Title: Capillary interactions drive the self-organization of bacterial colonies

Authors: Matthew E. Black[a,1], Chenyi Fei[a,1], Ricard Alert[b,c,d], Ned S. Wingreen[a,e,2], Joshua W. Shaevitz[a,f,2]

Affiliations:

[a]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA

[b]Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany

[c]Center for Systems Biology Dresden, 01307 Dresden, Germany

[d]Cluster of Excellence Physics of Life, 01062 Dresden, Germany

[e]Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA

[f]Joseph Henry Laboratories of Physics, Princeton University, Princeton, NJ 08544, USA

[1]These authors contributed equally.

[2]Correspondence to: wingreen@princeton.edu or shaevitz@princeton.edu

1

Abstract

Many bacteria inhabit thin layers of water on solid surfaces both naturally in soils or on hosts or textiles and in the lab on agar hydrogels. In these environments, cells experience capillary forces, yet an understanding of how these forces shape bacterial collective behaviors remains elusive. Here, we show that the water menisci formed around bacteria lead to capillary attraction between cells while still allowing them to slide past one another. We develop an experimental apparatus that allows us to control bacterial collective behaviors by varying the strength and range of capillary forces. Combining 3D imaging and cell tracking with agent-based modeling, we demonstrate that capillary attraction organizes rod-shaped bacteria into densely packed, nematic groups, and profoundly influences their collective dynamics and morphologies. Our results suggest that capillary forces may be a ubiquitous physical ingredient in shaping microbial communities in partially hydrated environments.

Introduction

Many bacteria live in thin layers of water wetted to solid surfaces. In partially saturated soils, cells inhabit the liquid bridges that bind soil particles together [1, 2, 3, 4]. In plants, bacteria colonize the surfaces of leaves and flowers, where thin water films and threads can form [5]. In textiles, cells adhere to wetted fibers, with the thickness of the surrounding liquid film changing as the material hydrates or dries [6]. Laboratory experiments typically use hydrogels containing liquid media to study growth and motility of microorganisms.

For nearly a century, substrate hydration has been recognized as a necessary and important factor in structuring the swarming behavior of flagellated bacteria [7, 8, 9, 10, 11, 12, 13]. For such cells, local water availability sets an allowable range within which the flagella can operate and cells can thus move. Non-flagellated, surface-adherent bacteria have demonstrated a variety of other motility mechanisms, *e.g.* through pilus-mediated twitching or focal-adhesion mediated gliding [14, 15, 16, 17]. Motility through these mechanisms also depends on water availability. For example, bulk twitching and gliding assays often need to be conducted in humid environments [15, 19, 21]. Recently, this link was more firmly established in *Myxococcus xanthus* where it was demonstrated that the wetting of surface heterogeneities affects cellular motility and aggregation [22]. Despite long-standing, albeit often tacit, knowledge that hydration affects the motility of surface-adherent bacteria, the mechanisms by which forces from the liquid affect these populations remain unknown.

Previous work on colloidal particles has shown that interfacial forces from water can have substantial effects on the micron scales typical of bacteria. Particles partially immersed in a thin layer of liquid experience capillary forces due to the deformation of the liquid-air interface [23, 24]. For a single hydrophilic particle, capillary forces push it into the hydrated substrate and may increase adhesion to or friction with the surface [25, 26]. When the menisci of two particles coalesce, capillary forces typically produce an attractive interaction between the particles [23]. Here, we show that this liquid-mediated interaction strongly affects the motility of gliding bacteria and governs their collective behaviors.

3

Results

Water menisci form around bacterial cells on hydrated substrates.

We first sought to measure the wetting meniscus around a single cell. We spotted bacteria onto 1.5% v/v agarose hydrogels and imaged the resulting liquid-air interface around single cells with a laser-scanning confocal microscope (Keyence VK-X1000). This instrument provides both an image of the reflectance of the sample, from which cells can be identified, and the corresponding height field of the sample's surface. We imaged the gliding bacteria *M. xanthus* and *Flavobacterium johnsoniae*, non-gliding *Escherichia coli*, and 1-μm-diameter polystyrene colloidal beads (Supplementary Table 1). In all cases, we observed a liquid meniscus surrounding the objects (Fig. 1a).

