Title

Cell-to-cell variation sets a tissue-rheology-dependent bound on collective gradient sensing.

Journal

Proceedings of the National Academy of Sciences, 114(47)

Authors

Camley, Brian

Rappel, Wouter-Jan

Publication Date

2017-11-21

DOI

10.1073/pnas.1712309114

Peer reviewed

Cell-to-cell variation sets a tissue-rheology–dependent bound on collective gradient sensing

Brian A. Camley[a,b,c,1] and Wouter-Jan Rappel[c]

[a]Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218; [b]Department of Biophysics, Johns Hopkins University, Baltimore, MD 21218; and [c]Department of Physics, University of California, San Diego, La Jolla, CA 92093

Edited by Andrea J. Liu, University of Pennsylvania, Philadelphia, PA, and approved October 5, 2017 (received for review July 10, 2017)

When a single cell senses a chemical gradient and chemotaxes, stochastic receptor–ligand binding can be a fundamental limit to the cell's accuracy. For clusters of cells responding to gradients, however, there is a critical difference: Even genetically identical cells have differing responses to chemical signals. With theory and simulation, we show collective chemotaxis is limited by cell-to-cell variation in signaling. We find that when different cells cooperate, the resulting bias can be much larger than the effects of ligand–receptor binding. Specifically, when a strongly responding cell is at one end of a cell cluster, cluster motion is biased toward that cell. These errors are mitigated if clusters average measurements over times long enough for cells to rearrange. In consequence, fluid clusters are better able to sense gradients: We derive a link between cluster accuracy, cell-to-cell variation, and the cluster rheology. Because of this connection, increasing the noisiness of individual cell motion can actually increase the collective accuracy of a cluster by improving fluidity.

fundamental bounds | collective motion | chemotaxis | cell-to-cell variability | rheology

Many cells follow signal gradients to survive or perform their functions, including white blood cells finding a wound, cells crossing a developing embryo, and cancerous cells migrating from tumors. Chemotaxis, sensing and responding to chemical gradients, is crucial in all of these examples (1, 2). Chemotaxis is traditionally studied by exposing single cells to gradients—but cells often travel in groups, not singly (3, 4). Collective cell migration is essential to development and metastasis (5) and can have remarkable effects on chemotaxis. Even when single cells cannot sense a gradient, a cluster of cells may cooperate to sense it. While collective chemotaxis is our primary focus, this "emergent" gradient sensing is found in response to many signals, including soluble chemical gradients (chemotaxis) (6–8), conditioned substrates (haptotaxis) (9), substrate stiffness gradients (durotaxis) (10), and electrical potential (galvanotaxis) (11, 12).

Cells can cooperate to sense gradients—but the physical principles limiting a cluster's sensing accuracy are not settled. For single cells, the fundamental bounds on sensing chemical concentrations and gradients are well studied (13–23), showing unavoidable stochasticity in receptor–ligand binding limits chemotactic accuracy. Is this true for cell clusters? Is a cell cluster simply equivalent to a larger cell? The answer is no! There is an essential difference between many clustered cells and a single large cell: Even clonal populations of cells can have highly variable responses to signals, due to many factors, including intrinsic variations in regulatory protein concentrations (24–26). These cell-to-cell variations (CCVs) can be persistent over timescales much larger than the typical motility timescale of the cell (26). CCV has not been addressed in models of collective chemotaxis and it is not clear whether collective gradient sensing is limited by CCV or by stochastic receptor–ligand binding (7, 8, 27–33).

Using a combination of analytics and simulations, we show that unless CCV is tightly controlled, collective guidance of a cluster of cells is limited by these variations: Gradient sensing is biased toward cells with intrinsically strong responses. This bias swamps the effects of stochastic ligand–receptor binding. Cell clusters may reduce this error by time averaging their gradient measurements only if the cells rearrange their positions, creating an unavoidable link between the mechanics of the cell cluster and its gradient-sensing ability. As a result, surprising new tradeoffs arise: Clusters must balance using motility to follow a biased signal with using motility to reduce error and compromise between reducing noise and increasing cluster fluidity.

Gradient-Sensing Error Is Dominated by Cell-to-Cell Variation, Not Receptor Noise

We study a 2D model of gradient sensing with CCV and ligand–receptor dynamics where cells sense a chemoattractant with concentration gradient $\mathbf{g}$. Each cell at position $\mathbf{r}$ measures local concentration, $c(\mathbf{r}) = c_0(1 + \mathbf{g}\cdot\mathbf{r})$, via ligand–receptor binding, which is stochastic. This noise leads to unavoidable errors in the cluster's estimate of $\mathbf{g}$. In addition, even if concentration is perfectly sensed, each cell responds differently to a fixed c, which models known CCV in signal response (34, 35). As a result, when the cluster combines measurements from its cells, it may develop a drift in the direction of stronger-responding cells (Fig. 1). To combine these effects, we specify the "measured" signal in cell i, M_i, which is what the cluster believes the chemoattractant signal in cell i to be, including ligand–receptor binding and CCV,

$$M^i = \left[c(\mathbf{r}^i) + \delta c^i \eta^i\right]/\bar{c} + \Delta^i, \qquad [1]$$

Significance

Cells cooperate to sense the direction of a chemical gradient by communicating with each other, which may be important when clumps of cancer cells metastasize or embryos develop. However, because each cell is distinct, we find these clumps are biased toward cells that are "loud"—sending inappropriately large signals. Cell clusters can reduce this bias by rearranging themselves so loud cells change their locations. This means the mechanical dynamics of the cluster matter—fluid, squishy clumps of cells are better at sensing than solid ones. If a single cell's motion gets noisier, it will make the cluster more fluid—so adding noise can actually make a cluster of cells a better sensor.

Author contributions: B.A.C. designed research; B.A.C. performed research; B.A.C. and W.-J.R. analyzed data; and B.A.C. and W.-J.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: MATLAB simulation code for the MLE method, self-propelled particle simulations, and subsequent analysis as well as simulation output sufficient to reproduce the relevant figures is available online via the github repository: https://github.com/bcamley/celltocellreproduce.

[1]To whom correspondence should be addressed. Email: bcamley@jhu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1712309114/-/DCSupplemental.

