Supplementary Information for
Collective behavior and non-genetic inheritance allow bacterial populations to adapt to changing environments

Henry Mattingly$^{1,2}$ and Thierry Emonet$^{1,2,3}$

$^1$ Department of Molecular, Cellular, and Developmental Biology, Yale University
$^2$ Quantitative Biology Institute, Yale University
$^3$ Department of Physics, Yale University

Thierry Emonet
Email: thierry.emonet@yale.edu

This PDF file includes:

- Supplementary text
- Figures S1 to S6
- Table S1
- SI References
**Supplementary Information Text**

**Model of collective migration without growth**

We model collective migration without using an extension of the Keller-Segel model, similar to one developed recently (1). The main differences from Keller-Segel are: 1) we model co-migration of *E. coli* with different swimming phenotypes, quantified by tumble bias, *TB*; 2) cells’ receptors have finite sensitivity for the chemotactant, set by the receptor-ligand dissociation constant when the receptors are in the inactive state, $K_i$ (2).

The full PDEs for cell density $\rho(x,t,TB)$ and attractant concentration $s(x,t)$ are (reproducing Eqns. 6 and 7 of the main text without the growth term):

\[
\frac{\partial t}{\partial x} \rho(x,t,TB) = -f(x) \rho(x,t,TB) \frac{1}{1 + \frac{s}{K_i}} - \mu(TB) \frac{\partial x}{\partial x} \rho(x,t,TB),
\]

\[
\frac{\partial t}{\partial x} s(x,t) = D_s \frac{\partial x}{\partial x} s(x,t) - k \frac{s}{K_m + s} \int_{TB} \rho(x,t,TB) \, dTB.
\]

**Motility models in liquid and porous environments**

Simulating collective migration requires models for $\chi(TB)$ and $\mu(TB)$ in each environment. Several previous works have shown that the diffusion coefficient $\mu(TB)$ resulting from run-and-tumble motion in 3D liquid is:

\[
\mu_{\text{liquid}}(TB) = \frac{1}{3} \frac{v_0^2}{(1 - \theta_T) \lambda_{RO}(TB) + 2 D_r (1 - TB)},
\]

where $v_0$ is the run speed, $\theta_T$ is the persistence of tumble (expectation of cosine of the angle between the cell’s heading before and after a tumble) (3), $\lambda_{RO}(TB)$ is the average run-to-tumble rate (hence “R”) in the absence of a gradient (hence “0”) (4), and $D_r$ is the rotational diffusion coefficient. We assumed that $v_0$, $\theta$, and $D_r$ are independent of tumble bias $TB$ (5).

$\lambda_{RO}(TB)$ was described by a motor model developed previously (4–7), consisting of a free energy barrier of height $F$ between run and tumble states when the cell is adapted to the background concentration of attractant and a basal switching frequency $\omega$:

\[
F(TB) = \log \left( \frac{1 - TB}{TB} \right),
\]

\[
\lambda_{RO/RO}(TB) = \omega \exp \left( \frac{1}{2} F(TB) \right),
\]

where $\lambda_{RO}$ is the average tumble-to-run transition rate. Choosing a tumble bias phenotype $TB$ immediately fixes $F$, $\lambda_{RO}$, and $\lambda_{T0}$. It also fixes the baseline free energy difference $f_0$ between receptor active and inactive states and baseline receptor-associated CheA kinase activity $a_0$ in the absence of a gradient.

In all main text simulations, $\chi_{\text{liquid}}(TB) \propto \frac{(1 - \theta_T) \lambda_{RO}(TB)}{(1 - \theta_T) \lambda_{RO}(TB) + 2 D_r} \mu_{\text{liquid}}(TB)$, with the proportionality constant chosen such that $a_{\text{liquid}} = \langle \chi_{\text{liquid}}(TB)/\mu_{\text{liquid}}(TB) \rangle = 10$ to match its value in a biophysical model, below. The expectation was taken with respect to the batch culture distribution of phenotypes (see section Growth model).
In porous environments, we used an expression for the diffusion coefficient derived by Licata et al. (8):

$$\mu_{\text{porous}}(TB) = \mu_{\text{liquid}}(TB) \left( 1 - \frac{1}{2} \theta_p \exp(-\theta_p) \right)$$  

(6)

Here, $\beta = l \frac{\lambda}{v_0}$, where $l$ is the typical pore size of the porous mesh. To arrive at this expression, the authors assumed that runs that displaced the cell by more than $l$ were truncated at $l$, and the cell stalled in place until its next tumble. $l$ was chosen such that the population average diffusivity was about 50 $\mu$m$^2$/s, in line with recent measurements (9), giving $l \sim 30 \mu$m. Finally, we assumed that $\chi_{\text{agar}}(TB) \propto \frac{(1-\theta_p)\lambda \alpha_g}{(1-\theta_p)\lambda \alpha_g + 2 D_r} \mu_{\text{agar}}(TB)$, with proportionality constant chosen such that $\alpha_{\text{agar}} = \frac{\chi_{\text{agar}}(TB)/\mu_{\text{agar}}(TB)}{\mu_{\text{agar}}(TB)} = 6.5$ (9,10).

To test whether the forms of $\chi(TB)$ we chose had qualitative effects on our results, we also modeled $\chi_{\text{liquid}}(TB)$ biophysically (4,5):

$$\chi_{\text{liquid}}(TB) = N \chi'_R(f) \left( \frac{\tau_m(TB)}{\lambda R_0(TB)} \right)$$

(7)

Here, $N$ is the cooperativity of the receptors that sense the ligand, $\tau_m(TB)$ is the chemotaxis pathway adaptation time, and $\chi'_R(f)$ is the gain of the tumble rate with respect to changes in receptor free energy $f$ in the adapted state $f_0$. The form of $\chi'_R(f)$ and associated parameter values were taken from refs (4,7), and a phenomenological model for $\tau_m(TB)$ was taken from refs (1,11).

Simulations of migrating groups using this model showed the same qualitative behavior as described in the main text.

**Quasi-steady state solutions**

To simplify the PDEs above, we first assume that $K_m$ is small, justified by measurements putting its value below 1 $\mu$M (1,12). To make analytical progress, we also assume that $s << K_s$, the concentration of attractant is much smaller than the receptor-ligand dissociation constant when the receptors are in the active form. With these assumptions, the PDEs become:

$$\partial_t \rho(x,t,TB) = -\partial_x \left( \chi(TB) \partial_x \log \left( \frac{1 + \frac{s(x,t)}{K_i}}{K_i} \right) \rho(x,t,TB) - \mu(TB) \partial_x \rho(x,t,TB) \right)$$  

(8)

$$\partial_t s(x,t) = D_s \partial_{xx} s(x,t) - k \theta(s(x,t)) \int_{TB} \rho(x,t,TB) \, dTB,$$

(9)

where $\theta(\ldots)$ is the Heaviside step function, which ensures that attractant can only be consumed where it is present.

Next, we define a moving reference frame, $z = x - \int_0^t c(t') \, dt'$ and convert to the moving reference frame:

$$\partial_t \rho(z,t,TB) - c(t) \partial_z \rho(z,t,TB) = -\partial_z \left( \chi(TB) \partial_z \log \left( 1 + \frac{s(z,t)}{K_i} \right) \rho(z,t,TB) - \mu(TB) \partial_z \rho(z,t,TB) \right)$$  

(10)

$$\partial_t s(z,t) - c(t) \partial_z s(z,t) = D_s \partial_{zz} s(z,t) - k \theta(s) \rho(z,t),$$

(11)

where $\rho(z,t) = \int_{TB} \rho(x,t,TB) \, dTB$.

For wave solutions, we seek a solution satisfying the following boundary conditions:
\[
\rho(z, t, TB) \to 0 \text{ as } z \to \infty; \partial_2 \rho(z, t, TB) \to 0 \text{ as } z \to \infty; s(z, t) \to s_\infty \text{ as } z \to \infty.
\] (12)

The equation for \( \rho(z, t, TB) \) can also be written in terms of the fluxes \( J(z, t, TB) \) of each phenotype in the moving reference frame:

\[
\partial_t \rho(z, t, TB) = -\partial_2 J(z, t, TB),
\] (13)

\[
J(z, t, TB) = \left( \chi(TB) \partial_z \log \left( \frac{\epsilon + s(z, t)}{s_\infty} \right) - c(t) \right) \rho(z, t, TB) - \mu(TB) \partial_2 \rho(z, t, TB),
\] (14)

where we have introduced \( \epsilon = \frac{K_i}{s_\infty} \ll 1 \).

Once the wave forms, it rapidly reaches a quasi-steady state; only on long time scales does the wave slow down due to loss of cells. Therefore, we can determine the quasi-steady wave solution by dropping the time derivatives (13).

First focusing on the equation for \( \rho(z, t, TB) \), we integrate once with respect to \( z \) and rearrange terms to get:

\[
\partial_z \rho(z, t, TB) - \left( \frac{\chi(TB)}{\mu(TB)} \right) \partial_z \log \left( \frac{\epsilon + s(z, t)}{s_\infty} \right) - \frac{c(t)}{\mu(TB)} \rho(z, t, TB) = 0
\] (15)

where, after applying the boundary conditions, the “constant” of integration is zero. This can be solved with an integrating factor

\[
\exp(U(z, t, TB)) = \exp \left( \int_0^z \left( \frac{\chi(TB)}{\mu(TB)} \partial_z \log \left( \frac{\epsilon + s(z, t)}{s_\infty} \right) - \frac{c(t)}{\mu(TB)} \right) dz \right)
\] (16)

\[
= \exp \left( - \frac{\chi(TB)}{\mu(TB)} \log \left( \epsilon + \frac{s(z, t)}{s_\infty} \right) - \frac{c(t)}{\mu(TB)} z \right)
\] (17)

\( U(z, t, TB) \) is analogous to an energy potential. Multiplying through by \( \exp(U(z, t, TB)) \) and completing the total derivative gives:

\[
\partial_z \left( \exp(U(z, t, TB)) \right) \rho(z, t, TB) = 0
\] (18)

Integrating with respect to \( z \) gives

\[
\rho(z, t, TB) = Q(t, TB) \exp \left( -U(z, t, TB) \right)
\] (19)

where \( Q(t, TB) \) comes from integration. The exact form of \( Q(t, TB) \) cannot be solved in closed form, but is constrained implicitly by the distribution of phenotypes traveling at time \( t \).