For a hydrated substrate such as agarose, we modeled the meniscus profile $h(\mathbf{r}) = h(x, y)$ by minimizing the energy of the wetting liquid with surface tension γ and an energy penalty for extracting water from the substrate due to the osmotic pressure difference P across the substrate surface (Fig. 1b). We computed the equilibrium height profile of the modeled water meniscus via numerical minimization of a total free energy $F[h(x, y)] = \int \left(\gamma \sqrt{1 + (\partial_x h)^2 + (\partial_y h)^2} + Ph \right) dx \, dy$, assuming that the water meniscus completely wets the cell surface (Supplementary Sec. I.A). These results are in agreement with the measured height profiles around all samples tested (Fig. 1c).

The shape of a water meniscus varies with osmotic pressure.

Our model predicts that the width of the water meniscus around the cell is controlled by the capillary length $l_\gamma = \frac{\gamma}{P}$. Specifically, for an infinitely long cylindrical cell of radius R_c, the non-zero principal curvature of the surrounding water meniscus is $l_\gamma^{-1} = \frac{P}{\gamma}$, and thus its width scales with $\sqrt{R_c l_\gamma}$, which arises from the geometry of two tangent circles of radii R_c and l_γ (Supplementary Figs. 1 and 2, and Sec. I.B).

To test this prediction for the meniscus width, we developed an apparatus to control the osmotic pressure difference P. Similar to devices for controlling water availability in model soils [27, 28], the apparatus works by connecting the agarose gel directly to a liquid

4

reservoir that, like the gel surface, is open to the atmosphere (Supplementary Fig. 3 and Sec. VI.B). The height of the gel surface (H_{gel}) relative to the height of the reservoir (H_{res}), $\Delta H = H_{\text{gel}} - H_{\text{res}}$, can then be used to apply a hydrostatic pressure to the gel $\Delta P = \rho_{\text{w}} g \Delta H$, where $\rho_{\text{w}} g = 9.8 \text{ kPa/m}$ is the specific weight of water (Fig. 1d, and Supplementary Fig. 4). We set $\Delta H = 0$ as the reference state for each condition and obtained the osmotic pressure, $P_0 \approx 2 \text{ kPa}$, of this state by fitting the modeled water height profiles to those measured in experiment (Supplementary Fig. 5, Sec. I.C,D, and Table 2). The osmotic pressure difference P across the substrate surface is given by $P = P_0 + \rho_{\text{w}} g \Delta H$, (Supplementary Fig. 6 and Sec. I.E). We varied ΔH and measured the meniscus shape around cells of variable size, shape, and aspect ratio. The surface tension of all media used in the experiments was measured by the pendant drop method (Supplementary Sec. VI.C and Table 3). We set $\gamma = 66 \text{ mN/m}$ when fitting the experimental profiles in simulation. As our model predicts, the width of the wetting meniscus, which correlates with the capillary length $l_\gamma = \frac{\gamma}{P}$, becomes smaller when P is increased due to the larger free-energy cost of extracting water from the substrate surface (Fig. 1e,f).

Capillary forces facilitate the formation of cell groups.

For an isolated cell, the lateral capillary forces around the cell balance one another. However, when the menisci of adjacent cells overlap with each other, the water profile around each cell becomes asymmetric. This gives rise to a lateral attractive capillary force between the cells (Fig. 2a). Higher osmotic pressures, P, corresponding to lower water availability, should lead to higher lateral forces but also smaller menisci and thus a shorter interaction range. We confirmed this intuition by coupling the free-energy model of water to an agent-based model of bacterial cells. We model the cells as horizontally oriented spherocylinders with radius R_{c} and cylindrical length l_{c} and compute the forces on the cells exerted by the simulated water menisci (Supplementary Figs. 7 and 8, and Sec. I.F,G). Using this model, we simulated the merging of two adjacent cells being pulled together by capillary forces for different values of the osmotic pressure (Fig. 2b,c).