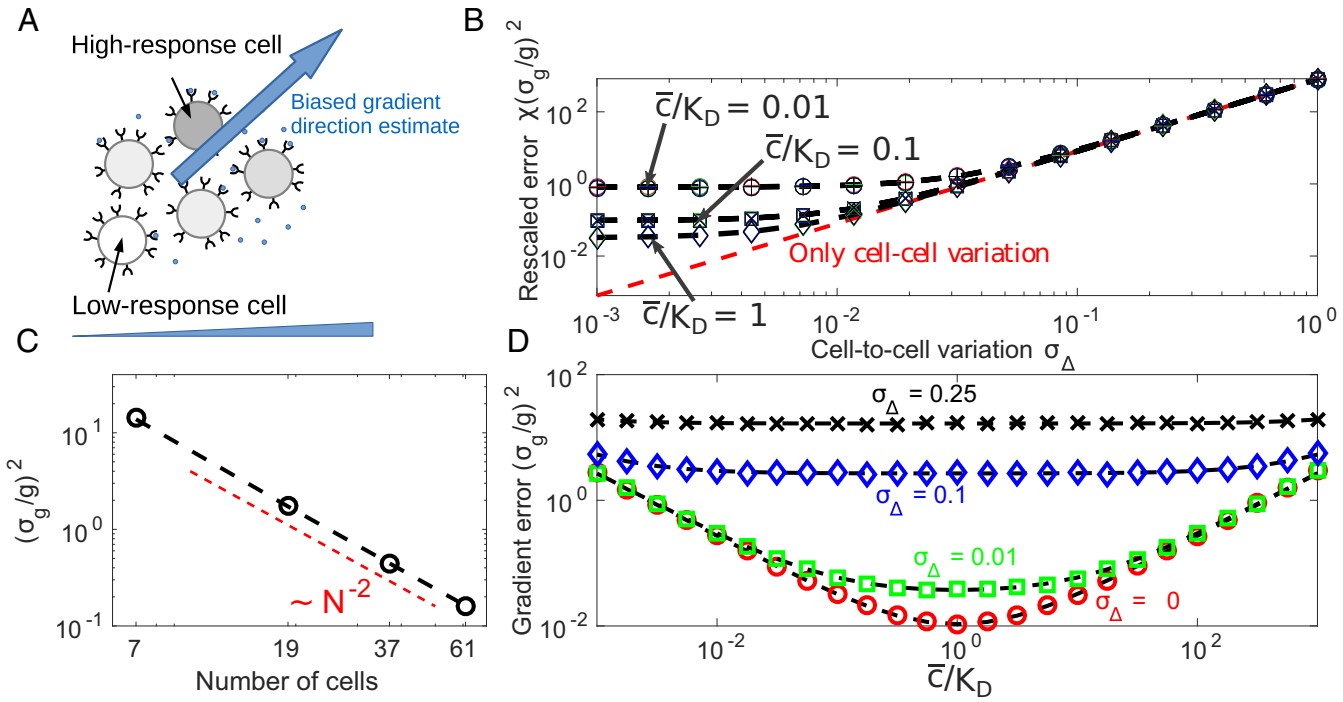

Fig. 1. Cell-to-cell variation creates systematic biases that can be significantly larger than the effects of receptor–ligand binding. (*A*) Schematic of how cell-to-cell variation can create bias in gradient sensing toward high-signaling and away from low-signaling cells. (*B*) Gradient-sensing error $\sigma_{\mathbf{g}}^2 = \left\langle |\hat{\mathbf{g}} - \mathbf{g}|^2 \right\rangle$, derived from numerical maximum likelihood (symbols), is well approximated by Eq. **2** (dashed lines) at low gradient strengths. Symbols are plotted for four cluster sizes: $N = 7$, 19, 37, and 61 cells (hexagonally packed clusters of unit spacing with $Q = 1$, 2, 3, 4 layers, illustrated in *Computation of χ for Cells in Hexagonally Packed Cluster*). In *A*–*D*, we use $n_r = 10^5$ and $g = 0.05$, in units where the cell–cell spacing is 1. (*C*) Gradient-sensing error decreases as cluster size increases as $\sigma_{\mathbf{g}}^2 \sim N^{-2}$. In *C*, $\sigma_\Delta = 0.23$ and $\bar{c} = K_D$. (*D*) Strong CCV can mask concentration dependence of accuracy. In absence of CCV, gradient-sensing accuracy is maximized when $\bar{c} \approx K_D$; this effect is screened when CCV dominates gradient sensing. *D* is shown for $N = 7$ cells.

where $\bar{c}$ is the mean concentration over the cluster, $\bar{c} = N^{-1}\sum_i c(\mathbf{r}^i)$, η^i are uncorrelated Gaussian noises with zero mean and unit variance, and Δ^i are uncorrelated Gaussian noises with zero mean and variance σ_Δ^2. Stochastic fluctuations in ligand–receptor binding are taken into account in the term δc^i, where $(\delta c^i / c(\mathbf{r}^i))^2 = \frac{1}{n_r}\frac{(c^i + K_D)^2}{c^i K_D}$. This is the error in concentration sensing from a single snapshot of n_r receptors with simple ligand–receptor kinetics and dissociation constant K_D (ref. 14 and *Review of Concentration Sensing Accuracy*). Eq. **1** assumes that cell–cell variance additively corrupts the measurement of the concentration $c(\mathbf{r})$ after an adaptation to the overall level of signal across the cluster $\bar{c}$. This is natural if the primary cell-to-cell variation is downstream of adaptation, as found to be a reasonable model in ref. 34. We expect similar results for CCV upstream of adaptation in certain limits (*Cell-to-Cell Variation Upstream of Adaptation*). We have also assumed that the CCV Δ^i is uncorrelated between different cells; this is a useful initial model describing large variations in protein levels that remain localized within each cell. Extensions of the model could potentially address correlations arising from, e.g., extracellular vesicle transport or cell division, where daughter cells' Δ^i may be correlated.

To determine gradient-sensing accuracy, we perform maximum-likelihood estimation (MLE) of $\mathbf{g}$ in Eq. **1**, as in past approaches for single-cell gradient sensing (16). We obtain the MLE $\hat{\mathbf{g}}$ numerically (*Materials and Methods* and *MLEs of Gradient Direction via Collective Guidance in the Presence of Cell–Cell Variation and Ligand–Receptor Noise*) and thus the uncertainty $\sigma_{\mathbf{g}}^2 \equiv \left\langle |\hat{\mathbf{g}} - \mathbf{g}|^2 \right\rangle$ (Fig. 1*B*, symbols), where $\langle \cdots \rangle$ is an average over CCV and ligand–receptor binding. For fixed and roughly circular (isotropic) cluster geometry, if the concentration change across the cluster is small, $gR_{\text{cluster}} \ll 1$, $\sigma_{\mathbf{g}}^2$ can be approximated by assuming δc^i is constant across the cluster, resulting in

$$\left\langle |\hat{\mathbf{g}} - \mathbf{g}|^2 \right\rangle \approx \frac{2}{\chi}\left(\sigma_\Delta^2 + \frac{1}{n_r}\frac{(\bar{c} + K_D)^2}{\bar{c} K_D}\right). \qquad [2]$$

Here, $\chi = \frac{1}{2}\sum_i |\boldsymbol{\delta}\mathbf{r}^i|^2$ is a shape parameter, and $\boldsymbol{\delta}\mathbf{r}^i = \mathbf{r}^i - \mathbf{r}_{\text{cm}}$ is cell position relative to cluster center of mass. Evaluating this expression reveals that it is an excellent approximation to the numerically obtained uncertainty (dashed lines, Fig. 1*B*).

The approximate expression for the uncertainty, Eq. **2**, allows us to quantify the relative contribution of receptor–ligand fluctuations and CCV to the gradient-sensing error. For background concentrations $\bar{c}$ near the receptor–ligand equilibrium constant K_D and for typical receptor numbers in eukaryotic cells [$n_r \sim 10^5$ (36, 37)], $\delta c/\bar{c}$ can be smaller than 0.01. Protein concentrations, on the other hand, often vary between cells to 10–60% of their mean (25)—hence we estimate $\sigma_\Delta \approx 0.1 - 0.6$. Thus, we expect CCV to dominate gradient-sensing error and that the error from concentration sensing and receptor binding can be neglected completely if $\sigma_\Delta > 0.1$ (Fig. 1*B*). Eq. **2** also reveals that CCV masks the impact of changing background concentration. When $\sigma_\Delta = 0$, gradient sensing is limited by ligand–receptor fluctuations and increases as $\bar{c}$ moves away from K_D (Fig. 1*D*)—accuracy decreases if either few receptors are bound or receptors are saturated. As CCV increases, $\sigma_{\mathbf{g}}^2$ no longer depends strongly on $\bar{c}$ (Fig. 1*D*). Finally, Eq. **2** shows that gradient-sensing accuracy depends on the shape parameter χ and, therefore, on cluster size. For hexagonally packed clusters of cells with unit spacing (we measure in units of the cell diameter; *Materials and*

Methods), a cluster with Q layers has $N = 1 + 3Q + 3Q^2$ cells and $\chi(Q) = (5/8)Q^4 + (5/4)Q^3 + (7/8)Q^2 + (1/4)Q$ (*Computation of χ for Cells in Hexagonally Packed Cluster*); i.e., $\chi(Q) \sim Q^4 \sim N^2$. Clusters of increasing size then have an error that decreases as $1/N^2$ (Fig. 1*C*); this scaling is similar to earlier results for single cells (*Computation of χ for Cells in Hexagonally Packed Cluster*).

Reducing Estimation Error by Time Averaging

If a cluster made n independent measurements, it could reduce $\sigma_{\mathbf{g}}^2$ by a factor of n. In single-cell gradient sensing, independent measurements can be made by averaging over time—improving errors by a factor $\sim T/\tau_{\text{corr}}$, where T is the averaging time and τ_{corr} the measurement correlation time. At first glimpse, time averaging seems unlikely to help with CCV, when correlation times for protein levels can be longer than cell division times, reaching 48 h in human cells (26). However, since gradient-sensing bias from CCV depends on the locations of strong- and weak-signaling cells within the cluster, time averaging can be successful if it is over a time long enough for the cluster to rearrange. This is true even if, as we initially assume, CCV biases Δ^i are time independent. We expect gradient sensing error with time averaging, $\sigma^2_{\mathbf{g},T}$, will decrease by a factor of T/τ_r from $\sigma^2_{\mathbf{g},0}$, where τ_r is a correlation time related to cell positions (Fig. 2). Is this true, and how should we define τ_r?

Our earlier results suggest that CCV dominates the gradient-sensing error. Ligand–receptor noise will also be even less relevant in the presence of time averaging, as the receptor relaxation time [seconds to minutes (38)] is much faster than that for cluster rearrangement (tens of minutes or longer). We therefore completely neglect ligand–receptor binding fluctuations, allowing an analytical solution for the MLE $\hat{\mathbf{g}}$ (*Materials and Methods* and *MLEs of Gradient Direction via Collective Guidance in the Presence of Cell–Cell Variation and Ligand–Receptor Noise*).

