ*, but it is reasonable to suppose that oxygen gradients (79, 80) and dense cellular packing in the wall (23) might limit $\lambda_w \lesssim \lambda_\ell$ for the most abundant species that are sampled in feces (26). Based on these considerations, we will restrict our attention to regimes where $\epsilon_w \ll 1$.

In a similar manner, the second parameter ($\epsilon_e$) can be interpreted as the ratio between the total number of cells that engraft into the wall from the lumen, and the total number produced by division within the wall itself. The empirical value of this parameter is largely unknown. Here we will focus on regimes where $\epsilon_e \ll 1$, since this corresponds most closely to our intuitive notion of a "refuge" (i.e., more cells produced within than without). We will see below that higher engraftment rates are already effectively well-mixed with the lumen, similar to the cross-sectional structure discussed above. In terms of the underlying parameters, the assumptions that $\epsilon_w \ll 1$ and $\epsilon_e \ll 1$ imply that

$$\varepsilon \ll \frac{\rho_w\lambda_w}{\rho_\ell\lambda_\ell} \ll 1\,. \tag{S231}$$

Note, however, that we make no assumptions about the *relative* values of $\epsilon_w$ and $\epsilon_e$. We will see that the relationship between these two small parameters will determine the transition between refuge-dominated and lumen-dominated dynamics below.

**Stochastic dynamics of a rare lineage.** We can obtain a diffusion model for the stochastic dynamics of a rare mutant lineage by repeating our derivation in Supplementary Information 1. Let $n_k$ denote the number of mutant cells in refuge $k$, and let $n_\ell$ the denote the corresponding number of cells in the lumen. Provided that $n_k \ll \overline{N}_r$ and

$n_\ell \ll \overline{N}_\ell$, the infinitesimal change in these quantities in a timestep $\Delta t$ is given by

$$\begin{aligned}\delta n_\ell &= \left(\lambda_\ell n_\ell \Delta t + \sqrt{\lambda_\ell n_\ell \Delta t} Z_{\lambda,\ell}\right) + \sum_k \left(m\lambda_w n_k \Delta t + \sqrt{m\lambda_w n_k \Delta t} Z_{m,k}\right) \\ &\quad - \left(\delta n_\ell \Delta t + \sqrt{\delta n_\ell \Delta t} Z_{\delta,\ell}\right) - \sum_k \left(\frac{\varepsilon \lambda_\ell n_\ell \Delta t}{K} + \sqrt{\frac{\varepsilon \lambda_\ell n_\ell \Delta t}{K}} Z_{e,k}\right) ,\end{aligned} \tag{S232a}$$

$$\begin{aligned}\delta n_k &= \left(\lambda_w n_k \Delta t + \sqrt{\lambda_w n_k \Delta t} Z_{\lambda,k}\right) - \left(m\lambda_w n_k \Delta t + \sqrt{m\lambda_w n_k \Delta t} Z_{m,k}\right) \\ &\quad + \left(\frac{\varepsilon \lambda_\ell n_\ell \Delta t}{K} + \sqrt{\frac{\varepsilon \lambda_\ell n_\ell \Delta t}{K}} Z_{e,k}\right) ,\end{aligned} \tag{S232b}$$

where the $\{Z_i\}$ again denote independent Gaussian random variables with zero mean and unit variance. After eliminating $\delta$ and $m$ using the steady-state conditions in Eq. (S229) and enforcing the limits that $\epsilon_w, \epsilon_e \ll 1$, these equations can be written in the compact form,

$$\delta n_\ell = -\epsilon_w \lambda_\ell n_\ell \Delta t + \sum_k \lambda_w n_k \Delta t + \sqrt{2\lambda_\ell n_\ell \Delta t} Z_\ell + \sum_k \left(\sqrt{\lambda_w n_k \Delta t} Z_{m,k} - \sqrt{\frac{\epsilon_e \epsilon_w \lambda_\ell n_\ell \Delta t}{K}} Z_{e,k}\right) , \tag{S233a}$$

$$\delta n_k = -\epsilon_e n_k \lambda_w \Delta t + \frac{\epsilon_e \epsilon_w n_\ell \lambda_\ell \Delta t}{K} + \sqrt{\lambda_w n_k \Delta t} Z_{\lambda,k} - \sqrt{\lambda_w n_k \Delta t} Z_{m,k} + \sqrt{\frac{\epsilon_e \epsilon_w \lambda_\ell n_\ell \Delta t}{K}} Z_{e,k} , \tag{S233b}$$