We next asked whether our model was sufficient to capture the merging dynamics of non-motile *M. xanthus* cells. We consider the balance of forces and torques exerted on each

5

cell, including those from cell-water interaction, cell-cell steric repulsion, and cell-substrate friction (Supplementary Sec. II). Following Copenhagen et al. [29], we model the friction between *M. xanthus* cells and the underlying substrate as anisotropic with smaller friction coefficient along the cell-body axis ($\xi_\parallel$) than perpendicular to it ($\xi_\perp$). We determined the values of these coefficients (Supplementary Table 2) by fitting the agent-based simulations to the dynamics of an experiment in which non-motile *M. xanthus* cells deposited onto an agarose gel were pulled together, into a side-by-side configuration, by their overlapping menisci (Fig. 2d, and Supplementary Video 1). The fitted model recapitulates both the translational (Fig. 2e) and rotational (Fig. 2f) dynamics of these merger events, and supports the anisotropic friction proposed in ref. [29] (Supplementary Fig. 9).

Similar rearrangements also occur for motile bacteria, giving rise to formation of pairs and larger groups when adjacent cells move near one another (Supplementary Video 2). Given the attractive capillary forces between adjacent cells (Supplementary Fig. 10), we reasoned that the strength of such forces would anti-correlate with the rate at which two motile cells would split apart from one another. Measurements of the splitting rate (number of events/total time tracked) of wild-type *M. xanthus* cell pairs across a range of capillary lengths confirmed this hypothesis (Fig. 2g,h). Similarly, for large groups, the magnitude of group shape fluctuations is inversely related to the strength of capillary forces (Supplementary Figs. 11 and 12).

For non-motile but growing bacteria, we hypothesized that similar dynamics driven by capillary forces could account for the side-by-side alignment of daughter cells after division. To test this, we deposited exponentially growing but non-motile *E. coli* and *F. johnsoniae* on agarose pads and collected timelapse micrographs of pairs of daughter cells formed by the division of an isolated, single cell (Fig. 2i, left). Subsequently, we segmented images using a convolutional neural network and tracked individual daughter cells over time. As hypothesized, pairs of cells of both species exhibited similar dynamics: cells that were initially arranged end-to-end first slid sideways and then longitudinally to align in a side-to-side configuration. To quantify this process, we tracked the centroid-to-centroid distance between pairs of cells, which first increased due to growth as the cells

slid sideways, and then decreased as the cells were pulled toward each other by the capillary force (Fig. 2i, right, and Supplementary Figs. 13 and 14).

Capillary forces and cell motility lead to different phases of collective motion.

We used the above data to calibrate an extended version of our agent-based model where cells self-propel by generating in-plane forces, $\boldsymbol{F}_{\mathrm{prop}}$, along their long axis, and reverse directions stochastically with mean reversal time τ_{r}. Since the capillary force acting laterally on a cell can be on the order of 10 nN (Fig. 2b), much stronger than the forces generated by bacterial motility motors (10~100 pN) [30, 31, 32], we explored how capillary forces might affect the collective dynamics of motile cells. We performed a series of simulations for motile but non-growing cells, fixing their density and aspect ratio while varying three dimensionless variables: a normalized reversal period $\tilde{\tau} = \frac{\tau_{\mathrm{r}}}{\tau_{\mathrm{c}}}$ where $\tau_{\mathrm{c}} = \frac{\xi_{\parallel} l_{\mathrm{c}}^2}{|\boldsymbol{F}_{\mathrm{prop}}|}$ is the time it takes for a self-propelled cell to travel its body length, a normalized self-propulsion strength $\tilde{F} = \frac{|\boldsymbol{F}_{\mathrm{prop}}|}{PR_{\mathrm{c}}l_{\mathrm{c}}}$ where $PR_{\mathrm{c}}l_{\mathrm{c}}$ is the scale of the horizontal capillary force acting on a cell, and a normalized capillary length $\tilde{\ell} = \frac{l_{\gamma}}{R_{\mathrm{c}}} = \frac{\gamma}{PR_{\mathrm{c}}}$.