The equation for \( s(z, t) \) is:

\[
c(t) \partial_z s(z, t) = D_s \partial_{zz} s(z, t) - k \Theta(s) \rho(z, t).
\] (20)

Plugging in Eqn. (19) and dropping the Heaviside function, for now:

\[
-c(t) \partial_z s(z, t) = D_s \partial_{zz} s(z, t) - k \int_{TB} Q(t, TB) \exp \left( -U(z, t, TB) \right) dTB
\] (21)
\[-c(t) \partial_z s(z,t) = D_s \partial_{zz}s(z,t) - k \int_{TB} Q(t,TB) \exp \left( \frac{\chi(TB)}{\mu(TB)} \log \left( \epsilon + \frac{s(z,t)}{s_\infty} \right) - c(t) \frac{z}{\mu(TB)} \right) dTB \]  

(22)

Here, in order to proceed, we have to make the assumption that \(\frac{\chi(TB)}{\mu(TB)} = \alpha\), a constant for all phenotypes. This assumption has been used before (9,10,14) and is based on the fact that directed and undirected motility both arise from the cells’ run-and-tumble swimming behavior. But it ignores several kinds of cell-to-cell variations, such as variations in receptor array composition (15–20) and adaptation time (19,21,22). While this assumption is necessary to make analytical progress, as mentioned in this discussion, it does not affect our qualitative conclusions.

With this assumption, \(s(z,t)\) can be separated from the integral over phenotypes:

\[-c(t) \frac{\partial_z s(z,t)}{\partial z} = D_s \frac{\partial_{zz}s(z,t)}{\partial z} - k \int_{TB} Q(t,TB) \exp \left( \alpha \log \left( \epsilon + \frac{s(z,t)}{s_\infty} \right) - c(t) \frac{z}{\mu(TB)} \right) dTB \]  

(23)

Rearranging:

\[\left( \epsilon + \frac{s(z,t)}{s_\infty} \right)^{-\alpha} \frac{\partial_z s(z,t)}{s_\infty} + \frac{D_s}{c(t)} \partial_z \log \left( \frac{s(z,t)}{s_\infty} \right) = \frac{k}{s_\infty c(t)} \int_{TB} Q(t,TB) \exp \left( - \frac{c(t) z}{\mu(TB)} \right) dTB \]  

(24)

The terms inside the derivative on the left-hand side can be written

\[\left( \epsilon + \frac{s(z,t)}{s_\infty} \right)^{-\alpha} \partial_z \left( \frac{s(z,t)}{s_\infty} \left( 1 + \frac{D_s}{c(t)} \partial_z \log \left( \frac{s(z,t)}{s_\infty} \right) \right) \right) = \frac{k}{s_\infty c(t)} \int_{TB} Q(t,TB) \exp \left( - \frac{c(t) z}{\mu(TB)} \right) dTB \]  

(25)

This indicates that when \(\frac{D_s}{c(t)} \partial_z \log \left( \frac{s(z,t)}{s_\infty} \right) \ll 1\), or the length scale of attractant diffusion in the wave is much shorter than the traveling gradient length scale, then attractant diffusion can be neglected. For now, we make this assumption.

Dropping the attractant diffusion term, and integrating with respect to \(z\) gives:

\[\frac{1}{1-\alpha} \left( \epsilon + \frac{s(z,t)}{s_\infty} \right)^{-\alpha} = - \frac{k}{s_\infty c(t)} \int_{TB} Q(t,TB) \exp \left( - \frac{c(t) z}{\mu(TB)} \right) dTB + A(t,TB) \]  

(26)

Applying the boundary conditions, we find:

\[\frac{1}{1-\alpha} \left( \epsilon + 1 \right)^{-\alpha} = A(t,TB). \]  

(27)

Using this and rearranging Eqn. (26):

\[\left( \frac{\epsilon + \frac{s(z,t)}{s_\infty}}{1 + \epsilon} \right)^{-\alpha} = 1 + \frac{1}{(1 + \epsilon)^{1-\alpha}} (\alpha - 1) - \frac{k}{s_\infty c(t)} \int_{TB} Q(t,TB) \exp \left( - \frac{c(t) z}{\mu(TB)} \right) dTB. \]  

(28)

Solving for \(s(z,t)\):

\[\frac{s(z,t)}{s_\infty} = (1 + \epsilon) \left( 1 + \frac{(\alpha - 1)}{1 + \epsilon} \frac{k}{s_\infty c(t)} \int_{TB} Q(t,TB) \exp \left( - \frac{c(t) z}{\mu(TB)} \right) dTB \right)^{\frac{1}{\alpha - 1}} \]  

(29)
Finally, we redefine \( Q(t, TB) \rightarrow \frac{c(t)}{\mu(TB)} \frac{s_\infty \ c(t)}{k} (1 + \epsilon)^{1-\alpha} Q(t, TB) \) to get:

\[
\frac{s(z, t)}{s_\infty} = (1 + \epsilon) \left( 1 + (\alpha - 1) \int_{TB}^\infty Q(t, TB) \exp \left( -\frac{c(t) z}{\mu(TB)} \right) dTB \right)^\frac{1}{\alpha-1} - \epsilon
\]  

(30)

Due to the Heaviside function in the approximate PDEs, this solution holds where \( s(z, t) > 0 \), i.e. for \( z > z_0 \); otherwise, \( s(z, t) = 0 \). Since we neglected attractant diffusion, this concentration profile has a kink at \( z = z_0 \), making its derivative discontinuous there. Eqn. (30) differs from the solution of Keller and Segel (23) and the zeroth-order solution of Novick-Cohen and Segel (13) in that \( \epsilon \), the normalized sensitivity, is finite. This causes the attractant concentration to reach zero at a finite value of \( z \).

As in Keller & Segel’s original work, we can also integrate the equation for the attractant from \( z = z_0 \) to \( z \to \infty \), giving:

\[
c(t) (s_\infty - 0) = k \int_{z_0}^\infty \rho(z, t) \, dz \to c(t) = \frac{k}{s_\infty} N(t),
\]

(31)

where \( N(t) = \int_{z_0}^\infty \rho(z, t) \, dz \) is the number of cells (per cross-sectional area) migrating at time \( t \). As in Keller and Segel’s model, the migration speed \( c(t) \) depends on the number of cells traveling because migration is driven by the cells’ consumption of the attractant.

Back to \( \rho(z, t, TB) \), in the region where \( s(z, t) > 0 \) and \( z > z_0 \), we have, with the new definition of \( Q(t, TB) \):

\[
\rho(z, t, TB) = \frac{c(t)}{\mu(TB)} \frac{s_\infty \ c(t)}{k} (1 + \epsilon)^{1-\alpha} Q(t, TB) \exp \left( -U(z, t, TB) \right)
\]

(32)

\[
= \frac{c(t)}{\mu(TB)} \frac{s_\infty \ c(t)}{k} (1 + \epsilon)^{1-\alpha} Q(t, TB) \left( \epsilon + \frac{s_0(z, t)}{s_\infty} \right)^{\alpha} \exp \left( -\frac{c(t) z}{\mu(TB)} \right)
\]

(33)

\[
= (1 + \epsilon) \frac{c(t)}{\chi(TB) - \mu(TB)} N(t) \left( \frac{z_\infty \ c(t)}{\chi(TB) - \mu(TB)} \right)^{-1} \left( 1 + (\alpha - 1) \int_{TB}^\infty Q(t, TB) \exp \left( -\frac{c(t) z}{\mu(TB)} \right) dTB \right)^{-\frac{1}{\alpha-1}}.
\]

(34)

We chose not lump \( (\alpha - 1) \) into the definition of \( Q(t, TB) \). Keeping them separate tends to make the peak cell density in the group roughly coincide with \( z = 0 \), separating a diffusive region in the front of the group where there is almost no gradient \( (z > 0) \) and a chemotactic region in the middle/back of the group \( (z < 0) \). The solution above for \( \rho(z, t, TB) \) is only defined for \( z > z_0 \). At \( z_0 \), where the attractant \( s = 0 \), cell density is finite. We will use this later to make predictions about the leakage flux at the back of the wave. The dimensions of \( \rho(z, t, TB) \) come from the prefactor, \( \frac{c(t)}{\chi(TB) - \mu(TB)} N(t) \), which is roughly the width of the density profile of phenotype \( TB \).

The distribution of phenotypes traveling is contained in \( Q(t, TB) \), which is constrained by \( \int_{z_0}^\infty \rho(z, t, TB) \, dz = N(t, TB) \), where \( N(t, TB) \) is the number of cells of each phenotype traveling at time \( t \). Without loss of generality, we can write \( Q(t, TB) = P(t, TB) / Z(t, TB) \),

\[
\rho(z, t, TB) = (1 + \epsilon) \frac{c(t)}{\chi(TB) - \mu(TB)} N(t) \left( \frac{z_\infty \ c(t)}{\chi(TB) - \mu(TB)} \right)^{-1} \left( 1 + (\alpha - 1) \int_{TB}^\infty \frac{P(t, TB)}{Z(t, TB)} \exp \left( -\frac{c(t) z}{\mu(TB)} \right) dTB \right)^{-\frac{1}{\alpha-1}}.
\]

(35)
Using $\int_{z_0}^{\infty} \rho(z,t,TB) \, dz = N(t,TB)$, and given a traveling distribution $P(t,TB) = N(t,TB)/N(t)$, the quantity $Z(t,TB)$ (and hence $Q(t,TB)$) is fixed by the implicit solution to:

$$\int_{z_0}^{\infty} (1 + \epsilon) N(t) \, c(t) \, \frac{(\alpha - 1) P(t,TB)}{\chi(TB) - \mu(TB)} \exp \left(- \frac{c(t) z}{\mu(TB)} \right) \frac{a}{a-1} \, dz = N(t,TB).$$  \hspace{1cm} (36)

**Traveling gradient steepness**

We must be clear that these solutions, Eqns. (30) and (35), have serious issues. Most importantly, the wave is not at steady state, and time derivatives at the back cannot be ignored. As a result, cell density cannot be described by an energy potential (Eqn. (19)). Still, they are useful for understanding the properties and organization of the wave, particularly in the region where chemotaxis dominates, roughly corresponding to $z < 0$. We use them here to derive the steepness of the traveling gradient the cells chase.