where we have merged the contributions of $Z_{\lambda,\ell}$ and $Z_{\delta,\ell}$ into a single variable $Z_\ell$. This update rule can also be expressed in terms of the local mutation frequencies, $f_k \equiv n_k/\overline{N}_r$ and $f_\ell \equiv n_\ell/\overline{N}_\ell$, so that

$$\delta f_\ell = \epsilon_w \lambda_\ell \left(\frac{1}{K}\sum_k f_k - f_\ell\right) \Delta t + \sqrt{\frac{2\lambda_\ell f_\ell \Delta t}{\overline{N}_\ell}} Z_\ell + \sum_k \left(\sqrt{\frac{\epsilon_w \lambda_\ell f_k \Delta t}{K\overline{N}_\ell}} Z_{m,k} - \sqrt{\frac{\epsilon_e \epsilon_w \lambda_\ell f_\ell \Delta t}{K\overline{N}_\ell}} Z_{e,k}\right) , \tag{S234a}$$

$$\delta f_k = -\epsilon_e \lambda_w \left(f_\ell - f_k\right) \Delta t + \sqrt{\frac{\lambda_w f_k \Delta t}{\overline{N}_r}} Z_{\lambda,k} - \sqrt{\frac{\lambda_w f_k \Delta t}{\overline{N}_r}} Z_{m,k} + \sqrt{\frac{\epsilon_e \lambda_w f_\ell \Delta t}{\overline{N}_r}} Z_{e,k} . \tag{S234b}$$

These frequency dynamics allow us to immediately identify two characteristic *mixing timescales* that equalize mutation frequencies across compartments. From Eq. (S234a), we can identify a characteristic *migration timescale*,

$$T_m \equiv \frac{1}{\lambda_\ell \epsilon_w} = \frac{1}{\lambda_\ell} \cdot \frac{\rho_\ell \lambda_\ell}{\rho_w \lambda_w} , \tag{S235}$$

which describes how rapidly the lumen population relaxes to the average composition of the wall. Similarly, Eq. (S234b) contains a corresponding *engraftment timescale*,

$$T_e \equiv \frac{1}{\lambda_w \epsilon_e} = \frac{1}{\lambda_w} \cdot \frac{1}{\varepsilon} \cdot \frac{\rho_w \lambda_w}{\rho_\ell \lambda_\ell} , \tag{S236}$$

which describes how rapidly the wall populations relax to a fixed composition of the lumen. The relationship between these two timescales will play a critical role in determining the emergent evolutionary dynamics below.

Using the update rule in Eq. (S232a), we can also derive an equation for the generalization of the moment generating function,

$$H(\phi_\ell, \phi_1, \ldots, \phi_K, t) \equiv \left\langle e^{-\phi_\ell n_\ell(t) - \sum_k \phi_k n_k(t)} \right\rangle , \tag{S237}$$

by repeating the derivation in Supplementary Information 1. This yields

$$\begin{aligned}\frac{\partial H}{\partial t} &= \lambda_\ell \left[-\epsilon_w \phi_\ell + \frac{\epsilon_e \epsilon_w}{K}\sum_k \phi_k - \phi_\ell^2 - \frac{\epsilon_e \epsilon_w}{2K}\sum_k (\phi_k - \phi_\ell)^2\right] \frac{\partial H}{\partial \phi_\ell} \\ &\quad + \sum_k \lambda_w \left[\phi_\ell - \epsilon_e \phi_k - \frac{\phi_k^2}{2} - \frac{(\phi_k - \phi_\ell)^2}{2}\right] \frac{\partial H}{\partial \phi_k} ,\end{aligned} \tag{S238}$$

with the initial condition $H(\vec{\phi},0) = e^{-\phi_\ell n_\ell(0) - \sum_k \phi_k n_k(0)}$.