For the simulation parameters we investigated, the collective dynamics are primarily determined by the cell-motility variables $\tilde{F}$ and $\tilde{\tau}$, with only a minor dependency on $\tilde{\ell}$. When the self-propulsion force is much stronger than the capillary force, $\tilde{F} \gg 1$, the cells mainly interact with one another sterically and thus behave like a nematic gas (Fig. 3a, gas, and Supplementary Video 3). In the opposite limit $\tilde{F} \ll 1$, when capillary forces dominate the dynamics, groups of cells cannot split apart. The population thus rapidly coarsens into "droplets" of cells (Fig. 3a, droplets, and Supplementary Video 4). However, for intermediate regimes $\tilde{F} \sim 1$, we observe dynamic cell groups which are capable of both merging and splitting. In this regime, the structure and dynamics of these groups are determined by the cell's reversal period. For very long reversal times $\tilde{\tau} > 10$, polar clusters form (Fig. 3a, polar clusters, and Supplementary Video 5). At shorter reversal times $\tilde{\tau} \lesssim 10$, cells propelling in opposite directions can pass by each other but are typically unable to escape the range of capillary attraction before reversing their direction of motion. Here, quasi-one-dimensional streams form (Fig. 3a, streams, and

7

Supplementary Video 6). When the reversal time becomes shorter than the time for a cell to move the length of its body, the droplet phase is recovered (Fig. 3a, droplets). We identified these different phases by computing the cell-axis-aligned, equal-time spatial correlation $\widetilde{C}_\rho(\Delta r, t) = \langle \rho(\boldsymbol{r}, t)\rho(\boldsymbol{r} + \Delta \boldsymbol{r}, t)\rangle_{\boldsymbol{r}} / \langle \rho(\boldsymbol{r}, t)\rho(\boldsymbol{r}, t)\rangle_{\boldsymbol{r}}$ of cell density ρ (Fig. 3c) and using the mean and variance of these correlations to classify the dynamics on a phase diagram (Fig. 3e, and Supplementary Fig. 15 and Sec. III).

To explore the phase diagram experimentally, we captured time-lapse micrographs of various strains of *M. xanthus* on agarose hydrogels filled with rich media (CTT) and maintained at constant osmotic pressure with the previously described device. We segmented the resulting images using a convolutional neural network, and used the resulting cell masks to compute the mean and variance of $\widetilde{C}_\rho(\Delta r, t)$ (Fig. 3b,d, and Supplementary Fig. 16) over 3-hour windows. To achieve $\tilde{F} \gg 1$, we used our device to reduce the osmotic pressure P and flood the hydrogel surface. In this case, wild-type populations show qualitative similarity to the gas phase (Fig. 3b,d, and Supplementary Video 7). In the opposite limit, a non-motile strain of *M. xanthus*, corresponding to $\tilde{F} = 0$, forms cellular droplets as predicted by our simulations (Fig. 3b,d, and Supplementary Video 8). Also consistent with the simulation results, wild-type cells form stable, elongated streams, while the hypo-reversing Δ*frzZ* cells form polar clusters (Fig. 3b,d, and Supplementary Videos 9 and 10). These phases of nematic streams and polar clusters have also been reported in previous studies of *M. xanthus* [33, 34, 35]. Overall, our experiments showed all the phases predicted by the simulations.

Additional experiments using a mutant strain of *M. xanthus* lacking pili (Δ*pilA*) did not reveal stable streams, suggesting that additional interactions may be necessary to maintain these structures in the wild-type [36, 37, 38]. Similar to the results of Anderson and Ordal [20] wherein gliding of flavobacteria was dependent on a high degree of water availability, experiments in *F. johnsoniae* revealed only "gas" and "droplet" phases with the population rapidly transitioning between the two as water was moved into or out of the gel (Supplementary Fig. 16).