How much does time averaging reduce error? If we average the MLE $\hat{\mathbf{g}}$ over a time T by applying a kernel $K_T(t)$, i.e., we define $\hat{\mathbf{g}}_T(t) \equiv \int_{-\infty}^{\infty} \hat{\mathbf{g}}(t') K_T(t - t')dt'$ and $\sigma^2_{\mathbf{g},T} \equiv \langle |\hat{\mathbf{g}}_T - \mathbf{g}|^2 \rangle$, we can derive (*Detailed Derivation of Time-Averaged Gradient-Sensing Error*)

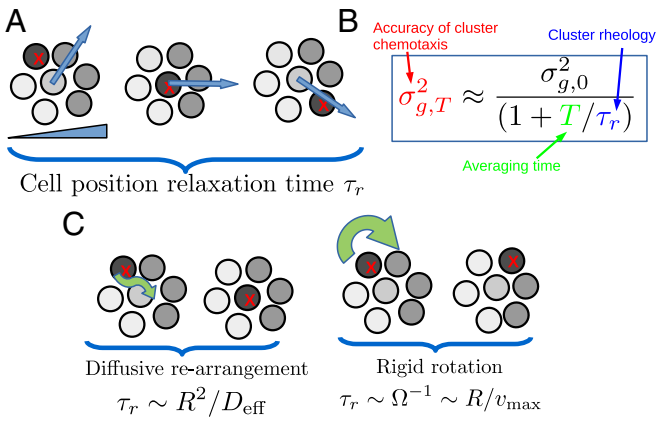

Fig. 2. Time averaging links fluidity and accuracy. (*A*) Schematic drawing of how cell–cell rearrangement can change bias due to CCV. Shades of gray indicate measured signal *M*; a cell with strong response (marked with *X*) moves through the cluster, leading to biases in gradient estimate (blue arrow). The characteristic relaxation time for this bias is τ_r (main text). (*B*) This leads to a link between the timescale τ_r, which is a measure of the cluster's rheology, and chemotactic accuracy $\sigma_{\mathbf{g},T}$ (box). (*C*) Different rearrangement mechanisms will depend on cluster size in different ways (main text and *Characteristic Timescales of Different Rearrangement Mechanisms*).

$$\sigma^2_{\mathbf{g},T} = \sigma^2_{\mathbf{g},0} \times \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} |K_T(\omega)|^2 C_{rr}(\omega), \quad [3]$$

where $C_{rr}(t' - t'') \equiv \langle \boldsymbol{\delta}\mathbf{r}(t') \cdot \boldsymbol{\delta}\mathbf{r}(t'') \rangle / \langle |\boldsymbol{\delta}\mathbf{r}|^2 \rangle$ is the normalized cell position–position correlation function, $C_{rr}(\omega)$ is its Fourier transform, and $\sigma^2_{\mathbf{g},0} = 2\sigma^2_{\Delta}/\chi$ is the error in the absence of time averaging. To derive Eq. **3**, we make two approximations: (*i*) The cluster has a constant and isotropic shape and (*ii*) rearrangement of cell positions relative to the center of mass is independent of the particular values of Δ. The first approximation is not necessary, but is a useful simplification; a generalized result is given in *Detailed Derivation of Time-Averaged Gradient-Sensing Error*. The second approximation assumes that averaging over CCV and averaging over cell positions are independent. This decoupling approximation is necessary to characterize cluster fluidity and mechanics separately from the details of signaling. It excludes, e.g., models where cells with larger-than-average Δ sort out from the cluster. We discuss potential errors due to this approximation later in this paper.

For exponential position–position correlation functions and averaging, $C_{rr}(t) = \exp(-t/\tau_r)$ and $K_T(t) = \theta(t)\frac{1}{T}e^{-t/T}$, where $\theta(t)$ is the Heaviside step function, Eq. **3** is simple:

$$\sigma^2_{\mathbf{g},T} = \frac{\sigma^2_{\mathbf{g},0}}{1 + T/\tau_r}. \quad [4]$$

In other words, gradient-sensing accuracy can be improved by taking T/τ_r independent measurements in a time T. [The assumption of exponential averaging here is not crucial, and many averaging functions can be used in Eq. **3**. We choose this average because it can be computed easily by a combination of accumulation and decay; i.e., $\partial_t \hat{\mathbf{g}}_T = -T^{-1}(\hat{\mathbf{g}}_T - \hat{\mathbf{g}})$.]

Crucial in this reduction is the position–position correlation time τ_r which depends on the cluster rearrangement mechanism. Two natural mechanisms are persistent cluster rotation and neighbor rearrangements within the cluster (Fig. 2*C*). These mechanisms may coexist, as when cells slide past one another during cluster rotation (39). τ_r can depend on cluster size; for diffusive rearrangements, we expect $\tau_r \sim R^2/D_{\text{eff}}$, and for persistently rotating clusters, $\tau_r \sim R/v_{\text{cell}}$ (*Characteristic Timescales of Different Rearrangement Mechanisms*).

We have assumed that the CCV is time independent over our scale of interest—consistent with the long memory found in ref. 26. If Δ changes faster than the cluster rearranges, our results can be straightforwardly modified. Generalizations of Eqs. **3** and **4** to this case are provided in *Detailed Derivation of Time-Averaged Gradient-Sensing Error*.

Tradeoffs in Collective Accuracy and Motility: Cluster Rotation

Our central result (Eq. **3**) shows that clusters can improve their chemotactic accuracy by changing cell positions. The simplest mechanism to do this is cluster rotation, which occurs in border cell clusters (40) and transiently in leukocyte clusters (7). When should a cluster actively rotate to increase its accuracy? Rotation creates an important tradeoff: More work must be put into rotating and therefore less into crawling up the gradient. [We note that this is most relevant if motility is a large portion of the cluster's energy budget, a complex question that may be cell-type dependent (41–43).]

For constant work of motility, the maximum speed of a cluster of radius R that rotates with angular speed Ω is (*Materials and Methods*)

$$v(\Omega) = v_{\max}\sqrt{1 - \frac{1}{2}\Omega^2\tau^2_{\text{rot}}}, \quad [5]$$

where $\tau_{\text{rot}} = R/v_{\max}$ and $v_{\max}$ the maximum cluster speed absent rotation. The cluster cannot rotate faster than $\Omega_{\max} = \sqrt{2}/\tau_{\text{rot}}$.

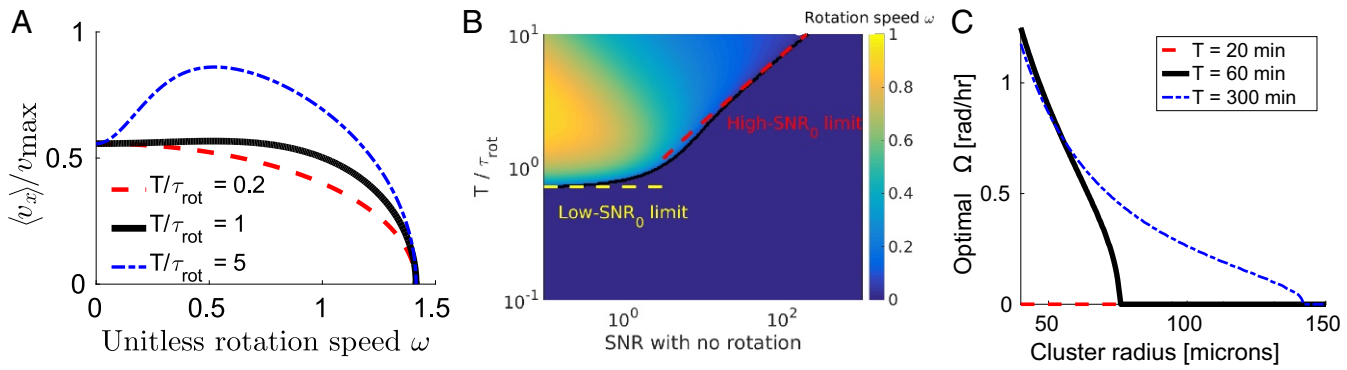

Fig. 3. Cluster rotation can improve directed cluster motility. (*A*) As the averaging time T is increased above the characteristic rotational timescale $\tau_{rot} = R/v_{max}$, the mean cluster velocity in the gradient direction $\langle v_x \rangle$ is maximized for nonzero rotational speed. $SNR_0 = 1$ in *A*. $\omega = \Omega\tau_{rot}$ is the unitless rotational speed. (*B*) Rotation improves chemotaxis at long averaging times T and low SNR_0 (bad gradient sensing in the absence of rotation). Color map shows the value of ω that maximizes $\langle v_x \rangle$, found by numerical evaluation; black line shows the $\omega \approx 0$ contour. (*C*) Cluster rotation is preferred at small cluster radii. In this graph, SNR_0 is estimated by using $\chi \approx \frac{\pi}{4}\rho_c R^4$, where $\rho_c \approx 3.2 \times 10^{-3}$ μm^{-2} is the number of cells per unit area in the cluster (*Materials and Methods*).

If a cluster follows its best estimate $\hat{\mathbf{g}}_T$ with speed $v(\Omega)$ given by Eq. **5**, it can improve its velocity in the gradient direction by rotating when the averaging time T is long compared with τ_{rot} (Fig. 3*A*). (The cluster's directionality is always improved by rotating, so there is no tradeoff unless speed of motion matters.) We find that the optimal rotation speed Ω that maximizes the up-gradient speed depends only on the signal-to-noise ratio (SNR) without rotation, $SNR_0 \equiv \frac{1}{2}g^2/\sigma^2_{\mathbf{g},0}$ and T/τ_{rot} (Fig. 3*B* and *Materials and Methods*). Mammalian cells have speeds in the range of micrometers per minute and radii of tens of micrometers, so to benefit from averaging ($T \gtrsim \tau_{rot}$), T must be longer than tens of minutes. The timescale τ_{rot} and SNR_0 both depend on cluster size—larger clusters with more cells are both better gradient sensors and more difficult to drive to large angular speeds. As a consequence, the optimal Ω is highly cluster-size dependent: As cluster size decreases, there is a continuous transition to nonzero optimal Ω if T is sufficiently long (Fig. 3*C*).