**Rate of genetic drift.** We can use Eq. (S238) to calculate the effective rate of genetic drift by generalizing the separation-of-timescales approach in Supplementary Information 2. In this case, the linearized versions of the characteristic curves, $u_\ell$ and $u_k = u_r$, satisfy the simple relation,

$$u_\ell - \epsilon_e u_r = 0 \tag{S239}$$

and are normalized so that $K\overline{N}_r u_r + \overline{N}_\ell u_\ell = 1$. This yields

$$u_r = \frac{u_\ell}{\epsilon_e}\,, \qquad u_\ell = \frac{1}{N_\ell}\left[1 + \frac{\lambda_\ell}{\lambda_w}\frac{\epsilon_w}{\epsilon_e}\right]^{-1}. \tag{S240}$$

We can recognize the second term in this expression as the ratio between the mixing timescales $T_e$ and $T_m$ above. When $T_e \ll T_m$, the value of $u_\ell$ is completely unchanged in the presence of the wall population. In both cases, the value of $u_r$ is significantly larger than $u_\ell$, consistent with the notion that the refuges constitute privileged spatial locations. Our results show that the magnitude of this advantage is directly tied to the engraftment ratio $\epsilon_e$, which is consistent with our intuition above.

We can use these results to calculate the effective rate of genetic drift by repeating the perturbation expansion in Supplementary Information 2. This yields

$$\Lambda = 2\lambda_\ell u_\ell^2 \overline{N}_\ell + \epsilon_e \epsilon_w \lambda_w (u_r - u_\ell)^2 \overline{N}_\ell + \lambda_w u_r^2 K\overline{N}_r + \lambda_w (u_r - u_\ell)^2 K\overline{N}_r\,, \tag{S241}$$

which reduces to

$$\Lambda \approx \frac{2}{N_\ell}\cdot\frac{1+\frac{\epsilon_w}{\epsilon_e^2}}{\left[1+\frac{\lambda_\ell\epsilon_w}{\lambda_w\epsilon_e}\right]^2} \approx \frac{2\lambda_\ell}{\overline{N}_\ell}\cdot\frac{1+\frac{1}{\varepsilon^2}\left(\frac{\rho_w\lambda_w}{\rho_\ell\lambda_\ell}\right)^3}{\left[1+\frac{1}{\varepsilon}\frac{\lambda_w}{\lambda_\ell}\left(\frac{\rho_w}{\rho_\ell}\right)^2\right]^2} \tag{S242}$$

in the limit that $\epsilon_w, \epsilon_e \ll 1$. The value of this expression critically depends on how the engraftment ratio $\epsilon_e$ compares to the internal scales $\sqrt{\epsilon_w}$ and $\rho_w/\rho_\ell$, which are both small compared to one. When $\epsilon_e \gg \sqrt{\epsilon_w} \gg \rho_w/\rho_\ell$, we find that

$$\Lambda \equiv \Lambda_\ell \approx \frac{2\lambda_\ell}{\overline{N}_\ell}\,, \quad (\epsilon_e \gg \sqrt{\epsilon_w} \gg \rho_w/\rho_\ell)\,, \tag{S243}$$

which is equivalent to the original rate of drift in the absence of the wall. In terms of the underlying parameters, this regime occurs when

$$\varepsilon \gg \varepsilon_c \equiv (\rho_w\lambda_w/\rho_\ell\lambda_\ell)^{3/2}\,. \tag{S244}$$

This shows that even tiny rates of engraftment are sufficient to ensure convergence to the lumen-dominated regime. In the opposite extreme, where $\epsilon_e \ll \sqrt{\epsilon_w} \ll \rho_w/\rho_\ell$, we find that

$$\Lambda \approx \Lambda_w \equiv \frac{2\lambda_w}{K\overline{N}_r}\,, \quad (\epsilon_e \ll \rho_w/\rho_\ell \ll \sqrt{\epsilon_w})\,, \tag{S245}$$

which is equivalent to the rate of genetic drift in an effective population that is the same size as the entire wall. This amplifies the rate of genetic drift by a factor of

$$\frac{\Lambda_w}{\Lambda_\ell} = \frac{\rho_\ell}{\rho_w}\frac{\lambda_w}{\lambda_\ell} \gg 1\,. \tag{S246}$$

This amounts to a 1000-fold increase when $\rho_w/\rho_\ell \sim 10^{-3}$ and $\lambda_w \sim \lambda_w$. In terms of the underlying parameters, this wall-dominated regime occurs when

$$\varepsilon \ll \varepsilon_w \equiv \frac{\lambda_w}{\lambda_\ell}\left(\frac{\rho_w}{\rho_\ell}\right)^2\,, \tag{S247}$$

which is identical to the condition that the engraftment timescale $T_e$ in Eq. (S236) is much greater than the migration timescale $T_m$ in Eq. (S235). In other words, wall-dominated drift occurs when fluctuations in the wall propagate to the lumen more rapidly than the other way around.