As would be expected for a "gas," gliding cells on saturated hydrogels can move freely relative to their nearest neighbors and appear to do so in an unstructured manner

(Supplementary Figs. 15 and 16). Conversely, cells in "droplets" are unable to readily pass each other because they are bound together by strong capillary forces (Supplementary Figs. 17 and 18). Interestingly, both the streams and polar clusters allow the cells to be tightly packed into spatially segregated groups while also moving relative to their neighbors (Fig. 3f,g, and Supplementary Video 11). These features might be crucial for the large-scale collective cell movement involved in the development of *M. xanthus* colonies (see Discussion).

Capillary forces influence the organization of bacterial populations.

To demonstrate how capillary forces influence the organization of bacterial populations, we used our device to impose changes in water availability on populations of *M. xanthus* and *F. johnsoniae* (Fig. 4a, and Supplementary Fig.19, and Videos 12 and 13). We tracked changes in population structure by computing the Delaunay triangulation between cell centroids (Fig. 4b, schematic). The distribution of edge lengths reveals both the relevant length scale(s) in the population – typically the approximate intercellular and intergroup spacings – as well as their heterogeneity across the population. Due to the initially low water availability in the substrate, cells are packed into small, dense groups. We then used our device to gradually saturate the substrate with water. This caused the surface to flood and the groups to break up, giving rise to a gaseous phase with a single apparent peak in the distribution of Delaunay edge lengths (Fig. 4b). Subsequently lowering the water reservoir below the substrate surface reimposed low water availability as in the start of the experiment. This change in water availability forced the cells back into densely packed groups (Fig. 4a,b).

Given the rod shape of individual cells, the packing into dense groups induces local nematic ordering (Fig. 4c). In the gas phase, nematic order is reduced as cells move in non-straight paths and only interact sterically for brief periods of time [39]. Our data shows that capillary attraction promotes dense cell packing and hence the emergence of nematic order in colonies of rod-shaped bacteria.

To further explore this idea, we asked how the morphology of both motile and non-motile cell groups was affected by capillary forces. To quantify the degree of elongation of cell

9

groups, we calculated a stream index, $S = \frac{1}{\sqrt{n}} \frac{\sum_{i=1}^{n} |sk(\mathcal{R}_i)|}{(\sum_{i=1}^{n} |\mathcal{R}_i|)^{1/2}}$ where n is the number of groups, $|\mathcal{R}_i|$ is the area of the ith group, and $|sk(\mathcal{R}_i)|$ is the length of the morphological skeleton of the ith group (Fig. 4d, schematic, and Supplementary Fig. 20 and Sec. IV). For both *M. xanthus* and *F. johnsoniae*, motile cell groups had higher stream indices than their non-motile counterparts, *M. xanthus* Δ*cglB*Ω*pilA* and *F. johnsoniae* Δ*sprB* (Fig. 4d, Supplementary Fig. 19). However, *M. xanthus*, whose cells exhibit periodic direction reversals, formed even more elongated structures than non-reversing *F. johnsoniae*, whose stream index was similar to that of the hypo-reversing Δ*frzZ* mutant of *M. xanthus* (Supplementary Fig. 19). The stream indices of the reversing, non-reversing, and non-motile cell groups are captured by simulations of a continuum model in the streams, polar clusters, and droplets phases, respectively (Fig. 4d and Supplementary Sec. IV).