Linking Chemotactic Accuracy and Fluidity

Clusters with more cell rearrangement are more accurate by Eq. **4**. To further quantify the consequences of cell rearrangements, we model a cluster of cells as self-propelled particles that follow the cluster estimate $\hat{\mathbf{g}}_T$ with a noise characterized by angular diffusion D_ψ and with cell–cell connections modeled as springs of strength κ between Delaunay neighbors (*Materials and Methods*). We emphasize that the angular diffusion parameterized by D_ψ is an additional source of noise: As D_ψ increases, cells are less accurate in following the cluster's estimate of the gradient. These two parameters are systematically varied to study the effects of cluster fluidity on chemotactic accuracy.

Cluster Fluidity Improves Cluster Chemotaxis. Within our model, increasing cell–cell adhesion κ makes clusters more ordered, moving between fluid-like and crystalline states (Fig. 4*A*). As a consequence, rearrangement slows significantly (Fig. 4*C*) with $\tau_r \sim \exp(\kappa/2)$ [$C_{rr}(t)$ is single exponential].

Cluster structure and size change when clusters fluidize (Fig. 4*A*), which may in principle affect the shape parameter χ, which also strongly affects the chemotactic accuracy (Eq. **2**). However, in our simulations χ is not significantly dependent on κ, changing by under 10% (Fig. 4*D*). Averaging time T also has only a weak effect on cluster shape and dynamics—changes in τ_r and χ when the averaging time T is increased by orders of magnitude are small (Fig. 4). This is consistent with our assumption decoupling the gradient estimate and cell rearrangements, suggesting clusters should obey the bound [**3**].

We can, using the results in *Reducing Estimation Error by Time Averaging*, predict the cluster chemotactic index, $CI \equiv \langle V_x/|\mathbf{V}| \rangle$, where $\mathbf{V}$ is the cluster velocity. Assuming $\mathbf{V} \sim \hat{\mathbf{g}}_T$, we can compute the CI from $\sigma_{\mathbf{g},T}$ given by Eq. **4** (*Materials and Methods*). This requires parameters τ_r and χ (measured from simulations) and $\mathbf{g}$, T, and σ_Δ (known). We note that our approach, which extracts τ_r and χ from cell trajectories, could also be applied to experimental data; in that case, $\mathbf{g}$ would still be known, but the extent of time averaging (T) and the error due to CCV (σ_Δ) would have to be determined by fitting to the data. This prediction should be an upper bound to the measured CI, because our model includes additional noise beyond the assumptions of Eq. **4**, via D_ψ. As expected, the cluster CI decreases significantly as clusters solidify and the relaxation time τ_r increases. The simulation data qualitatively follow the predicted upper bound (Fig. 4*B*). When the averaging time T is reduced below typical relaxation times, the CI significantly decreases. In addition, for this short time averaging, changing cluster stiffness no longer strongly affects the CI.

Our model describes cell–cell adhesion by controlling the stiffness of springs connecting circular cells. Within this scheme, increasing adhesion reduces cluster fluidity; other models and experiments have suggested adhesion increases jamming and decreases rearrangement (44). However, descriptions of epithelial sheets using models that resolve cell shape, including vertex and cellular Potts models, suggest that increasing adhesion actually promotes a more fluid state (45–48). With this in mind, whether adhesion increases or decreases cluster fluidity may be an open question. We use adhesion in our model as a simple control knob to adjust cluster fluidity; our results in Eq. **4** do not depend on the origin of the tissue's fluidity, only on its characteristic relaxation time τ_r.

Increasing the Stochasticity of Single-Cell Motility Can Increase Cluster Accuracy. Any mechanism that fluidizes the cluster can decrease the correlation time τ_r. Because of this, increasing motility noise can improve cluster chemotactic accuracy (Fig. 5). We increase single-cell angular noise D_ψ and see an initial sharp increase in cluster CI as $D_\psi > 0$ (Fig. 5*B*). At larger values of D_ψ, cluster CI decreases below the bound set by Eq. **4**, as the additional noise added degrades the gradient-following behavior. Without significant time averaging ($T = 0.2$), additional noise primarily impedes chemotactic accuracy.

Why can extra noise D_ψ help sensing? For $D_\psi = 0$, all cells follow the best estimate $\hat{\mathbf{g}}_T$ precisely, leading to an ordered cluster

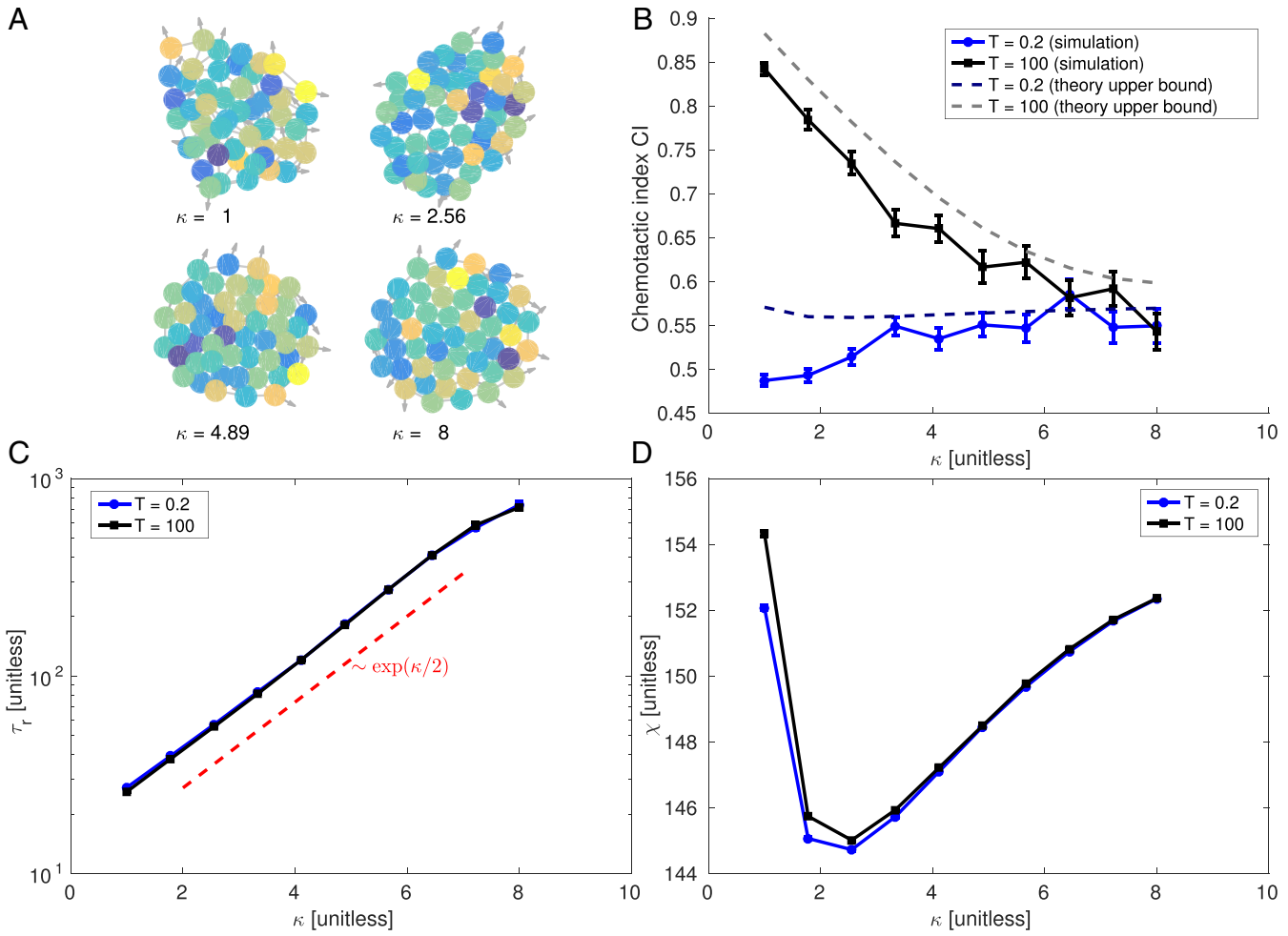

Fig. 4. Cluster fluidity increases cluster accuracy. (*A*) Typical configurations of cell clusters (plotted from simulations with $T = 0.2$). Color indicates measured signal M^i, lines connect neighboring cells, and the arrows indicate the polarity **p**. (*B*) When time averaging is significant, chemotactic index of clusters decreases as cell–cell adhesion stiffness is increased; this behavior is captured by the upper bound computed using Eq. **3**. (*C*) Positional relaxation time τ_r increases with stiffness κ roughly as $\tau_r \sim e^{\kappa/2}$; τ_r does not strongly depend on averaging time T. (*D*) Cluster size parameter χ is not strongly dependent on κ in this range of κ. All plots are computed by averaging over 600 simulations of $N = 50$ cells, each composed of 2×10^4 time steps with $\Delta t = 0.02$. $\sigma_\Delta = 0.3$, $D_\psi = 1$, $\tau = 1$, $\ell = 1$, and $g = 0.025$. The first $2 \times \max(\tau, T)$ time units of the simulation are discarded, to allow the system to reach a steady state. Error bars indicate 68% CIs.