In between these two extremes ($\rho_w/\rho_\ell \ll \epsilon_e \ll \sqrt{\epsilon_w}$), the rate of genetic drift scales as

$$\Lambda \approx \frac{2\lambda_\ell}{\overline{N}_\ell}\left(\frac{\epsilon_w}{\epsilon_e^2}\right) \approx \frac{2\lambda_\ell}{\overline{N}_\ell}\left(\frac{\rho_w\lambda_w}{\rho_\ell\lambda_\ell}\right)^3 \frac{1}{\varepsilon^2}\,, \quad (\rho_w/\rho_\ell \ll \epsilon_e \ll \sqrt{\epsilon_w})\,, \tag{S248}$$

which interpolates between the $\Lambda_\ell$ and $\Lambda_w$ limits above.

**Breakdown of the separation-of-timescales approximation.** The perturbation expansion used to derive Eq. (S242) will be valid on timescales much longer than the mixing timescales $T_m$ and $T_e$ in Eqs. (S235) and (S236). At a minimum, this requires that the drift timescale in the entire gut, $T_{\text{drift}}{\sim}1/\Lambda$, must be long compared to $T_m$ and $T_e$ above. When $T_e \gg T_m$, this yields the condition,

$$\varepsilon \gg \frac{\lambda_w}{\lambda_\ell \overline{N}_\ell}\,. \tag{S249}$$

Intuitively, this condition requires that large numbers of cells must successfully engraft somewhere in the wall population during a single wall generation time. However, in addition to this global constraint, our model in Eq. (S234b) also assumed that the mutant lineage remains rare in individual refugia as well. This requires that the mixing timescale $T_e$ must also be fast compared to the drift timescale within an *individual* subpopulation, $T_{\text{drift},k} \sim \overline{N}_r/\lambda_w$. This leads to a more stringent condition,

$$\varepsilon \gg \frac{\epsilon_w}{\overline{N}_r}\,, \tag{S250}$$

which depends on the absolute population size of an individual refuge. (Note that this is the first time where we have observed an explicit dependence on $\overline{N}_r$, as opposed to the entire wall population, $K\overline{N}_r$.)

When $\varepsilon \lesssim \epsilon_w/\overline{N}_r$, the separation-of-timescales approximation breaks down, and the dynamics are instead dominated by the slow spread of mutations across many approximately clonal refugia. We can understand these dynamics with a simple heuristic picture that is valid in the limit that $\varepsilon \ll \epsilon_w/\overline{N}_r$. At any individual time step, each subpopulation will be fixed for either the mutant or the wildtype, and the lumen will consist of a weighted mixture of these $K$ clonal populations. New lineages will enter a given subpopulation from the lumen at rate $\varepsilon\lambda_\ell\overline{N}_\ell/K$. The vast majority of these lineages will go extinct; however, with probability $1/\overline{N}_r$, a lucky migrant will drift to fixation within the refuge on a timescale $\sim\overline{N}_r/\lambda_w$ and will then propagate back into the lumen. The net effect of this event is to replace the focal subpopulation with a single individual drawn at random from one of the other $K$ refuges. Thus, the dynamics in this regime will be equivalent to an effective model with $K$ "individuals" that reproduce with an effective "division rate" $\lambda_{\text{e}} \equiv \varepsilon\lambda_\ell\overline{N}_\ell/K\overline{N}_r \approx T_e^{-1}$. On timescales $t \gg T_e \gg T_m$, this leads to an effective rate of genetic drift,

$$\Lambda_e \sim \frac{\varepsilon\lambda_\ell\overline{N}_\ell}{K\overline{N}_r}\,, \tag{S251}$$

which can be substantially smaller than the wall-dominated regime above. We can now see that the origin of this slow behavior is the long time that it takes for mutations to sequentially fix across many individual refugia. This highlights an important way in which many isolated refugia are qualitatively different than a single one.