First, the density of each phenotype can be written as:

$$\rho(z,t,TB) = \frac{N(t) \, c(t)}{\chi(TB) - \mu(TB)} \frac{(\alpha - 1) P(t,TB)}{\chi(TB) - \mu(TB)} \exp \left(- \frac{c(t) z}{\mu(TB)} \right) \left( \epsilon + \frac{s(z,t)}{s_\infty} \right).$$  \hspace{1cm} (37)

Towards the back of the group where chemotaxis dominates, the exponential term in Eqn. (35) is much larger than 1, giving:

$$\rho(z,t,TB) \sim \frac{c(t)}{\chi(TB) - \mu(TB)} N(t) \frac{P(t,TB)}{Z(t,TB)} \exp \left(- \frac{c(t) z}{\mu(TB)} \right) \left( \epsilon + \frac{s(z,t)}{s_\infty} \right).$$  \hspace{1cm} (38)

Next, the total cell density profile in the wave is

$$\rho(z,t) = \int_{TB} \rho(z,t,TB) \, dTB$$

$$\sim c(t) N(t) \frac{1}{\chi(TB) - \mu(TB)} \frac{P(t,TB)}{Z(t,TB)} \exp \left(- \frac{c(t) z}{\mu(TB)} \right) \left( \epsilon + \frac{s(z,t)}{s_\infty} \right).$$  \hspace{1cm} (40)

Using Eqns. (38) and (40), the local composition of phenotypes at position $z$, $P(TB|z,t)$, is:

$$P(TB|z,t) = \frac{\rho(z,t,TB)}{\rho(z,t)} = \frac{1}{\chi(TB) - \mu(TB)} \frac{P(t,TB)}{Z(t,TB)} \exp \left(- \frac{c(t) z}{\mu(TB)} \right) \frac{1}{\int_{TB} \frac{P(t,TB)}{\chi(TB) - \mu(TB)} \frac{P(t,TB)}{Z(t,TB)} \exp \left(- \frac{c(t) z}{\mu(TB)} \right) \, dTB}$$

and the mean value of $\chi(TB) - \mu(TB)$ among cells at position $z$ is:

$$\langle \chi - \mu | z, t \rangle = \int_{TB} (\chi(TB) - \mu(TB)) P(TB|z,t) \, dTB$$  \hspace{1cm} (42)
traveling together, relative abundances of different phenotypes at the back of the group. Say there are two phenotypes where the density of high performers is exponentially reduced by low performers. The gradient is steeper as the wave speed increases. This is because the cell density gradient scales with the wave speed, and the cell creates a "shallow" and "wide" dent. This is reflected above, where the perceived gradient steepness is inversely proportional to the length scale of the migrating group.

From Eqn. (35), we can readily read off that $L(z, t) = \frac{s(z, t)}{(\chi - \mu |z, t|) / c(t)}$ is the typical length scale of the group at position $z$. Furthermore, this form of the quasi-steady state solution makes it straightforward to determine the traveling gradient steepness. We will find that the local phenotype composition at position $z$ sets the local gradient steepness.

$$\lambda(z, t) = \partial_z \log \left( \frac{\rho(z, t)}{(\chi - \mu |z, t|) / c(t)} \right) = \frac{1}{(\epsilon + \frac{s(z, t)}{s_\infty})} \partial_z \left( \frac{s(z, t)}{s_\infty} \right)$$

From Eqn. (19) (with $\partial_z \to 0$), we have $\partial_z \left( \frac{s(z, t)}{s_\infty} \right) = \frac{1}{N(t)} \rho(z, t)$. Together with Eqn. (45), this means the local gradient steepness $\lambda(z, t)$ towards the back where chemotaxis dominates is:

$$\lambda(z, t) = \frac{c(t)}{(\chi - \mu |z, t|)}$$

which is exactly the inverse length scale of the migrating group, $L(z, t)$. The gradient length scale is the distance over which concentration drops, which is the width of the wave.

To understand this, remember that the gradient is created by the cells’ own consumption. Since all cells are assumed to consume the attractant at the same rate, a cell with low diffusivity creates a "deep" and "narrow" dent in the attractant concentration profile, while a high diffusivity cell creates a "shallow" and "wide" dent. This is reflected above, where the perceived gradient steepness at each location is inversely proportional to $(\chi - \mu |z, t|)$. Cells at the back have smaller $\chi(TB)$ and $\mu(TB)$; therefore, they make the gradient steeper there. Additionally, $\partial_z \log \left( \epsilon + \frac{s(z, t)}{s_\infty} \right)$ scales with the wave speed $c(t)$ – with a fixed phenotype composition, migrating faster requires that gradient is steeper.

**Low performers exponentially reduce the density of high performers at the back of the group**

We can also use the quasi-steady state solutions Eqns. (70) and (75) to get a sense for the relative abundances of different phenotypes at the back of the group. Say there are two phenotypes traveling together, $TB_1$ and $TB_2$. The density of $TB_1$ is:

$$\rho(z, t, TB_1) = \frac{(1 + c) \gamma(TB_1, t)}{\chi(TB_1, t) - \mu(TB_1)} \left[ (1 + (a - 1) \frac{P(TB_1)}{Z(TB_1)} \exp \left( - \frac{c(t) z}{\mu(TB_1)} \right) \right]$$

where $\gamma(TB_1, t)$ is the survival probability of $TB_1$ and $P(TB_1)$ is the probability of phenotype $TB_1$. To find the density of $TB_2$, we replace $TB_1$ with $TB_2$ in the above equation.
Towards the back of the wave, the exponential terms in the denominator dominate, so we can drop the 1:

\[
\rho(z, t, TB_1) \sim \frac{(1 + \epsilon) N(t) c(t)}{(\chi(TB_1) - \mu(TB_1)) (\alpha - 1)^{\frac{1}{\pi - 1}}} \left( \frac{P(t, TB)}{Z(t, TB)} \exp \left( - \frac{c(t) z}{\mu(TB)} \right) + \frac{(1 + \epsilon) N(t) c(t)}{(\chi(TB_1) - \mu(TB_1)) (\alpha - 1)^{\frac{1}{\pi - 1}}} \left( \frac{P(t, TB)}{Z(t, TB)} \exp \left( - \frac{c(t) z}{\mu(TB)} \right) \right) \right)
\]

(49)

We can write this as:

\[
\rho(z, t, TB_1) = \frac{(1 + \epsilon) N(t) c(t)}{(\chi(TB_1) - \mu(TB_1)) (\alpha - 1)^{\frac{1}{\pi - 1}}} \left( \frac{P(t, TB)}{Z(t, TB)} \exp \left( - \frac{c(t) z}{\mu(TB)} \right) + \frac{(1 + \epsilon) N(t) c(t)}{(\chi(TB_1) - \mu(TB_1)) (\alpha - 1)^{\frac{1}{\pi - 1}}} \left( \frac{P(t, TB)}{Z(t, TB)} \exp \left( - \frac{c(t) z}{\mu(TB)} \right) \right) \right)
\]

(50)

or:

\[
\rho(z, t, TB_1) = \frac{(1 + \epsilon) N(t) c(t)}{(\chi(TB_1) - \mu(TB_1)) (\alpha - 1)^{\frac{1}{\pi - 1}}} \left( \frac{P(t, TB)}{Z(t, TB)} \exp \left( - \frac{c(t) z}{\mu(TB)} \right) + \frac{(1 + \epsilon) N(t) c(t)}{(\chi(TB_1) - \mu(TB_1)) (\alpha - 1)^{\frac{1}{\pi - 1}}} \left( \frac{P(t, TB)}{Z(t, TB)} \exp \left( - \frac{c(t) z}{\mu(TB)} \right) \right) \right)
\]

(51)

Then, say \( \chi(TB_1) < \chi(TB_2) \) and \( \mu(TB_1) < \mu(TB_2) \). Since \( \frac{1}{\mu(TB_1)} > \frac{1}{\mu(TB_2)} \), towards the back the exponential in the denominator goes to zero, leaving:

\[
\rho(z, t, TB_1) \sim \frac{(1 + \epsilon) N(t) c(t)}{(\chi(TB_1) - \mu(TB_1)) (\alpha - 1)^{\frac{1}{\pi - 1}}} \left( \frac{P(t, TB)}{Z(t, TB)} \exp \left( - \frac{c(t) z}{(\alpha - 1) \mu(TB)} \right) \right)
\]

(52)

This is the behavior of cell density if there is just one phenotype traveling: towards the back, cell density decays with length scale

\[
L(TB_1) = \frac{(\alpha - 1) \mu}{c} = \frac{(\chi - \mu)}{c}
\]

(53)

Then, the density of the high-performing phenotype is:

\[
\rho(z, t, TB_2) = \frac{(1 + \epsilon) N(t) c(t)}{(\chi(TB_2) - \mu(TB_2)) (\alpha - 1)^{\frac{1}{\pi - 1}}} \left( \frac{P(t, TB)}{Z(t, TB)} \exp \left( - \frac{c(t) z}{(\alpha - 1) \mu(TB)} \right) \right)
\]

(54)

Towards the back, the exponential in the denominator gets large, giving:
\begin{equation}
\rho(x, t, TB_2) \sim \frac{(1 + e) N(t) c(t)}{(\chi(TB) - \mu(TB))((a - 1)z^{-1})} \left( \frac{P(t, TB)}{Z(t, TB)} \right)^{\frac{1}{a-1}} \exp \left( \frac{c(t) z}{(a - 1)\mu(TB)} \right)^2 \frac{\alpha}{a-1}.
\end{equation}

which simplifies to:

\begin{equation}
\rho(x, t, TB_2) = \frac{(1 + e) N(t) c(t) \left( \frac{P(t, TB)}{Z(t, TB)} \right)^{\frac{a}{a-1}} \exp \left( \frac{\alpha}{a - 1} \frac{c(t)}{\mu(TB)} - \frac{c(t)}{\mu(TB_2)} \right)^{1/2}}{(\chi(TB) - \mu(TB))((a - 1)^{1/2})^{1/2}} \exp \left( \frac{\alpha}{a - 1} \frac{c(t)}{\mu(TB_1) - \mu(TB_2)} \right)^{1/2}.
\end{equation}

Since \( \frac{1}{\mu(TB_2)} > \frac{1}{\mu(TB_1)} \) and \( \frac{\alpha}{a-1} > 1 \), this means that \( \left( \frac{\alpha}{a - 1} \frac{c(t)}{\mu(TB_1)} - \frac{c(t)}{\mu(TB_2)} \right) > 0 \), and the density of the high-performing phenotype decays towards the back (as \( z \to -\infty \)). The length scale of this decay is:

\begin{equation}
L(TB_2) = \frac{1}{(\alpha - 1) \frac{c(t)}{\mu(TB_2)}} = \frac{1}{\alpha - 1} \frac{c(t)}{\mu(TB_2)}.
\end{equation}

where the last step follows because \( \mu(TB_2) > \mu(TB_1) \) and \( \alpha > 1 \). This means that density profile of the high performing phenotype decays much faster when a lower performer is present, which exponentially reduces its density at the back.