(Fig. 5*A*) with τ_r effectively infinite. As D_ψ is increased, the cluster fluidizes, and the relaxation time decreases strongly (Fig. 5*C*), resulting in more independent measurements. As in Fig. 4, this fluidization is relevant only if the averaging time T exceeds the relaxation time, so when $T = 0.2$, the effect of increasing D_ψ is solely detrimental to chemotaxis.

In deriving our bound, we made two key approximations: cluster isotropy and decoupling. These approximations are exact for the rigid cluster rotation in *Tradeoffs in Collective Accuracy and Motility: Cluster Rotation*, but only approximate for this self-propelled particle model. As a consequence, at small D_ψ, simulated clusters have chemotactic indexes slightly exceeding our predictions (Fig. 5*B*). This error likely arises from emergent couplings between cluster shape and Δ_i—clusters may spread perpendicular to $\hat{\mathbf{g}}_T$, weakening the decoupling approximation (Fig. 5*A*). The approximation of cluster isotropy can be removed (*Detailed Derivation of Time-Averaged Gradient-Sensing Error*) and does not resolve the violation of the bound (Fig. 5*B*). Despite this potential error source, the model captures CI variation over a broad range of parameters (*Bounds Capture Variation of CI over a Large Range of Parameters*).

Discussion

Our study results in several predictions and suggestions for experiments that investigate collective chemotaxis. For example, we predict that, when CCV is large, gradient-sensing error is insensitive to background concentration (Fig. 1*C*). This is consistent with recent measurements on developing organoids that show that the up-gradient bias is not strongly dependent on mean concentration (8), although in contrast with results on lymphocyte clusters (7). Furthermore, if CCV limits collective chemotaxis, clusters that gradient sense in vivo should have tightly regulated expression of proteins relevant to the signal response. Interestingly, measurements of zebrafish posterior lateral line primordium (49) show tightly controlled Sdf1 signaling, measured by Cxcr4b internalization, suggesting that CCV may be small enough to allow for accurate collective gradient sensing.

We also show that there is a direct link between fluidity and chemotaxis as illustrated by Eq. **3**. Verifying this expression in experiments requires simultaneous measurement of several quantities, including cluster size, cluster rearrangement, and signal gradient. Therefore, care has to be taken when modifying experimental conditions as these might change many key

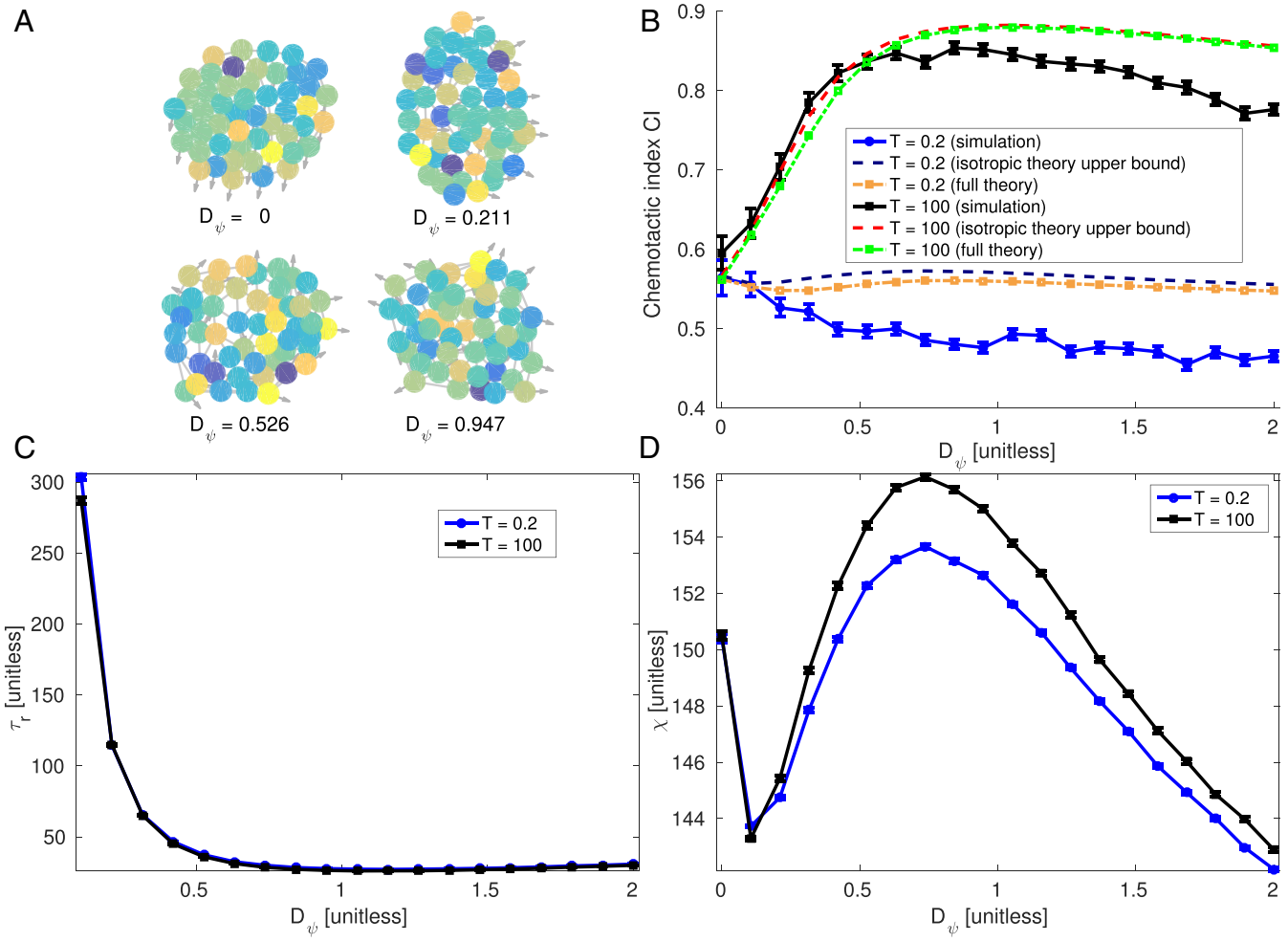

Fig. 5. Finite levels of stochasticity in single-cell motility increase accuracy. (*A*) Typical configurations of cell clusters (plotted from simulations with $T = 100$). Color indicates measured signal M^i, lines connect neighboring cells, and the arrows indicate the polarity $\mathbf{p}$. (*B*) When time averaging is significant, the CI of clusters first increases as increasing single-cell noise D_ψ fluidizes the cluster and then falls below the upper bound computed using Eq. **3**. We also plot an extended theory not assuming cluster isotropy, derived in *Detailed Derivation of Time-Averaged Gradient-Sensing Error*. Simulations for small D_ψ may slightly exceed the upper bound (main text). (*C*) Positional relaxation time τ_r first decreases and then increases as D_ψ is increased; the point $D_\psi = 0$, which has τ_r effectively infinite, is not shown. τ_r does not strongly depend on averaging time T. (*D*) Cluster size parameter χ weakly depends on fluidization by D_ψ. All plots are computed by averaging over 600 simulations of $N = 50$ cells, each composed of 2×10^4 time steps with $\Delta t = 0.02$. $\sigma_\Delta = 0.3$, $\kappa = 1$, $\tau = 1$, $\ell = 1$, and $g = 0.025$. The first $2 \times \max(\tau, T)$ time units of the simulation are discarded, to allow the system to reach a steady state. Error bars indicate 68% CIs.

quantities simultaneously. Altering adhesion, for example, changes both cluster fluidity and spreading as shown in a recent study using neural crest clusters (50), creating a confounding factor. Nevertheless, these types of experiments may be successful in setting bounds on possible time averaging and the link between fluidity and chemotaxis. In principle, given detailed cell position tracks, our analysis could be performed on experimental data. However, there are significant technical details to be worked out. In comparing theory and simulation, we have full tracks of every cell in the cluster over every time point, allowing simple computations; experimental data, however, may have only transient or partial tracks or show clusters that fragment or merge. Critical in this comparison will be assessing potential sources of bias from missing data or including more complex cluster dynamics.

Our results suggest that many recent experiments may need reinterpretation. Measured chemotactic accuracies can depend on cluster size (6, 7, 31); these results have been modeled without time averaging or CCV (7, 27, 28, 31). Our results show that rearrangement times τ_r also influence chemotaxis—and that τ_r depends on cluster size. Cluster relaxation dynamics are therefore an unexplored potential issue for interpreting collective gradient-sensing experiments.

Essential in the reduction of gradient-sensing errors due to CCV is the existence of a biochemical or mechanical memory that can perform a time average over tens of minutes. There are several possibilities. First, memory could be external to the cluster—e.g., stored in extracellular matrix structure or a long-lived trail (51). Second, supracellular structures like actin cables influence cell protrusion and leader cell formation (52), suggesting that collective directional memory could be kept by regulating actin cable formation and maintenance. Third, memory may be kept at the individual cell level by cells attempting to estimate their own bias level Δ^i and compensating for it. This contrasts with our straightforward average of the collective estimate $\hat{\mathbf{g}}$, but could be an important alternative mechanism.

Our model states that the cluster can take its information about signal levels, modulated only by CCV and ligand–receptor noise, and make the optimal computation of the gradient estimator $\hat{\mathbf{g}}$. We do not explicitly identify a biochemical or mechanical mechanism by which the cluster can compute this estimate.