Discussion

In total, our work suggests that lateral capillary forces promote the aggregation of terrestrial bacteria by condensing individual cells into tightly packed groups while allowing for relative motion of neighboring cells. These features allow motile bacterial cell groups to behave as active nematic liquid crystals, whose properties cells can exploit as part of their lifestyle [29, 40]. In *M. xanthus*, cells form collective groups to swarm, predate, and aggregate under different environmental conditions [41]. When nutrients are limited, topological defects of cell alignment in the starving swarm promote the formation of new cell layers [29], which stack on top of one another to form a nascent fruiting body [42]. While we focused on lateral motion in the current work, the capillary force also contains a strong vertical component (~100 nN on isolated cells) that pushes cells downward and may deform the underlying substrate. The same vertical capillary force opposes the extrusion of cells into new layers above the substrate, such that layer formation requires collective fluctuations in the vicinity of topological defects [40].

Nascent fruiting bodies are connected by quasi-one-dimensional streams [43, 44]. These streams may accelerate the growth of fruiting bodies by providing pathways that

guide cells to coalesce [35, 45, 46]. Our results indicate that capillary attraction may play a role in the developmental process by facilitating the population's condensation into nematic phases and the formation of inter-fruiting body streams. In the future, we will further investigate how vertical capillary forces influence the three-dimensional structure and developmental dynamics of bacterial colonies.

Our combined experimental and modeling results showed that water availability, which tunes the range and strength of capillary forces, controls the organization of bacterial colonies and gives rise to various collective phases formed by gliding bacteria on agarose hydrogels. Bacteria are known to produce surfactants that reduce the surface tension of the surrounding liquid [47, 48, 49], although whether or not this process is regulated and how it could affect the capillary forces we explored in this work remain unknown. More broadly, future work will be needed to understand how capillary forces affect bacterial growth, motion, and population morphology in more complex geometries such as those found in textiles and soil.

Acknowledgments

Bibliography

[1] D. Or, B. F. Smets, J. M. Wraith, A. Dechesne and S. P. Friedman, "Physical constraints affecting bacterial habitats and activity in unsaturated porous media – a review," *Advances in Water Resources,* vol. 30, no. 6-7, pp. 1505-1527, 2007.

[2] A. Stovicek, A. Azatyan, M. I. M. Soares and O. Gillor, "The Impact of Hydration and Temperature on Bacterial Diversity in Arid Soil Mesocosms," *Fronters in Microbiology,* vol. 8, 2017.

[3] M. Tuller and D. Or, "Water films and scaling of soil characteristic curves at low water contents," *Water Resources Research,* vol. 41, no. 9, 2005.

[4] H. Reichenbach, "The ecology of the myxobacteria," *Environmental Microbiology,* vol. 1, no. 1, pp. 15-21, 1999.

[5] G. A. Beattie, "Water Relations in the Interaction of Foliar Bacterial Pathogens with Plants," *Annual Review of Phytopathology,* vol. 49, pp. 533-555, 2011.

[6] A. Møllebjerg, L. G. Palmén, K. Gori and R. L. Meyer, "The bacterial life cycle in textiles is governed by fiber hydrophobicity," *Microbiology Spectrum,* vol. 9, p. e01185–21, 2021.

[7] L. E. Shinn, "A Cinematographic Analysis of the Motion of Colonies of B. alvei," *Journal of Bacteriology,* vol. 36, p. 419–422, 1938.

[8] H. C. Berg, "Swarming Motility: It Better Be Wet," *Current Biology,* vol. 15, no. 15, pp. R599-R600, 2005.

[9] L. Ping, Y. Wu, B. G. Hosu, J. X. Tang and H. C. Berg, "Osmotic Pressure in a Bacterial Swarm," *Biophysical Journal,* vol. 107, no. 4, pp. 871-878, 2014.

[10] Q. Wang, A. Suzuki, S. Mariconda, S. Porwollik and R. M. Harshey, "Sensing wetness: a new role for the bacterial flagellum," *The EMBO Journal,* vol. 24, pp. 2034-2042, 2005.

[11] A. Yang, W. S. Tang, T. Si and J. X. Tang, "Influence of Physical Effects on the Swarming Motility of Pseudomonas aeruginosa," *Biophysical Journal,* vol. 112, no. 7, pp. 1462-1471, 2017.