**Leakage flux**

To derive an approximate leakage flux, we must break our assumption that the wave is at (quasi) steady state. By integrating the PDE for cell density from \( z = z_0 \) to \( z \to \infty \), keeping the time derivative term, we get that the rate of change of the number of cells in the wave is:

\begin{equation}
\partial_t N(t, TB) = \int_{z_0}^{\infty} \partial_z \rho(z, t, TB) \, dz = J(z_0, t, TB)
\end{equation}

At \( z_0 \), there is no attractant, so physically the chemotactic flux should be zero. This is the first break from the quasi-steady solution, which requires that there be chemotactic flux at \( z_0 \) in order for there to be no changes in time. Furthermore, in simulations we noticed that at \( z_0 \) there is a local minimum in cell density: behind \( z_0 \), are cells that were leaked at an earlier time, when the leakage flux was higher; ahead of \( z_0 \) is the increasing cell density of the wave. Therefore, we also neglect the diffusive flux at \( z_0 \). This leaves:

\begin{equation}
\partial_t N(t, TB) \sim - c(t) \rho(z_0, t, TB).
\end{equation}

From Eqns. (44) and (45), we can readily derive an approximate expression for this leakage using the quasi-steady solutions for \( \rho(z_0, t, TB) \). Using the fact that \( s(z_0, t) = 0 \) by definition, and plugging in Eqn. (44) gives:

\begin{equation}
\partial_t N(t, TB) \sim - \epsilon \frac{c(t)^2}{(\chi(TB) - \mu(TB)|z_0, t|)} P(TB|z_0, t) N(t).
\end{equation}
With the arguments in the previous section, the presence of low-performing phenotypes in the migrating group dramatically reduces the density of high performers at the back, thus dramatically reducing their leakage flux.

Finally, the total leakage flux is:

$$\partial_t N(t) = \int \partial_t F(t) \, dTB = -c(t) \rho(z_0, t) \sim -\epsilon \frac{c(t)^2}{\langle \chi(TB) - \mu(TB) \rangle(z_0, t)} N(t).$$  \hspace{1cm} (61)

**Growth model**

When a cell with phenotype $TB$ divides, what are the phenotypes of its daughter cells? We model the probability of a cell dividing in time $dt$ during exponential growth as a Poisson process. One could then construct a master equation with microscopic transitions that are proportional to the rates at which mother cells divide and the relative probabilities of producing two daughter cells with some phenotypes. Ultimately, however, only the average net growth matrix appears in the PDE for cell density through an additional term on the right-hand side:

$$\int R(TB, TB', z) \rho(z, t, TB') \, dTB'.$$  \hspace{1cm} (62)

For a discrete set of phenotypes, element $R_{ij}(z) = R(TB_i, TB_j, z)$ is the average net rate at which cells of phenotype $TB_j$ produce cells of phenotype $TB_i$. For example, dropping the $z$-dependence for now, on the diagonal, $R_{ii} = r_i \left( 2 \, p_{ii} + \sum_j p_{ij} + \sum_j p_{ji} - 1 \right)$ has four contributions. Here, $r_i$ is the rate at which phenotype $TB_i$ undergoes divisions. The first term is the probability that both daughter cells have the same phenotype as the mother, with the factor of 2 accounting for two cells being produced. The second and third terms are the probability that only one daughter cell has the same phenotype as the mother. The last term accounts for removal of the mother cell upon division. This formulation allows for cases where the two daughter cells are distinguishable, such as the "old pole" cell and the "new pole", and could have different, possibly correlated, statistics. The other entries of the matrix $R$ have terms that include the possibility that division of a cell with phenotype $TB_i$ produces two cells with phenotype $TB_j$. The sum of elements in each column of $R$ is $r_j$, the division rate of phenotype $TB_j$.

The structure of $R$ that arises from the microscopic transitions described above places constraints on its elements. The diagonal elements can be less than zero, but the remaining elements must be non-negative because division of one phenotype cannot cause removal of another phenotype. Additionally, the sum of all elements in a column, excluding the diagonal element, must be less than or equal to $2 \, r_j$, with equality occurring only when phenotype $j$ never produces offspring with the same phenotype as itself.

In general, the growth matrix $R$ depends on position $z$ through its dependence on the local concentration of nutrients. This can be simplified by assuming that the growth nutrients are abundant and distinct from the attractant, making the growth matrix independent of position: $R = R(TB, TB')$. Furthermore, we assume that all phenotypes grow exponentially with the same maximal rate, $r_j = r$. Additionally, the dynamics of cell density do not distinguish which daughter cell is which, and therefore the joint distribution of daughter phenotypes is symmetric. In this case, even if the daughter cell phenotypes are correlated, $R(TB, TB')$ only depends on the marginal distributions of each daughter cell's phenotype. To see this, we write

$$R_{ij} = r \left( 2 \, P(TB_1 = TB_i, TB_2 = TB_j | TB_i) + P(TB_1 = TB_i, TB_2 = TB_j | TB_i) \right)$$

$$+ P(TB_1 = TB_i, TB_2 = TB_j | TB_i) - \delta(TB_i, TB_j)$$  \hspace{1cm} (63)
with \( \delta(TB_i, TB_j) = 1 \) if \( TB_i = TB_j \) and \( \delta(TB_i, TB_j) = 0 \) otherwise. If the distribution of the two daughter phenotypes is symmetric, then \( P(TB_i = TB_i, TB_2 \neq TB_i|TB_i) = P(TB_i \neq TB_i, TB_2 = TB_i|TB_i) \). Then \( R_{ij} \) simplifies to:

\[
R_{ij} = r \left( 2 \left( P(TB_i = TB_i, TB_2 = TB_i|TB_j) + P(TB_i = TB_i, TB_2 \neq TB_i|TB_j) \right) - \delta(TB_i, TB_j) \right).
\]

But \( P(TB_i = TB_i, TB_2 = TB_i|TB_j) + P(TB_i = TB_i, TB_2 \neq TB_i|TB_j) \) is the marginalization of \( P(TB_i, TB_2|TB_j) \) over \( TB_2 \). Due to the symmetry between \( TB_1 \) and \( TB_2 \), this leaves:

\[
R_{ij} = r \left( 2 P(TB_i|TB_j) - \delta(TB_i, TB_j) \right).
\]

Symmetry of the joint distribution of the two daughter cell phenotypes makes their marginal distributions the same. As a result, even if the phenotypes of daughter cells are correlated, only their marginal distributions appear in \( R(TB, TB') \).

The PDE for cell density in the moving reference frame (Eqn. (10)) becomes:

\[
\frac{\partial \rho(x,t,TB)}{\partial t} - c(t) \frac{\partial \rho(x,t,TB)}{\partial x} = -\frac{\partial}{\partial x} \left( \frac{\partial}{\partial \phi} \log \left( e^{\frac{\phi(x,t)}{\sigma^2}} \rho(x,t,TB) - \mu(TB) \rho(x,t,TB) \right) \right) + r \int \left( 2 P(TB|TB') - \delta(TB - TB') \right) \rho(x,t,TB') \, dTB'.
\]

Next, we need a model for \( P(TB|TB') \). We expect that daughter cells should have phenotypes similar to their mothers, with some variation. Possibly the simplest model of growth with local inheritance of phenotype is an autoregressive (AR) process, where the phenotypes of daughter cells are Gaussian-distributed with means that revert back to the batch culture mean phenotype. This model is more appropriate for phenotypes defined over \(( -\infty, \infty) \), rather than the finite region \([0,1]\) on which \( TB \) is defined. Therefore, we consider an AR-type growth-inheritance model for \( F \) (7), defined by:

\[
TB = \frac{1}{1 + \exp(F)}, \quad F = \log \left( \frac{1 - TB}{TB} \right).
\]

\( F \) can be thought of as an effective free energy difference between the run and tumble states. As \( F \to \infty, TB \to 0 \); as \( F \to -\infty, TB \to 1 \). Since the mapping from \( F \) to \( TB \) is sigmoidal, cells with \( TB \) approaching zero or one are increasingly difficult to generate.