These results show that wall growth can only produce elevated rates of genetic drift within a limited parameter regime,

$$\frac{1}{\overline{N}_r}\left(\frac{\rho_w\lambda_w}{\rho_\ell\lambda_\ell}\right) \ll \varepsilon \ll \left(\frac{\rho_w\lambda_w}{\rho_\ell\lambda_\ell}\right)^{3/2}\,, \tag{S252}$$

which sensitively depends on the value of $\overline{N}_r$. In particular, if $\overline{N}_r \lesssim (\rho_w\lambda_w/\rho_\ell\lambda_\ell)^{1/2}$, wall growth can only *reduce* the apparent rates of genetic drift. Larger rates of drift are possible above this threshold, up to a maximum possible amplification factor

$$\frac{\Lambda_{\max}}{\Lambda_\ell} \sim \left(\frac{\rho_w\lambda_w}{\rho_\ell\lambda_\ell}\right)\overline{N}_r^2\,. \tag{S253}$$

Unlike Eq. (S246) above, this bound becomes smaller for smaller values of $\rho_w \lambda_w / \rho_\ell \lambda_\ell$, and can only be counteracted by larger local refuge sizes $\overline{N}_r$. While the relevant magnitudes of $\overline{N}_r$ are largely unknown, these analytical results nevertheless place strong constraints on the ability of wall growth (or other spatial refugia) to enhance the rate of genetic drift. Rough estimates of the relevant parameters suggest that these effects are unlikely to explain the large shifts in frequency observed in the human fecal samples in Fig. 5. Our results also highlight the importance of migration through the lumen. This suggests that it will be impossible to fully separate the impact of wall growth from the longitudinal dynamics within the luminal fluid above. We can attain a rough correspondence between these two pictures by taking $\overline{N}_\ell \approx \rho_\ell \ell_g$, where $\ell_g$ is the typical length scale of the ancestral region, $g(x) \propto u(x)\rho(x)$, in Figs 2 and 4. However, we caution that our current analysis has focused only on the simplest possible caricatures of wall growth, and omits many important features of real microbiota. A detailed understanding of these surface-growth dynamics (and their interactions with the luminal fluid above) remains an important topic for future work.

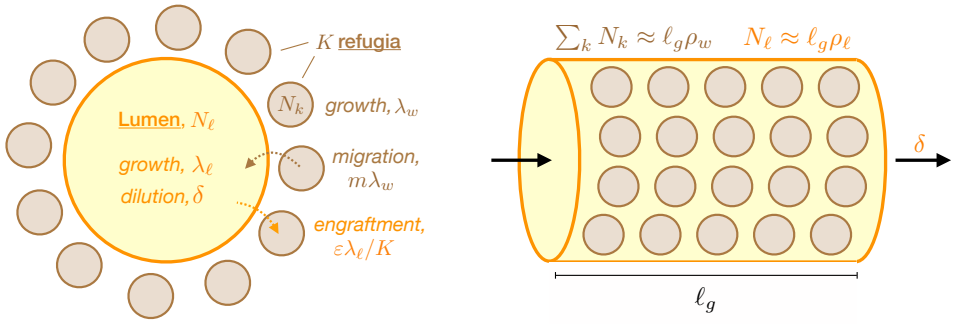


FIG. S1 **Schematic of simplified model of wall growth.** A well-mixed luminal compartment exchanges cells with $K$ semi-isolated subpopulations (or "refugia") arranged on the walls. Cells move from the refugia to the lumen with a per capita migration rate $m\lambda_w$, where $\lambda_w$ is the growth rate on the wall; conversely, cells move from the lumen into the wall with a per capita engraftment rate $\epsilon\lambda_\ell$, where $\lambda_\ell$ is the growth rate in the lumen. Cells in the lumen experience constant dilution at rate $\delta$, while cells in the refugia can only be diluted if they happen to migrate into the lumen. The balance between growth, dilution, and migration leads to steady state population sizes $N_\ell \approx \overline{N}_\ell$ and $\sum_k N_k \approx K\overline{N}_r$ in the lumen and wall, respectively. If the luminal compartment has a cylindrical geometry with length $\ell_g$, these steady-state population sizes are proportional to the longitudinal densities $\rho_\ell$ and $\rho_w$ of the lumen and wall, respectively.