[12] R. Zhang, L. Turner and H. C. Berg, "The upper surface of Escherichia coli swarm is stationary," *Proceedings of the National Academy of Sciences,* vol. 107, no. 1, pp. 288-290, 2009.

[13] J. D. Partridge and R. M. Harshey, "More than Motility: Salmonella Flagella Contribute to Overriding Friction and Facilitating Colony Hydration during Swarming," *Journal of Bacteriology,* vol. 195, no. 5, pp. 919-929, 2013.

[14] A. M. Spormann, "Gliding motility in bacteria: insights from studies of Myxococcus xanthus," *Microbiology and Molecular Biology Reviews,* vol. 63, p. 621–641, 1999.

[15] J. Henrichsen, "Bacterial Surface Translocation: a Survey and a Classification," *Bacteriological Reviews,* vol. 36, no. 4, pp. 478-503, 1972.

[16] T. Mignot, J. W. Shaevitz, P. W. Hartzell and D. R. Zusman, "Evidence That Focal Adhesion Complexes Power Bacterial Gliding Motility," *Science,* vol. 315, no. 5813, pp. 853-856, 2007.

[17] A. J. Merz, M. So and M. P. Sheetz, "Pilus retraction powers bacterial twitching motility," *Nature,* vol. 407, pp. 98-102, 2000.

[18] R. G. E. Murray and R. H. Elder, "The predominance of counterclockwise rotation during swarming of Bacillus species," *Journal of Bacteriology,* vol. 58, p. 351–359, 1949.

[19] A. B. Semmler, C. B. Whitchurch and J. S. Mattick, " A re-examination of twitching motility in Pseudomonas aeruginosa," *Microbiology,* vol. 145, no. 10, 1999.

[20] R. L. Anderson and E. J. Ordal, "Cytophaga succinicans sp. n., a facultatively anaerobic, aquatic myxobacterium," *Journal of bacteriology,* vol. 81, p. 130–138, 1961.

[21] L. Turnbull and C. B. Whitchurch, "Motility Assay: Twitching Motility," in *Pseudomonas Methods and Protocols*, Humana Press, 2014, pp. 73-86.

[22] C. H. Ramos, E. Rodriguez-Sanchez, J. A. Arias del Angel, A. V. Arzola, M. Benitez, A. Escalante, A. Franci, G. Volpe and N. Rivera-Yoshida, "The environment topography alters the way to multicellularity in Myxococcus xanthus," *Science Advances,* vol. 7, no. 35, 2021.

[23] P. A. Kralchevsky and K. Nagayama, "Capillary forces between colloidal particles," *Langmuir,* vol. 10, p. 23–36, 1994.

[24] M. Cavallo, L. Botto, E. P. Lewandowsi, M. Wang and K. J. Stebe, "Curvature-driven capillary migration and assembly of rod-like particles," *Proceedings of the National Academy of Sciences,* vol. 108, no. 52, pp. 20923-20928, 2011.

[25] I. Gudavadze and E.-L. Florin, "Normal capillary forces on a spherical particle protruding from a thin liquid film and its application to swarming bacteria," *Colloids and Surfaces A: Physicochemical and Engineering Aspects,* vol. 636, p. 128176, 2022.

[26] A. De Lazzer, M. Dreyer and H. J. Rath, "Particle- surface capillary forces," *Langmuir,* vol. 15, p. 4551–4559, 1999.

[27] R. Tecon and D. Or, "Bacterial flagellar motility on hydrated rough surfaces controlled by aqueous film thickness and connectedness," *Scientific Reports,* vol. 6, p. 19409, 2016.

[28] A. Dechesne, D. Or, G. Gulez and B. F. Smets, "The porous surface model, a novel experimental system for online quantitative observation of microbial processes under unsaturated conditions," *Applied and Environmental Microbiology,* vol. 74, p. 5195–5200, 2008.

[29] K. Copenhagen, R. Alert, N. S. Wingreen and J. W. Shaevitz, "Topological defects promote layer formation in Myxococcus xanthus colonies," *Nature Physics,* vol. 17, p. 211–215, 2021.