The AR model for the distribution of offspring phenotypes \( F \) given mother’s phenotype \( F' \) has 3 parameters: the mean \( \langle F \rangle \) and variance \( \sigma^2 \) of \( F \) in the batch culture distribution, and \( \phi \) the correlation between mother and daughter phenotypes. In this model, the offspring distribution is:

\[
P(F|F') = N \left( (1 - \phi) \langle F \rangle + \phi F', \sigma^2 (1 - \phi^2) \right)
\]

where \( N(\mu, \sigma^2) \) is a Gaussian distribution with mean \( \mu \) and variance \( \sigma^2 \). Again, since we assume the two daughter cell phenotype statistics are symmetric, we only need to specify their marginal distribution \( P(F|F') \). As \( \phi \to 1 \), daughter cells become identical to their mothers, and the population never reaches a steady state composition in batch culture. When \( \phi \to 0 \), daughter cells are sampled directly from the batch culture distribution.
To derive the relaxation time of the population’s composition during exponential growth in batch culture, we start with the dynamics of the numbers of cells of each phenotype \( N(F, t) \):

\[
\frac{dN(F, t)}{dt} = r \int_{-\infty}^{\infty} R(F, F') N(F', t) dF'.
\]  

(69)

where \( R(F, F') = 2 P(F|F') - \delta(F - F') \). Since the phenotype composition is \( P(F, t) = N(F, t)/N(t) \), where \( N(t) = \int_{-\infty}^{\infty} N(F', t) dF' \) is the total number of cells, the dynamics of composition follow:

\[
\frac{dP(F, t)}{dt} = 2 r \int_{-\infty}^{\infty} (P(F|F') - \delta(F - F')) N(F', t) dF'.
\]  

(70)

The stationary phenotype distribution during exponential growth in batch culture, \( P_0(F) \), is given by the leading eigenmode of \( P(F|F') \), which is a Gaussian distribution and has leading eigenvalue \( \nu_0 = 1 \). The relaxation time of \( P(F|F') \) is governed by the next eigenvalue, which can be shown to be \( \nu_1 = \phi \), i.e. the level of inheritance, with corresponding eigenmode \( P_1(F) = \frac{d}{dF} P_0(F) \). The relaxation time of composition dynamics in Eqn. (70) is therefore \( \tau^{-1} = 2 r (1 - \phi) \), and thus inheritance sets the relaxation time scale in batch culture.

**Steady state migration speed**

Steady state migration is reached when the number of cells of each phenotype produced per unit time is equal to the number lost per unit time. The production of cells in the group is integral of the growth term over the width of the wave, i.e. from \( z = z_0 \) to \( z \to \infty \). Since we assumed the growth matrix is independent of position, this reduces to

\[
[\partial_t N(t, TB)]_{growth} = \int_{z_0}^{\infty} \int R(TB, TB') \rho(z, t, TB') dTB' \, dz
\]  

(71)

\[
= \int R(TB, TB') N(t, TB') \, dTB'.
\]  

(72)

Then, the time evolution of \( N(t, TB) \) in the wave during travel has contributions from growth and leakage:

\[
\frac{dN(t, TB)}{dt} = [\partial_t N(t, TB)]_{growth} + [\partial_t N(t, TB)]_{leakage}.
\]  

(73)

Since the time scale of growth, \( 1/r \), is typically slow compared to the time scale of wave migration, \( (\chi(TB) - \mu(TB))/c(t)^2 \) (14), the quasi-stationary expression for \( \rho(z, t, TB) \) and the leakage expression derived above are expected to be a good approximation in the presence of growth. Using these, we get:

\[
\frac{dN(t, TB)}{dt} = \int R(TB, TB') N(t, TB') \, dTB' - \frac{c(t)^2}{(\chi(TB) - \mu(TB))|z_0, t|} \frac{P(TB|z_0, t)}{P(TB, t)} N(t, TB)
\]  

(74)

\[
= \int r \left( 2 P(TB|TB') - \delta(TB - TB') \right) N(t, TB') \, dTB'
\]  

(75)

\[
- \frac{c(t)^2}{(\chi(TB) - \mu(TB))|z_0, t|} \frac{P(TB|z_0, t)}{P(TB, t)} N(t, TB)
\]

The total number of cells is the integral of the above expression over phenotypes \( TB \):

13
\[
\frac{dN(t)}{dt} = \int 2 r \int P(TB|TB') N(t, TB') dTB' \ dTB - r N(t) c(t)^2 \epsilon \frac{\langle \chi(TB) - \mu(TB) \rangle_{z_0, t}}{z_0, t} N(t).
\]

Integrating the first term with respect to \(TB\) first results in the following simple expressions:

\[
\frac{dN(t)}{dt} = 2 r \int N(t, TB') dTB' - r N(t) - \epsilon \frac{c(t)^2}{\langle \chi(TB) - \mu(TB) \rangle_{z_0, t}} N(t). \tag{77}
\]

At steady state, growth and leakage balance

\[
\frac{dN(t)}{dt} = 0 \rightarrow r N_{ss} = \epsilon \frac{c_{ss}^2}{\langle \chi(TB) - \mu(TB) \rangle_{z_0, ss}} N_{ss}, \tag{79}
\]

and solving for \(c_{ss}\):

\[
c_{ss} = \left( \frac{r}{\epsilon} \langle \chi(TB) - \mu(TB) \rangle_{z_0, ss} \right)^{1/2}. \tag{80}
\]

We find that the steady state traveling speed is set by the average performance of phenotypes at the back of the wave at steady state, \(\langle \chi(TB) - \mu(TB) \rangle_{z_0, ss}\). Although this quantity is difficult to predict, it is clear that if the composition of the migrating group shifts such that minimum performance in higher, then the wave speed at steady state will be higher.

This theoretical expression is consistent with recent measurements (9) for the dependence of wave speed on growth rate \(r\) and its lack of dependence on the consumption rate of attractant \(k\) (assuming that the wave composition was similar across those experiments). Recent theory that treated all cells as identical and considered attractant diffusion found the same dependence on \(r\), \(\epsilon\), and \(k\) (14). We extend the expression in Eqn. (80) to account for attractant diffusion in the following section, **Steady state migration speed with attractant diffusion**.

Finally, as recently found for migrating populations consisting of a single phenotype, the density of cells left behind diverse migrating waves at steady state also has a simple form. At steady state, the flux of cells off the back is balanced by the total flux of cells produced by growth:

\[
c_{ss} \rho_{ss}(z_0, TB) = r N_{ss} \rightarrow \rho_{ss}(z_0, TB) = r \frac{N_{ss}}{c_{ss}} = r \frac{s_{ss}}{k}, \tag{81}
\]

where we used \(c_{ss} = k \ N_{ss} / s_{ss}\). Surprisingly, the total density of cells left behind is independent of the steady state phenotype composition of cells traveling when growth nutrients are in excess within the group, and therefore is also independent of phenotypic inheritance and independent of which physical environment the cells are migrating through.

**Steady state migration speed with attractant diffusion**

As noted recently (14), attractant diffusion can have significant effects on the migrating wave when the attractant diffusivity \(D_e\) approaches or exceeds \(\chi(TB) - \mu(TB)\) among the cells traveling. This can occur in porous environments, where all phenotypes’ chemotactic coefficients and diffusivities are lower than in liquid, but the attractant diffusivity is essentially unchanged. The
steady state ODEs we will consider are slightly different from those in earlier sections. In addition to keeping attractant diffusion, the form of the consumption term will be that used in (14):

\[-c_{ss} \partial_z \left( \frac{s}{s_{\infty}} \right) = D_s \partial_{zz} \left( \frac{s}{s_{\infty}} \right) - k \frac{s}{s_{\infty}} \left( \frac{s}{s_{\infty}} + \epsilon \right) \rho.\]

(82)

Throughout, the $z$-dependence of $\rho$ and $s$ are understood, but not shown, $\rho$ is total cell density, $\rho(TB)$ is the density of cells with phenotype $TB$, and $c_{ss}$ is the steady state migration speed. Importantly, this consumption term has a Monod-like form with a half-maximum consumption constant $K_{m}$ that we set equal to $K_1$, the chemotactic sensitivity for the attractant. Thus, $K_{m}/s_{\infty} = K_1/s_{\infty} = \epsilon$.

Our approach in this section will closely follow that of (14), but with non-genetic diversity of swimming phenotypes. We focus on the back region of the migrating group, where chemotaxis dominates over cell diffusion, and make the ansatz that the attractant has the following form in the chemotaxis region:

\[\frac{s}{s_{\infty}} = A \exp \left( \int \lambda(z) \, dz \right).\]

(83)

Plugging the attractant ansatz into the attractant ODE, Eqn. (81) gives:

\[c_{ss} \lambda(z) \frac{s}{s_{\infty}} + D_s (\partial_z \lambda(z) + \lambda(z)^2) \frac{s}{s_{\infty}} = \frac{k}{s_{\infty}} \frac{\rho}{\epsilon + \frac{s}{s_{\infty}}} \frac{s}{s_{\infty}}\]

(84)

\[c_{ss} \lambda(z) \frac{s}{s_{\infty}} + D_s \lambda(z) \left( \partial_z \log(\lambda(z)) + \lambda(z) \right) \frac{s}{s_{\infty}} = \frac{k}{s_{\infty}} \frac{\rho}{\epsilon + \frac{s}{s_{\infty}}} \frac{s}{s_{\infty}}.\]

(85)

Next, we argue that $\lambda(z) \gg \partial_z \log(\lambda(z))$. Later, we’ll see that $\lambda(z) = \frac{c_{ss}}{\lambda - \mu |z|}$, making $\partial_z \log(\lambda(z)) = -\partial_z \log(\mu |z| - \langle \mu |z| \rangle)$. With this, the assumption that $\lambda(z) \gg \partial_z \log(\lambda(z))$ means that although chemotaxis abilities in the group vary with position, but they vary on a length scale that is longer than the width of group itself, $\sim 1/\lambda(z)$. If the opposite were true, i.e. $\lambda(z) \ll \partial_z \log(\lambda(z))$, then cells throughout the group would experience almost the same gradient. In this case, there would be no compensatory mechanism keeping them traveling together, so a stable traveling wave cannot form in this regime. We also argue that it is not the case $\lambda(z) \sim \partial_z \log(\lambda(z))$, because empirically in simulations, chemotaxis abilities in the group do not change by orders of magnitude in the region where chemotaxis dominates. Thus, we take $\lambda(z) \gg \partial_z \log(\lambda(z))$, and we simplify the ODE above to:

\[c_{ss} \lambda(z) + D_s \lambda(z)^2 = \frac{k}{s_{\infty}} \frac{\rho}{\epsilon + \frac{s}{s_{\infty}}}.\]

(86)

We’ll return to this equation later.