However, at least in the limit where receptor noise is negligible and $\hat{\mathbf{g}} = (\chi^{-1}) \sum_i M^i \boldsymbol{\delta}\mathbf{r}^i$, the estimate depends simply on the cluster geometry. To compute $\hat{\mathbf{g}}$, the cluster would have to be able to estimate only the cell location relative to the cluster center of mass $\boldsymbol{\delta}\mathbf{r}^i$; this could potentially be done by secreting and sensing a signal. Other possibilities could include generalizing the tug-of-war scheme proposed in ref. 27 so that cells farther away from the cluster center exert larger forces. We also note that, even if cell clusters use a different estimator than the optimal MLE $\hat{\mathbf{g}}$ we have described here [e.g., estimators derived from local excitation, global inhibition (LEGI), or tug-of-war models (7, 27, 28, 30)], similar qualitative results will hold: Relaxation of any position-dependent estimator $\hat{\mathbf{g}}$ will be faster if the cluster is more fluid.

We have so far addressed gradient sensing in only two dimensions. In higher dimensions, we expect some aspects of our results to be relatively unchanged; for instance, Eq. **2** will change only by a prefactor, and the results of Eq. **3** straightforwardly generalize to three dimensions. However, the scaling laws that determine how the shape factor χ changes with cluster size and the relaxation times τ_r change with cluster size will naturally differ in three dimensions. For instance, we would expect $\chi \sim \sum_i |\boldsymbol{\delta}\mathbf{r}^i|^2 \sim R^5_{\text{cluster}} \sim N^{5/3}$.

Our results are critical for understanding the ubiquitous phenomenon of collective gradient sensing. The importance of CCV provides a valuable design principle: CCV must either be tightly controlled or be mitigated by time averaging. We also established a surprising link between a central mechanical property of a cluster—its rheology—and its sensing ability. This connects mechanical transitions like unjamming (47) to sensing, opening up more areas of study. In addition, our results show cluster accuracy depends strongly on cluster rearrangement mechanism. Finally, our results show that noise in cell motility can be beneficial for collective sensing.

Materials and Methods

MLE of Gradient Direction in the Presence of Cell–Cell Variation and Ligand–Receptor Noise. We compute the MLE of gradient direction given the measured signal at cell i, M^i, given by Eq. **1**. If the cluster of cells is in a shallow linear gradient, with concentration c_0 at the cluster's center of mass $\mathbf{r}_{\text{cm}} = N^{-1} \sum_i \mathbf{r}^i$, then $c(\mathbf{r}) = c_0\,[1 + \mathbf{g} \cdot (\mathbf{r} - \mathbf{r}_{\text{cm}})]$ and thus $\bar{c} = c_0$. M_i is then $M^i = 1 + \mathbf{g} \cdot \boldsymbol{\delta}\mathbf{r}^i + (\delta c^i / c_0)\eta^i + \Delta^i$ with $\boldsymbol{\delta}\mathbf{r}^i = \mathbf{r} - \mathbf{r}_{\text{cm}}$ and $(\delta c^i / c^i)^2 = \frac{1}{n_r} \frac{(c^i + \kappa_D)^2}{c^i \kappa_D}$; i.e., $(\delta c^i / c_0)^2 = \frac{1}{n_r}(1 + \mathbf{g} \cdot \boldsymbol{\delta}\mathbf{r}^i + K_D/c_0)^2 \frac{1 + \mathbf{g} \cdot \boldsymbol{\delta}\mathbf{r}^i}{K_D/c_0}$.

Δ^i are uncorrelated between cells, with a Gaussian distribution of zero mean and SD σ_Δ; i.e., $\left\langle \Delta^i \Delta^j \right\rangle = \sigma^2_\Delta \delta^{ij}$ with δ^{ij} the Kronecker delta function. As η^i and Δ^i are both Gaussian, the sum of these variables is also Gaussian, and the likelihood of observing a configuration of measured signals $\{M^i\}$ is $\mathcal{L}(\mathbf{g}; \{M^i\}) = P(\{M^i\}|\mathbf{g})$, where $P(\{M^i\}|\mathbf{g})$ is the probability density function of observing the configuration $\{M^i\}$ given parameters $\mathbf{g}$. The likelihood is

$$\mathcal{L}(\mathbf{g}; \{M^i\}) = \prod_i \frac{1}{\sqrt{2\pi h^i}} \exp\left[-\frac{(M^i - \mu^i)^2}{2h^i} \right], \qquad [6]$$

where $\mu^i = 1 + \mathbf{g} \cdot \boldsymbol{\delta}\mathbf{r}^i$ is the mean value of M^i and $h^i = (\delta c^i / c_0)^2 + \sigma^2_\Delta$ is its variance. We want to apply the method of maximum likelihood by finding the gradient parameters $\hat{\mathbf{g}}$ that maximize this likelihood; i.e.,

$$\hat{\mathbf{g}} = \arg \max_{\mathbf{g}} \mathcal{L}(\mathbf{g}; \{M^i\}). \qquad [7]$$

However, because of the complex dependence of h^i on the gradient $\mathbf{g}$, this is not possible analytically. We perform this optimization numerically, using a Nelder–Mead method (Matlab's fminsearch), with an initial guess set by the maximum for $n_r \to \infty$ (i.e., neglecting concentration-sensing noise), which can be found exactly. For numerical convenience, we maximize the log-likelihood $\ln \mathcal{L}(\mathbf{g}; \{M^i\})$, $\ln \mathcal{L}(\mathbf{g}; \{M^i\}) = -\frac{1}{2} \sum_i \ln h^i - \sum_i \frac{(M^i - \mu^i)^2}{2h^i}$ up to an additive constant.

In the limit of $n_r \to \infty$ (neglecting concentration noise), our model becomes a simple linear regression, and the log likelihood can be maximized analytically by finding $\hat{\mathbf{g}}$ such that $\partial_{\mathbf{g}} \ln \mathcal{L}(\mathbf{g}; \{M^i\})|_{\hat{\mathbf{g}}} = 0$ (*MLEs of Gradient Direction via Collective Guidance in the Presence of Cell–Cell Variation and Ligand–Receptor Noise*). The result is $\hat{\mathbf{g}} = \mathcal{A}^{-1} \cdot \sum_i M^i \boldsymbol{\delta}\mathbf{r}^i$, where $\mathcal{A}_{\alpha\beta} \equiv \sum_i \delta r^i_\alpha \delta r^i_\beta$. This estimator is simplest in the limit of roughly circular (isotropic) clusters, where $\sum_i (\delta x^i)^2 \approx \sum_i (\delta y^i)^2 \gg \sum_i \delta x^i \delta y^i$. In this case, $\hat{\mathbf{g}} = (\chi^{-1}) \sum_i M^i \boldsymbol{\delta}\mathbf{r}^i$, where $\chi = \frac{1}{2} \sum_i |\boldsymbol{\delta}\mathbf{r}^i|^2$.

Cluster Rotation Dynamics. How much speed does a cluster lose by rotating? One possibility is to assume the power expended in generating motility is constant. Consider a circular cluster propelling itself over a surface, with the cells having velocity $\mathbf{v}(\mathbf{r})$; we expect that the frictional force per unit area between the cluster and substrate will be $\mathbf{f}_{\text{drag}} = -\xi \mathbf{v}$, where ξ is a friction coefficient with the substrate. If all of the power available for motility is going into driving the cluster over the substrate, then we can write $P = -\int d^2r\, \mathbf{v} \cdot \mathbf{f}_{\text{drag}} = \xi \int d^2r |\mathbf{v}|^2$. If the cluster is traveling as a rigid, circular cluster with its maximum possible velocity, $\mathbf{v} = v_{\max}\hat{\mathbf{x}}$, then $P = \xi \pi R^2 v^2_{\max} \equiv \gamma_t v^2_{\max}$, where $\gamma_t = \xi \pi R^2$ is the translational drag coefficient of the cluster. If, instead, the cluster puts its entire power into rigid-body rotation with $\mathbf{v}(\mathbf{r}) = \Omega_{\max} r(-\sin\theta, \cos\theta)$ (in polar coordinates), then $P = \xi \Omega^2_{\max} \int d^2r r^2 = \xi \frac{\pi}{2} R^4 \Omega^2_{\max} \equiv \gamma_r \Omega^2_{\max}$, where $\gamma_r = \frac{\xi\pi}{2} R^4$ is the rotational drag coefficient of the cluster. In general, the power dissipated if the cluster is moving rigidly with velocity $\mathbf{v}$ and angular speed Ω is $P = \gamma_t v^2 + \gamma_r \Omega^2$ and hence we find that the speed $v(\Omega)$ that a cluster rotating with angular velocity Ω is able to travel to obtain is

$$v(\Omega) = \sqrt{v^2_{\max} - \frac{\gamma_r}{\gamma_t} \Omega^2}. \qquad [8]$$

This quantifies one reasonable tradeoff between speed and angular velocity for a cluster. If the power available for cell motility is a small amount of the cell's energy budget (41–43), other tradeoffs may be more important and additional modeling will be necessary.