Next, we consider the ODE for $\rho(TB)$, reproduced here:

\[-c_{ss} \partial_z \rho(TB) = -\partial_z \left( \chi(TB) \frac{\rho(TB)}{\epsilon + \frac{s}{s_{\infty}}} \partial_z \left( \frac{s}{s_{\infty}} \right) - \mu(TB) \partial_z \rho(TB) \right)\]

\[+ r \int R(TB|TB') \rho(TB') \, dTB'.\]

(87)
We first integrate this equation over phenotypes $TB$.

\[
-c_{ss} \partial_z \rho = -\partial_z \left( \int \chi(TB) \rho(TB) \, dTB \frac{\partial_z \left( \frac{s}{S_{so}} \right)}{\epsilon + \frac{s}{S_{so}}} - \int \mu(TB) \partial_z \rho(TB) \, dTB \right) + r \int \int R(TB|TB') \rho(TB') \, dTB' \, dTB.
\]  \hspace{1cm} (88)

To simplify this equation, we first note that $P(TB|z) = \rho(TB)/\rho$. With this,

\[
\int \chi(TB) \rho(TB) \, dTB = \rho \int \chi(TB) \rho(TB) \rho \, dTB = \rho \langle \chi|z \rangle.
\]  \hspace{1cm} (89)

Furthermore, we can simplify the last term by changing the order of integration with respect to $TB$ and $TB'$, and using the fact that $\int R(TB|TB') \, dTB = 1$ for all $TB'$:

\[
\int \left( \int R(TB|TB') \rho(TB') \, dTB' \right) \, dTB = \int \rho(TB') \, dTB' = \rho.
\]  \hspace{1cm} (90)

Using these simplifications, Eqn. (88) becomes:

\[
0 = -\partial_z \left( \langle \chi|z \rangle \frac{\rho}{\epsilon + \frac{s}{S_{so}}} \partial_z \left( \frac{s}{S_{so}} \right) - c_{ss} \rho - \int \mu(TB) \partial_z \rho(TB) \, dTB \right) + r \rho.
\]  \hspace{1cm} (92)

To simplify the term $\int \mu(TB) \partial_z \rho(TB) \, dTB$, note that:

\[
\partial_z \rho(TB|z) = \partial_z \frac{\rho(TB)}{\rho} = \frac{1}{\rho} \partial_z \rho(TB) - \frac{\rho(TB)}{\rho^2} \partial_z \rho = \frac{1}{\rho} \partial_z \rho(TB) - \frac{1}{\rho} P(TB|z) \partial_z \rho
\]  \hspace{1cm} (93)

And therefore

\[
\partial_z \rho(TB) = \rho \partial_z P(TB|z) + P(TB|z) \partial_z \rho.
\]  \hspace{1cm} (94)

With this, we get

\[
\int \mu(TB) \partial_z \rho(TB) \, dTB = \int \mu(TB) \left( \rho \partial_z P(TB|z) + P(TB|z) \partial_z \rho \right) \, dTB
\]  \hspace{1cm} (95)

\[
= \rho \int \mu(TB) \partial_z P(TB|z) \, dTB + \langle \mu(TB) \rangle \partial_z \rho
\]  \hspace{1cm} (96)

\[
= \rho \partial_z \langle \mu(z) \rangle + \langle \mu(z) \rangle \partial_z \rho
\]  \hspace{1cm} (97)

Plugging this into Eqn. (92):

\[
0 = -\partial_z \left( \langle \chi|z \rangle \frac{\rho}{\epsilon + \frac{s}{S_{so}}} \partial_z \left( \frac{s}{S_{so}} \right) - c_{ss} \rho - \left( \rho \partial_z \langle \mu(z) \rangle + \langle \mu(z) \rangle \partial_z \rho \right) \right) + r \rho
\]  \hspace{1cm} (98)

\[
0 = -\partial_z \left( \langle \chi|z \rangle \frac{\rho}{\epsilon + \frac{s}{S_{so}}} \partial_z \left( \frac{s}{S_{so}} \right) - c_{ss} \rho - \langle \mu(z) \rangle \rho (\partial_z \log(\langle \mu(z) \rangle) + \partial_z \log(\rho)) \right) + r \rho
\]  \hspace{1cm} (99)
Plugging in the attractant ansatz:

\[ 0 = -\partial_z \left( \langle \chi | z \rangle \frac{\rho}{\epsilon + \frac{s}{s_{\infty}}} \lambda(z) \frac{s}{s_{\infty}} - c_{ss} \rho - \langle \mu | z \rangle \rho \left( \partial_z \log(\langle \mu | z \rangle) + \partial_z \log(\rho) \right) \right) + r \rho. \]  

(100)

Next, we make the following ansatz for total cell density: \( \rho = \beta(z) \left( \epsilon + \frac{s}{s_{\infty}} \right) \). With this ansatz, \( \partial_z \log(\rho) = \partial_z \log(\beta(z)) + \frac{1}{(\epsilon + \frac{s}{s_{\infty}})} \lambda(z) \frac{s}{s_{\infty}} \). Later, we will see that \( \partial_z \log(\beta(z)) \ll \lambda(z) \), and therefore:

\[ \partial_z \log(\rho) \approx \frac{1}{(\epsilon + \frac{s}{s_{\infty}})} \lambda(z) \frac{s}{s_{\infty}} = O(\lambda(z)). \]  

(101)

We anticipate, from our derivations in earlier sections, that the gradient steepness \( \lambda(z) \) will be:

\[ \lambda(z) = \frac{c_{ss}}{\langle \chi - \mu | z \rangle}. \]  

(102)

Furthermore, if \( \chi/\mu \) is constant among all phenotypes, which we assumed earlier, we have:

\[ \partial_z \log(\langle \mu | z \rangle) = \partial_z \log \left( \frac{\langle \chi - \mu | z \rangle}{c_{ss}} \right) \]  

(103)

\[ = \partial_z \log \left( \frac{\langle \chi | z \rangle - \langle \mu | z \rangle}{c_{ss}} \right), \]  

(104)

\[ = -\partial_z \log(\langle \chi | z \rangle). \]  

(105)

With our claim that \( \partial_z \log(\lambda(z)) \ll \lambda(z) \), we then have \( \partial_z \log(\langle \mu | z \rangle) \ll \lambda(z) \), and therefore \( \partial_z \log(\langle \mu | z \rangle) \ll \partial_z \log(\rho) \). As a result, we drop \( \partial_z \log(\langle \mu | z \rangle) \) from the ODE for \( \rho \):

\[ 0 = -\partial_z \left( \langle \chi | z \rangle \frac{\rho}{\epsilon + \frac{s}{s_{\infty}}} \lambda(z) \frac{s}{s_{\infty}} - c_{ss} \rho - \langle \mu | z \rangle \rho \partial_z \log(\rho) \right) + r \rho. \]  

(106)

Plugging in the ansatz \( \rho = \beta(z) \left( \epsilon + \frac{s}{s_{\infty}} \right) \) and \( \partial_z \log(\rho) \sim \frac{1}{(\epsilon + \frac{s}{s_{\infty}})} \lambda(z) \frac{s}{s_{\infty}} \):

\[ 0 = -\partial_z \left( \langle \chi | z \rangle \beta(z) \lambda(z) \frac{s}{s_{\infty}} - c_{ss} \beta(z) \left( \epsilon + \frac{s}{s_{\infty}} \right) - \langle \mu | z \rangle \beta(z) \lambda(z) \frac{s}{s_{\infty}} \right) + r \beta(z) \left( \epsilon + \frac{s}{s_{\infty}} \right) \]  

(107)

Rearranging terms:

\[ 0 = -\partial_z \left( \beta(z) \left( (\langle \chi | z \rangle - \langle \mu | z \rangle) \lambda(z) \frac{s}{s_{\infty}} - c_{ss} \left( \epsilon + \frac{s}{s_{\infty}} \right) \right) \right) + r \beta(z) \left( \epsilon + \frac{s}{s_{\infty}} \right) \]  

(108)

Next, we plug in the expected form of the gradient steepness, \( \lambda(z) = \frac{c}{\langle \chi | z \rangle - \langle \mu | z \rangle} \):

\[ 0 = -\partial_z \left( \beta(z) \left( (\langle \chi | z \rangle - \langle \mu | z \rangle) \frac{s}{s_{\infty}} - c_{ss} \left( \epsilon + \frac{s}{s_{\infty}} \right) \right) \right) + r \beta(z) \left( \epsilon + \frac{s}{s_{\infty}} \right) \]
\[ 0 = -\partial_z \left( \beta(z) \left( \frac{c_{ss}}{S_{ss}} - \frac{c_{ss}}{S_{ss}} \left( \epsilon + \frac{S}{S_{ss}} \right) \right) + r \beta(z) \left( \epsilon + \frac{S}{S_{ss}} \right) \right) + r \beta(z) \geq \frac{S}{S_{ss}} \]  
\quad (109)

\[ \partial_z \beta(z) = - \frac{r}{c_{ss} \epsilon} \beta(z) \left( \epsilon + \frac{S}{S_{ss}} \right) \]  
\quad (110)

\[ \partial_z \log(\beta(z)) = - \frac{r}{c_{ss} \epsilon} \left( \epsilon + \frac{S}{S_{ss}} \right) \]  
\quad (111)

Eqn. (111) above indicates that \( \partial_z \log(\beta(z)) \ll \lambda(z) \) specifically in the region where \( \frac{s}{S_{ss}} \approx \epsilon \), because in that region \( \partial_z \log(\beta(z)) \sim \frac{r}{c_{ss}} \ll \lambda(z) \) (as in (14)). This is enough for our purposes, because this is the region of interest for leakage. This also means that our guess \( \lambda(z) = \frac{c_{ss}}{(\chi[z]-\mu[z])} \) is valid in that region – there are no contradictions.