We consider a circular cluster traveling toward its best estimate of the gradient with speed $v(\Omega)$ given by Eq. **8** and traveling in the direction of the estimator $\hat{\mathbf{g}}_T$. We can then determine when the cluster maximizes its mean velocity in the direction of the increasing gradient, which we choose to be x, $\langle v_x \rangle$, as a function of Ω. This average is

$$\langle v_x \rangle = \sqrt{v^2_{\max} - \frac{\gamma_r}{\gamma_t} \Omega^2} \times \left\langle \frac{\hat{\mathbf{g}}_{T,x}}{|\hat{\mathbf{g}}_T|} \right\rangle. \qquad [9]$$

We know from our results above and in *Detailed Derivation of Time-Averaged Gradient-Sensing Error* that, for a fixed configuration, $\hat{\mathbf{g}}_T$ has a Gaussian distribution with mean $\mathbf{g} = g\hat{\mathbf{x}}$ and variance given by Eq. **3**. The average of $\frac{\hat{\mathbf{g}}_{T,x}}{|\hat{\mathbf{g}}_T|}$ depends only on $\text{SNR}_T \equiv \frac{1}{2}(g^2/\sigma^2_{\mathbf{g},T})$, with $\left\langle \frac{\hat{\mathbf{g}}_{T,x}}{|\hat{\mathbf{g}}_T|} \right\rangle = C(\text{SNR}_T^{-1/2})$ (*Computing Chemotactic Indexes*, below).

Given the angular velocity Ω, we can work out the distribution of $\hat{\mathbf{g}}_T$ by Eq. **3**. We know $\mathbf{r}(t) \cdot \mathbf{r}(0) = |\mathbf{r}(0)|^2 \cos \Omega t$, and hence $C_{rr}(t) = \cos \Omega t$ and its Fourier transform $C_{rr}(\omega) = \pi\,[\delta(\omega - \Omega) + \delta(\omega + \Omega)]$, and thus $\sigma^2_{\mathbf{g},T} = \sigma^2_{\mathbf{g},0} \times |K_T(\Omega)|^2 = \sigma^2_{\mathbf{g},0}/(1 + \Omega^2 T^2)$.

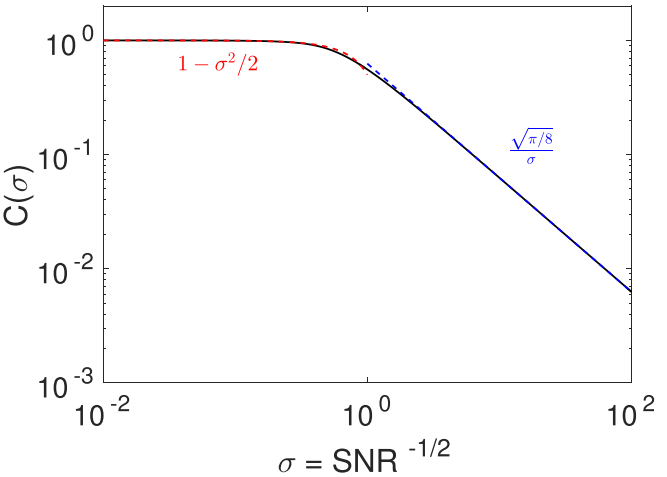

Fig. 6. Relationship between instantaneous CI and SNR. $C(\sigma)$ is plotted numerically from definition in Eq. **18**.

By rescaling to unitless parameters, we then find that

$$\langle v_x\rangle/v_{\max} = \sqrt{1-\frac{1}{2}\omega^2}\times C\left(\left[\mathrm{SNR}_0(1+\omega^2\widetilde{T}^2)\right]^{-1/2}\right), \quad [10]$$

where $\omega = \Omega R/v_{\max}$ is the unitless angular velocity, $\mathrm{SNR}_0 = \sigma_\Delta^{-2} g^2 \chi$ is the usual SNR with no averaging, $\widetilde{T} = Tv_{\max}/R$ is the ratio of the averaging time to the characteristic rotational time $R/v_{\max}$, and $C(\sigma)$ is the function given by Eq. **18**. When SNR_0 is sufficiently small and $\widetilde{T}$ sufficiently large, $\langle v_x\rangle/v_{\max}$ has a maximum at finite ω (Fig. 3).

In the limit of low SNR, $C(\sigma)\approx\sqrt{\pi/8}\sigma^{-1}$, and we find $\langle v_x\rangle/v_{\max}$ is maximized by $\omega = \pm\sqrt{1-\frac{1}{2}\widetilde{T}^{-2}}$ when $\widetilde{T} > 1/\sqrt{2}$ and $\omega = 0$ otherwise. For the large SNR limit, $C(\sigma)\approx 1-\sigma^2/2$ and rotation will increase the mean velocity in the direction of the gradient when $\widetilde{T}^2 > \mathrm{SNR}_0/2 - 1/4$. More generally, it is possible to find the value of ω that maximizes $\langle v_x\rangle$ numerically. We show the complete phase diagram in Fig. 3*B*.

Particle-Based Model of Collective Cell Migration. We use a minimal model of collective cell migration, describing cells as self-propelled particles connected by springs,

$$\frac{d}{dt}\mathbf{r}^i = \mathbf{p}^i + \sum_{j\sim i}\mathbf{F}^{ij} \quad [11]$$

$$\mathbf{p}^i = (\cos\theta^i,\ \sin\theta^i) \quad [12]$$

$$\theta^i = \arctan(\hat{g}_{T,y}/\hat{g}_{T,x}) + \psi^i \quad [13]$$

$$\frac{d}{dt}\psi^i = -\tau^{-1}\sin\psi^i + \sqrt{2D_\psi}\xi^i(t), \quad [14]$$

where $\xi^i(t)$ is a Gaussian Langevin noise with zero mean and $\langle\xi^i(t)\xi^j(t')\rangle = \delta^{ij}\delta(t-t')$, with δ^{ij} the Kronecker delta. In this model, the orientation of an individual cell θ^i is the cluster's best estimate of the gradient direction, $\arctan(\hat{g}_{T,y}/\hat{g}_{T,x})$, plus a noise ψ^i which varies from cell to cell. τ here controls the persistence of this noise and D_ψ its amplitude; when D_ψ is increased, each individual cell is worse at following the estimate $\hat{\mathbf{g}}_T$. The cell–cell forces are

$$\mathbf{F}^{ij} = -\kappa(|d^{ij}-\ell|)\hat{\mathbf{r}}^{ij}, \quad [15]$$

where $d^{ij} = |\mathbf{r}^i - \mathbf{r}^j|$ and $\hat{\mathbf{r}}^{ij} = (\mathbf{r}^j - \mathbf{r}^i)/d^{ij}$. The forces are only between neighboring cells $j\sim i$, where we define neighboring cells as any cells connected by the Delaunay triangulation of the cell centers (Fig. 4*A*); this approach resembles that of ref. 53. We use the Euler–Maruyama method to integrate Eqs. **11**–**14**.

Simulation Units. We have chosen our parameters in the simulation and throughout the paper to be measured in units where the equilibrium cell–cell separation $\ell = 1$ (i.e., the cell diameter is unity), and the velocity of a single cell in the absence of cell–cell forces $\mathbf{v} = \mathbf{p} = (\cos\theta,\ \sin\theta)$ has unit magnitude. For, e.g., neural crest cells, the cell diameters are of order 20 μm and the cell speeds on the order of micrometers per minute—so a unitless time of T corresponds to roughly 20 min $\times$ T in real time. However, cell size and speed vary strongly from cell type to cell type, so we prefer to present these results in their unitless form so that they can be more easily converted.

Computing Chemotactic Indexes. If we use the maximum-likelihood method to make an estimate for the direction in which the cell moves, how do we translate between the uncertainty $\sigma_{\mathbf{g}}^2$ and the distribution of velocities? We found that the MLE for the gradient is $\hat{\mathbf{g}} = \mathbf{g} + \Lambda$, with Λ a Gaussian random variable with zero mean and variance $\langle\Lambda_x^2\rangle = \langle\Lambda_y^2\rangle = \sigma_{\mathbf{g}}^2/2$—and similar results for the time-average $\hat{\mathbf{g}}_T$. One measure of this estimate's accuracy is the instantaneous chemotactic index—or the cosine of the angle between the estimate and the gradient direction. To compute this, if $\mathbf{g} = g\hat{\mathbf{x}}$ without loss of generality, we find

$$\left\langle\frac{\hat{\mathbf{g}}_x}{|\hat{\mathbf{g}}|}\right\rangle = \left\langle\frac{g+\Lambda_x}{(g+\Lambda_x)^2+\Lambda_y^2}\right\rangle \quad [16]$$

$$= \int\frac{dxdy}{2\pi\sigma}\frac{1+x}{\left[(1+x)^2+y^2\right]^{1/2}}e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad [17]$$

$$\equiv C(\sigma), \quad [18]$$

where $\sigma = \mathrm{SNR}^{-1/2}$, with $\mathrm{SNR} = \frac{1}{2}(g^2/\sigma_{\mathbf{g}}^2)$ (Fig. 6). These results carry over naturally to the time-averaged case if $\hat{\mathbf{g}}_T$ remains Gaussian—we find $\left\langle\frac{\hat{\mathbf{g}}_{T,x}}{|\hat{\mathbf{g}}_T|}\right\rangle = C(\mathrm{SNR}_T^{-1/2})$.

The integral for $C(\sigma)$ cannot be solved analytically, but we can find asymptotic forms for $C(\sigma)$ or evaluate it numerically. For $\sigma\gg 1$, we find $C(\sigma)\approx\sqrt{\pi/8}\sigma^{-1}$, and $C(\sigma)\approx 1-\frac{1}{2}\sigma^2$ for $\sigma\ll 1$.

ACKNOWLEDGMENTS. We thank Albert Bae and Monica Skoge for useful discussions and many scientists from the Gordon Research Conference on Directed Cell Motility for interesting questions and reference suggestions. B.A.C. also thanks Kristen Flowers for several useful suggestions. This work was supported by NIH Grant P01 GM078586.

1. Swaney KF, Huang CH, Devreotes PN (2010) Eukaryotic chemotaxis: A network of signaling pathways controls motility, directional sensing, and polarity. *Annu Rev Biophys* 39:265–289.
2. Levine H, Rappel WJ (2013) The physics of eukaryotic chemotaxis. *Phys Today* 66: 24–30.
3. Hakim V, Silberzan P (2017) Collective cell migration: A physics perspective. *Rep Prog Phys* 80:076601.
4. Camley BA, Rappel WJ (2017) Physical models of collective cell motility: From cell to tissue. *J Phys D Appl Phys* 50:113002.
5. Friedl P, Gilmour D (2009) Collective cell migration in morphogenesis, regeneration and cancer. *Nat Rev Mol Cell Biol* 10:445–457.
6. Theveneau E, et al. (2010) Collective chemotaxis requires contact-dependent cell polarity. *Dev Cell* 19:39–53.
7. Malet-Engra G, et al. (2015) Collective cell motility promotes chemotactic prowess and resistance to chemorepulsion. *Curr Biol* 25:242–250.
8. Ellison D, et al. (2016) Cell–cell communication enhances the capacity of cell ensembles to sense shallow gradients during morphogenesis. *Proc Natl Acad Sci USA* 113:E679–E688.
9. Winklbauer R, Selchow A, Nagel M, Angres B (1992) Cell interaction and its role in mesoderm cell migration during Xenopus gastrulation. *Dev Dyn* 195:290–302.
10. Sunyer R, et al. (2016) Collective cell durotaxis emerges from long-range intercellular force transmission. *Science* 353:1157–1161.
11. Li L, et al. (2012) E-cadherin plays an essential role in collective directional migration of large epithelial sheets. *Cell Mol Life Sci* 69:2779–2789.
12. Lalli ML, Asthagiri AR (2015) Collective migration exhibits greater sensitivity but slower dynamics of alignment to applied electric fields. *Cell Mol Bioeng* 8: 247–257.
13. Berg HC, Purcell EM (1977) Physics of chemoreception. *Biophys J* 20:193–219.
14. Kaizu K, et al. (2014) The Berg-Purcell limit revisited. *Biophys J* 106:976–985.
15. Hu B, Chen W, Rappel WJ, Levine H (2010) Physical limits on cellular sensing of spatial gradients. *Phys Rev Lett* 105:048104.
16. Hu B, Chen W, Rappel W-J, Levine H (2011) How geometry and internal bias affect the accuracy of eukaryotic gradient sensing. *Phys Rev E* 83:021917.
17. Endres RG, Wingreen NS (2009) Maximum likelihood and the single receptor. *Phys Rev Lett* 103:158101.
18. Endres RG, Wingreen NS (2008) Accuracy of direct gradient sensing by single cells. *Proc Natl Acad Sci USA* 105:15749–15754.
19. Fuller D, et al. (2010) External and internal constraints on eukaryotic chemotaxis. *Proc Natl Acad Sci USA* 107:9656–9659.
20. Ueda M, Shibata T (2007) Stochastic signal processing and transduction in chemotactic response of eukaryotic cells. *Biophys J* 93:11–20.
21. Andrews BW, Iglesias PA (2007) An information-theoretic characterization of the optimal gradient sensing response of cells. *Plos Comput Biol* 3:e153.
22. Bialek W, Setayeshgar S (2005) Physical limits to biochemical signaling. *Proc Natl Acad Sci USA* 102:10040–10045.
23. Segota I, et al. (2013) High fidelity information processing in folic acid chemotaxis of dictyostelium amoebae. *J R Soc Interface* 10:20130606.
24. Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci USA* 99:12795–12800.
25. Niepel M, Spencer SL, Sorger PK (2009) Non-genetic cell-to-cell variability and the consequences for pharmacology. *Curr Opin Chem Biol* 13:556–561.
26. Sigal A, et al. (2006) Variability and memory of protein levels in human cells. *Nature* 444:643–646.
27. Camley BA, Zimmermann J, Levine H, Rappel W-J (2016) Emergent collective chemotaxis without single-cell gradient sensing. *Phys Rev Lett* 116:098101.
28. Camley BA, Zimmermann J, Levine H, Rappel WJ (2016) Collective signal processing in cluster chemotaxis: Roles of adaptation, amplification, and co-attraction in collective guidance. *PLoS Comput Biol* 12:e1005008.
29. Varennes J, Han B, Mugler A (2016) Collective chemotaxis through noisy multicellular gradient sensing. *Biophys J* 111:640–649.
30. Mugler A, Levchenko A, Nemenman I (2016) Limits to the precision of gradient sensing with spatial communication and temporal integration. *Proc Natl Acad Sci USA* 113:E689–E695.

31. Cai D, et al. (2016) Modeling and analysis of collective cell migration in an in vivo three-dimensional environment. *Proc Natl Acad Sci USA* 113:E2134–E2141.
32. Copenhagen K, et al. (2017) Frustration induced phases in migrating cell clusters. arXiv:1705.00025.
33. Varennes J, Fancher S, Han B, Mugler A (2017) Emergent versus individual-based multicellular chemotaxis. *Phys Rev Lett* 119:188101.
34. Wang CJ, Bergmann A, Lin B, Kim K, Levchenko A (2012) Diverse sensitivity thresholds in dynamic signaling responses by social amoebae. *Sci Signal* 5:ra17.
35. Samadani A, Mettetal J, van Oudenaarden A (2006) Cellular asymmetry and individuality in directional sensing. *Proc Natl Acad Sci USA* 103:11549–11554.
36. Hesselgesser J, et al. (1998) Identification and characterization of the CXCR4 chemokine receptor in human T cell lines: Ligand binding, biological activity, and HIV-1 infectivity. *J Immunol* 160:877–883.
37. Macdonald JL, Pike LJ (2008) Heterogeneity in EGF-binding affinities arises from negative cooperativity in an aggregating system. *Proc Natl Acad Sci USA* 105:112–117.
38. Wang K, Rappel WJ, Kerr R, Levine H (2007) Quantifying noise levels of intercellular signals. *Phys Rev E Stat Nonlin Soft Matter Phys* 75:061905.
39. Rappel WJ, Nicol A, Sarkissian A, Levine H, Loomis WF (1999) Self-organized vortex state in two-dimensional Dictyostelium dynamics. *Phys Rev Lett* 83:1247.
40. Combedazou A, et al. (2016) Myosin II governs collective cell migration behaviour downstream of guidance receptor signalling. *J Cell Sci* 130:97–103.
41. Purcell EM (1977) Life at low Reynolds number. *Am J Phys* 45:3–11.
42. Flamholz A, Phillips R, Milo R (2014) The quantified cell. *Mol Biol Cell* 25:3497–3500.
43. Katsu-Kimura Y, Nakaya F, Baba SA, Mogami Y (2009) Substantial energy expenditure for locomotion in ciliates verified by means of simultaneous measurement of oxygen consumption rate and swimming speed. *J Exp Biol* 212:1819–1824.
44. Garcia S, et al. (2015) Physics of active jamming during collective cellular motion in a monolayer. *Proc Natl Acad Sci USA* 112:15314–15319.
45. Chiang M, Marenduzzo D (2016) Glass transitions in the cellular Potts model. *Europhys Lett* 116:28009.
46. Bi D, Lopez J, Schwarz J, Manning ML (2015) A density-independent rigidity transition in biological tissues. *Nat Phys* 11:1074–1079.
47. Bi D, Yang X, Marchetti MC, Manning ML (2016) Motility-driven glass and jamming transitions in biological tissues. *Phys Rev X* 6:021011.
48. Park JA, et al. (2015) Unjamming and cell shape in the asthmatic airway epithelium. *Nat Mater* 14:1040–1048.
49. Venkiteswaran G, et al. (2013) Generation and dynamics of an endogenous, self-generated signaling gradient across a migrating tissue. *Cell* 155:674–687.
50. Kuriyama S, et al. (2014) In vivo collective cell migration requires an lpar2-dependent increase in tissue fluidity. *J Cell Biol* 206:113–127.
51. Lim K, et al. (2015) Neutrophil trails guide influenza-specific CD8+ T cells in the airways. *Science* 349:aaa4352.
52. Reffay M, et al. (2014) Interplay of RhoA and mechanical forces in collective cell migration driven by leader cells. *Nat Cell Biol* 16:217–223.
53. Meineke FA, Potten CS, Loeffler M (2001) Cell migration and organization in the intestinal crypt using a lattice-free model. *Cell Prolif* 34:253–266.