Next, we integrate the ODE for total cell density \( \rho \) Eqn. (98) over space from \( z \) to \( z \to \infty \). Although the ansatz is no longer valid as \( z \to \infty \), it is only a local approximation of the actual function \( \rho \). Therefore, we apply the boundary conditions in Eqn. (12) to get:

\[ c_{ss} \rho = (\chi[z]) \frac{\rho}{\epsilon + \frac{S}{S_{ss}}} \partial_z \frac{S}{S_{ss}} - (\rho \partial_z \mu[z] + (\mu[z] \partial_z \rho) + r N(z). \]  
\quad (112)

All of the calculations that we made inside the \( z \)-derivative in the section above still apply. With them, we get:

\[ 0 = \beta(z) \left( (\chi[z] - \mu[z]) \lambda(z) \right) \frac{\rho}{\epsilon + \frac{S}{S_{ss}}} \partial_z \frac{S}{S_{ss}} - \rho \beta(z) \lambda(z) \partial_z \left( \epsilon + \frac{S}{S_{ss}} \right) + r N(z), \]  
\quad (113)

where \( N(z) \) is the cumulative number of cells in the traveling group ahead of position \( z \).

Substituting for \( \lambda(z) \):

\[ 0 = \beta(z) \left( (\chi[z] - \mu[z]) \frac{S}{S_{ss}} - c_{ss} \left( \epsilon + \frac{S}{S_{ss}} \right) \right) + r N(z), \]  
\quad (114)

\[ \beta(z) = \frac{r}{c_{ss} \epsilon} N(z) \]  
\quad (115)

Again, this is valid at the back, where \( \frac{s}{S_{ss}} \approx \epsilon \). Evaluating both sides at the back of the wave \( z = z_0 \) and use \( N(z_0) = N \), the total number of cells traveling:

\[ \beta(z_0) \approx \frac{r}{c_{ss} \epsilon} \frac{N}{s_{ss}} \]  
\quad (116)

Here, we define \( z_0 \) as the location where \( \frac{s(z_0)}{S_{ss}} = \epsilon \). We will see below that this captures the total number of cells to leading order.

Next, we integrate the attractant ODE from \( z = z_0 \) to \( z \to \infty \) and apply the boundary conditions in Eqn. (12) again:

\[ -c_{ss} (1 - \epsilon) = D_s \int_{z_0}^{\infty} \left( \partial_z \left( \frac{S}{S_{ss}} \right) - k \frac{S}{S_{ss}} \left( \frac{S}{S_{ss}} + \epsilon \right) \rho \right) dz. \]  
\quad (117)
Using the boundary conditions, the ansatz for the attractant, and the definition \( s(z_0) = \epsilon \), we can simplify the first term on the right-hand side:

\[-c_{ss} \left( 1 - \epsilon \right) = -D_s \lambda(z_0) \epsilon - \int_{z_0}^\infty k \frac{S}{S_{\infty}} \left( \frac{S}{S_{\infty}} + \epsilon \right) \rho \, dz. \tag{118}\]

The last term can be split into two:

\[-c_{ss} \left( 1 - \epsilon \right) = -D_s \lambda(z_0) \epsilon - \left( \int_{z_0}^\infty k \frac{S}{S_{\infty}} \rho \, dz - \int_{z_0}^\infty k \frac{\epsilon}{S_{\infty}} \rho \, dz \right) \tag{119}\]

\[-c_{ss} \left( 1 - \epsilon \right) = -D_s \lambda(z_0) \epsilon - \left( \int_{z_0}^\infty k \frac{S}{S_{\infty}} N - \int_{z_0}^\infty k \left( \frac{S}{S_{\infty}} + \epsilon \right) \rho \, dz \right) \tag{120}\]

The last term here is small (i.e. \( O(\epsilon) \)) because in regions where \( \frac{S}{S_{\infty}} \sim O(\epsilon) \), we also have \( \rho \sim O(\epsilon) \), and in regions where \( \frac{S}{S_{\infty}} \gg \epsilon \), the term in the parenthesis is \( O(\epsilon) \). Keeping only leading-order terms, we get:

\[c_{ss} = \frac{k}{S_{\infty}} N, \tag{121}\]

as in Keller and Segel’s analysis.

Plugging this into Eqn. (116), we eliminate \( N \) and solve for \( \beta(z_0) \):

\[\beta(z_0) = \frac{r \, S_{\infty}}{\epsilon \, k}. \tag{122}\]

Finally, we return to the ODE for the attractant \( s \), Eqn. (86). Plugging in the ansatz for \( \rho \):

\[c_{ss} \lambda(z) + D_s \lambda(z)^2 = \frac{k}{S_{\infty}} \beta(z) \tag{123}\]

Evaluating both sides at \( z = z_0 \) and plugging in expression for \( \lambda(z_0) \):

\[\frac{c_{ss}^2}{\langle \chi - \mu | z_0 \rangle} + D_s \left( \frac{c_{ss}}{\langle \chi - \mu | z_0 \rangle} \right)^2 = \frac{r}{\epsilon} \tag{124}\]

Solving for the steady state migration speed \( c_{ss} \), we find that

\[c_{ss} = \left( \frac{r}{\epsilon} \frac{\langle \chi - \mu | z_0 \rangle^2}{\langle \chi - \mu | z_0 \rangle + D_s} \right)^{1/2}. \tag{125}\]

Note that both Eqn. (31) and Eqn. (121) show that the migration speed \( c \) at any given time is proportional to the number of cells traveling \( N \), but they differ in how the back of the group \( z_0 \) is defined, and therefore differ in terms of which cells count towards \( N \) because \( N \equiv \int_{z_0}^\infty \rho \, dz \). When \( K_m \rightarrow 0 \), the back of the group \( z_0 \) is defined by \( s(z_0) = 0 \), and in simulations this definition of the back does indeed make \( N = c \, s_{\infty} / k \). When \( K_m = K_i \), instead defining \( z_0 \) such that \( s(z_0) = K_m = K_i \) makes \( N = c \, s_{\infty} / k \). For more general \( K_m \) and \( K_i \), we find empirically in simulations that \( N = c \, s_{\infty} / k \).
results when the back of the group \( z_0 \) is defined by \( s(z_0) = \sqrt{K_m K_i} \). This captures both of the case we considered analytically here (\( K_i = K_m \)), as well as the Keller-Segel case of \( K_i \to 0 \) (in which \( z_0 \to -\infty \)).
**Fig. S1. Simulation profiles without growth.** Shown are simulation results for A) total cell density, B) attractant concentration, and C) perceived gradient steepness \((\partial_z \log(1 + s/K_i))\), in the moving frame of reference \(z = x - \int_0^t c(t') \, dt'\), in liquid (top) and porous (bottom) environments, at the last time point of simulations without growth. These are the same simulations used to generate the plots in Fig. 1 of the main text.
**Fig. S2. Spatial organization is environment-dependent.** Shown are simulation results for the spatial organization of various quantities. **A)** Cells in the migrating group sort by chemotactic performance, regardless of environment (throughout, top: liquid, bottom: porous). Plotted is the position of the peak cell density of each phenotype $T_B$ versus its chemotactic performance $\chi$. **B)** Peak cell density of each phenotype $T_B$ versus $TB$, in each environment. **C)** Mean phenotype, $TB$, at each position $z$ in liquid and agar. **D)** Mean motility $\langle \chi - \mu \rangle$ among cells at position $z$. These are the same simulations used to generate the plots in Fig. 1 of the main text.
Fig. S3. Comparison of total leakage flux in theory and simulations. The theoretical expression for total leakage $\partial_t N(t)$ (main text Eqn. (1), SI Eqn. (61); black dashed lines) captures dependence of total leakage flux on time-dependent quantities (green line: liquid simulation; orange line: porous simulation). Colors of dots match the time points at which phenotype distributions are shown in main text Fig. 1C. To generate the theoretical curves (black dashed line), $Z(t, TB)$ was computed numerically using the implicit Eqn. (36) at each point in time, taking the migration speed $c(t)$ and composition $P(t, TB)$ as given from the simulation. $N(t, TB)$, and thus $P(t, TB)$, in simulations was computed by integrating cell density over the wave from $z = z_0$ to $z \rightarrow \infty$ (where $s(z_0) = \sqrt{K_m K_i}$ defines the back of the wave; see Steady state migration speed with attractant diffusion). This determined $Z(t, TB)$ up to a constant, which sets the location of $z = 0$ in the moving reference frame. To fix the constant, $Z(t, TB)$ for the highest-performing phenotype was pinned at 1. We used Matlab’s `fmincon` function to find $\log(N(t, TB))$ numerically, since $Z(t, TB)$ varied by orders of magnitude at each time point. To help the optimization routine find solutions, we imposed a constraint on $Z(t, TB)$ for consecutive values of $TB$: if $\chi(TB_i) > \chi(TB_{i-1})$, then we required that $Z(t, TB_i) < Z(t, TB_{i-1})$. Optimization was performed in two rounds for each time point. First, the sum of squared errors between $N(t, TB)$ in the simulation and the analytical solution was minimized; then the sum of squared errors of $\log(N(t, TB))$ was minimized to better capture $N(t, TB)$ for low-abundance phenotypes, especially those at the back. For the first time point of the simulation, the optimization routine was initialized at $Z(t, TB) = 1$ for all $TB$, since this is the solution for one phenotype. In following time points, the optimization was initialized at the solution $Z(t, TB)$ found for the previous time point, since $Z(t, TB)$ is expected to change smoothly with time.
Fig. S4. Wild type distribution of phenotypes $TB$ in batch culture. The tumble bias distribution $P(TB)$ of wild-type *E. coli* cells is well-described by a Gaussian distribution of the transformed phenotype $F = \log \left( \frac{1 - TB}{TB} \right)$. Tumble bias distributions were measured by tracking RP437 cells (blue, left; shading throughout is standard error found by bootstrapping). Cells picked from a -80°C freezer stock were grown in 1X M9 salts + 0.1% w/v casamino acids + 0.4% v/v glycerol + 1 mM MgSO4 + 0.05% w/v PVP-40 (1) at 30°C overnight, and then diluted 100X in the same media the following morning, and harvested at OD600 = 0.25. The cells were then washed three times in chemotaxis buffer (1X M9 salts + 10 µM methionine + 100 µM EDTA + 0.05% w/v PVP-40) (1), injected into a 100 um-deep microfluidic chamber, and tracked with a 4X objective by phase contrast imaging at 20 frames per second. Three biological replicates were performed, and only trajectories longer than 10 seconds were kept, resulting in 5,825 trajectories with mean duration of 20 seconds. Tumbles were detected using angular speed and swimming speed variations, as described in refs (21,24). From $P(TB)$ (blue, A), the distribution $P(F)$ was computed using a change of variables formula (blue, B). $P(F)$ was approximated by a Gaussian distribution with the same mean and variance (orange, B). This distribution was then transformed to a distribution of tumble bias by inverting the change of variable formula (left, A). The mean of $P(F)$ was $\langle F \rangle = 1.2$ and the standard deviation was $\sigma_F = 0.46$. Simulations throughout used a different mean phenotype (i.e. the generalist phenotype described in the main text), but the same amount of variation, $\sigma_F$. 
Fig. S5. Composition adaptation rate in migrating groups versus composition relaxation rate in batch culture. Black dashed line shows $y = x$, and $r$ is the growth rate. Batch culture relaxation rate is $\tau = 1/(2r (1 - \phi))$. 
Fig. S6. Average migration speeds in changing environments. A) Normalized average migration speeds of a diverse population in changing environments where the standard deviation of the population’s tumble bias distribution is twice that in the main text (\( \sigma_F \) increased by \( \sim 1.5 \) times). \( r \) is the growth rate and \( T \) is the environment duration. B) Same as the middle panel in main text Fig. 3, but with the color limits of (A) for comparison. C) The same as (A), but plotted with inheritance \( \phi \) on the y-axis. D) Same as the middle panel in main text Fig. 3, but plotted with inheritance \( \phi \) on the y-axis.
**Table S1.** Parameter values used in simulations.

| Parameter | Value                  | Meaning                                             | Source          |
|-----------|------------------------|-----------------------------------------------------|-----------------|
| \( k \)  | 0.126 \( \mu M \) 0 \( \text{D}^{-1} \) \( \text{s}^{-1} \) | Attractant consumption rate                         | (1)             |
| \( D_s \) | 0 or 800 \( \mu m^2/s \) | Attractant diffusivity                              | (25,26)         |
| \( K_m \) | 0.5 \( \mu M \)       | Monod constant                                      | (1,12)          |
| \( s_\infty \) | 100 \( \mu M \)    | Attractant concentration ahead of the wave          | (1,9,27)        |
| \( K_i \) | 3.5 \( \mu M \)       | Receptor lower sensing limit                        | This study, (1,2)|
| \( K_a \) | 55 \( mM \)           | Receptor upper sensing limit                        | (1)             |
| \( v_0 \) | 25 \( \mu m/s \)      | Cell swimming speed                                 | E.g. (5,21,28) |
| \( \omega \) | 1.3 \( \text{s}^{-1} \) | Cell tumble rate parameter                         | (4,5,11,29)    |
| \( D_r \) | 0.062 \( \text{rad}^2/\text{s} \) | Cell rotational diffusion coefficient               | (30)            |
| \( \theta_p \) | 0.33 \( \text{s}^{-1} \) | Persistence of tumbles                              | (30)            |
| \( l \) | 30 \( \mu m \)        | Average pore size                                   | This study      |
| \( \alpha_{\text{liquid}} \) | 10 \( (\chi(TB)/\mu(TB)) \) \( \text{in liquid} \) |                                          | This study      |
| \( \alpha_{\text{porous}} \) | 6.25 \( (\chi(TB)/\mu(TB)) \) \( \text{in porous media} \) |                                      | (10)            |
| \( OD_0 \) | 0.5, 3, or 6 OD        | Initial cell density                                | This study      |
| \( \langle F \rangle \) | 1.8 \( \langle F \rangle \) \( \text{in liquid} \) |                                          | This study      |
| \( \sigma \) | 0.46 \( \text{Batch culture phenotype standard deviations} \) |                        | This study; Fig. S4 |
| \( \phi \) | Varying                | Mother-daughter correlation of phenotype \( F \)   | This study      |
| \( r \) | 0.83 \( hr^{-1} \)    | Growth rate                                         | E.g. (9)       |
SI References

1. Fu X, Kato S, Long J, Mattingly HH, He C, Vural DC, et al. Spatial self-organization resolves conflicts between individuality and collective migration. Nature Communications. 2018 Jun 5;9(1):2177.

2. Mello BA, Tu Y. Quantitative modeling of sensitivity in bacterial chemotaxis: The role of coupling among different chemoreceptor species. PNAS. 2003 Jul 8;100(14):8223–8.

3. Lovely PS, Dahlquist FW. Statistical measures of bacterial motility and chemotaxis. Journal of Theoretical Biology. 1975 Apr 1;50(2):477–96.

4. Dufour YS, Fu X, Hernandez-Nunez L, Emonet T. Limits of Feedback Control in Bacterial Chemotaxis. PLOS Computational Biology. 2014 Jun 26;10(6):e1003694.

5. Waite AJ, Frankel NW, Dufour YS, Johnston JF, Long J, Emonet T. Non-genetic diversity modulates population performance. Molecular Systems Biology. 2016 Dec 1;12(12):895.

6. Sneddon MW, Pontius W, Emonet T. Stochastic coordination of multiple actuators reduces latency and improves chemotactic response in bacteria. PNAS. 2012 Jan 17;109(3):805–10.

7. Long J, Zucker SW, Emonet T. Feedback between motion and sensation provides nonlinear boost in run-and-tumble navigation. PLOS Computational Biology. 2017 Mar 6;13(3):e1005429.

8. Licata NA, Mohari B, Fuqua C, Setayeshgar S. Diffusion of Bacterial Cells in Porous Media. Biophysical Journal. 2016 Jan 5;110(1):247–57.

9. Cremer J, Honda T, Tang Y, Wong-Ng J, Vergassola M, Hwa T. Chemotaxis as a navigation strategy to boost range expansion. Nature. 2019 Nov;575(7784):658–63.

10. Liu W, Cremer J, Li D, Hwa T, Liu C. An evolutionarily stable strategy to colonize spatially extended habitats. Nature. 2019 Nov;575(7784):664–8.

11. Frankel NW, Pontius W, Dufour YS, Long J, Hernandez-Nunez L, Emonet T. Adaptability of non-genetic diversity in bacterial chemotaxis. Marder E, editor. eLife. 2014 Oct 3;3:e03526.

12. Schellenberg GD, Furlong CE. Resolution of the multiplicity of the glutamate and aspartate transport systems of Escherichia coli. J Biol Chem. 1977 Dec 25;252(24):9055–64.

13. Novick-Cohen A, Segel LA. A gradually slowing travelling band of chemotactic bacteria. J Math Biology. 1984 Jan 1;19(1):125–32.

14. Narla AV, Cremer J, Hwa T. A Traveling-Wave Solution for Bacterial Chemotaxis with Growth. arXiv:210308100 [q-bio] [Internet]. 2021 Mar 14 [cited 2021 Mar 17]; Available from: http://arxiv.org/abs/2103.08100

15. Kollmann M, Lovdok L, Bartholomé K, Timmer J, Sourjik V. Design principles of a bacterial signalling network. Nature. 2005 Nov;438(7067):504–7.

16. Salman H, Libchaber A. A concentration-dependent switch in the bacterial response to temperature. Nat Cell Biol. 2007 Sep;9(9):1098–100.
17. Kalinin Y, Neumann S, Sourjik V, Wu M. Responses of Escherichia coli Bacteria to Two Opposing Chemoattractant Gradients Depend on the Chemoreceptor Ratio. Journal of Bacteriology. 2010 Apr 1;192(7):1796–800.

18. Yoney A, Salman H. Precision and Variability in Bacterial Temperature Sensing. Biophysical Journal. 2015 May 19;108(10):2427–36.

19. Keegstra JM, Kamino K, Anquez F, Lazova MD, Emonet T, Shimizu TS. Phenotypic diversity and temporal variability in a bacterial signaling network revealed by single-cell FRET. Barkai N, editor. eLife. 2017 Dec 12;6:e27455.

20. Kamino K, Keegstra JM, Long J, Emonet T, Shimizu TS. Adaptive tuning of cell sensory diversity without changes in gene expression. Science Advances. 2020;

21. Mattingly HH, Kamino K, Machta BB, Emonet T. E. coli chemotaxis is information-limited. arXiv:210211732 [physics, q-bio] [Internet]. 2021 Feb 23 [cited 2021 Mar 29]; Available from: http://arxiv.org/abs/2102.11732

22. Park H, Pontius W, Guet CC, Marko JF, Emonet T, Cluzel P. Interdependence of behavioural variability and response to small stimuli in bacteria. Nature. 2010 Dec;468(7325):819–23.

23. Keller EF, Segel LA. Traveling bands of chemotactic bacteria: A theoretical analysis. Journal of Theoretical Biology. 1971 Feb 1;30(2):235–48.

24. Masson J-B, Voisinne G, Wong-Ng J, Celani A, Vergassola M. Noninvasive inference of the molecular chemotactic response using bacterial trajectories. PNAS. 2012 Jan 31;109(5):1802–7.

25. Hazel JR, Sidell BD. A method for the determination of diffusion coefficients for small molecules in aqueous solution. Analytical Biochemistry. 1987 Nov 1;166(2):335–41.

26. Cremer J, Segota I, Yang C, Arnoldini M, Sauls JT, Zhang Z, et al. Effect of flow and peristaltic mixing on bacterial growth in a gut-like channel. PNAS. 2016 Oct 11;113(41):11414–9.

27. Saragosti J, Calvez V, Bournaveas N, Perthame B, Buguin A, Silberzan P. Directional persistence of chemotactic bacteria in a traveling concentration wave. PNAS. 2011 Sep 27;108(39):16235–40.

28. Wong-Ng J, Melbinger A, Celani A, Vergassola M. The Role of Adaptation in Bacterial Speed Races. PLOS Computational Biology. 2016 Jun 3;12(6):e1004974.

29. Cluzel P, Surette M, Leibler S. An Ultrasensitive Bacterial Motor Revealed by Monitoring Signaling Proteins in Single Cells. Science. 2000 Mar 2;287(5458):1652–5.

30. Berg HC, Brown DA. Chemotaxis in Escherichia coli analysed by Three-dimensional Tracking. Nature. 1972 Oct;239(5374):500–4.