[30] B. Sabass, M. D. Koch, G. Liu, H. A. Stone and J. W. Shaevitz, "Force generation by groups of migrating bacteria," *Proceedings of the National Academy of Sciences,* vol. 114, no. 28, pp. 7266-7271, 2017.

[31] M. Clausen, V. Jakovljevic, L. Søgaard-Andersen and B. Maier, "High-Force Generation Is a Conserved Property of Type IV Pilus Systems," *Journal of Bacteriology,* vol. 191, no. 14, pp. 4633-4638, 2009.

[32] R. Balagam, D. B. Litwin, F. Czerwinski, M. Sun, H. B. Kaplan, J. W. Shaevitz and O. A. Igoshin, " Myxococcus xanthus gliding motors are elastically coupled to the substrate as predicted by the focal adhesion model of gliding motility," *PLoS Computational Biology,* vol. 10, no. 5, 2014.

[33] F. Peruani, J. Starruß, V. Jakovljevic, L. Søgaard-Andersen, A. Deutsch and M. Bär, "Collective motion and nonequilibrium cluster formation in colonies of gliding bacteria," *Physical review letters,* vol. 108, p. 098102, 2012.

[34] J. Starruß, F. Peruani, V. Jakovljevic, L. Søgaard-Andersen, A. Deutsch and M. Bär, "Pattern-formation mechanisms in motility mutants of Myxococcus xanthus," *Interface focus,* vol. 2, p. 774–785, 2012.

[35] S. Thutupalli, M. Sun, F. Bunyak, K. Palaniappan and J. W. Shaevitz, "Directional reversals enable Myxococcus xanthus cells to produce collective one-dimensional streams during fruiting-body formation," *Journal of The Royal Society Interface,* vol. 12, p. 20150049, 2015.

[36] R. Balagam and O. A. Igoshin, "Mechanism for collective cell alignment in Myxococcus xanthus bacteria," *PLoS computational biology,* vol. 11, p. e1004474, 2015.

[37] R. Balagam, P. Cao, G. P. Sah, Z. Zhang, K. Subedi, D. Wall and O. A. Igoshin, "Emergent Myxobacterial Behaviors Arise from Reversal Suppression Induced by Kin Contacts," *mSystems,* vol. 6, no. 6, 2021.

[38] A. Janulevicius, M. van Loosdrecht and C. Picioreanu, "Short-Range Guiding Can Result in the Formation of Circular Aggregates in Myxobacteria Populations," *PLoS Computational Biology,* vol. 11, no. 8, 2015.

[39] G. Liu, A. Patch, F. Bahar, D. Yllanes, R. D. Welch, M. C. Marchetti, S. Thutupalli and J. W. Shaevitz, "Self-Driven Phase Transitions Drive Myxococcus xanthus Fruiting Body Formation," *Physical Review Letters,* vol. 122, no. 248102, 2019.

[40] E. Han, C. Fei, R. Alert, K. Copenhagen, M. D. Koch, N. S. Wingreen and J. W. Shaevitz, "Local polar order controls mechanical stress and triggers layer formation in developing Myxococcus xanthus colonies," *arXiv,* 2023.

[41] J. Muñoz-Dorado, F. J. Marcos-Torres, E. García-Bravo, A. Moraleda-Muñoz and J. Pérez, "Myxobacteria: Moving, Killing, Feeding, and Surviving Together," *Fronteris in Microbiology,* vol. 7, 2016.

[42] P. D. Curtis, R. G. Taylor, R. D. Welch and L. J. Shimkets, "Spatial Organization of Myxococcus xanthus during Fruiting Body Formation," *Journal of Bacteriology,* vol. 189, no. 24, pp. 9126-9130, 2007.

[43] L. Søgaard-Andersen, M. Overgaard, S. Lobedanz, E. Ellehauge, L. Jelsbak and A. A. Rasmussen, " Coupling gene expression and multicellular morphogenesis during fruiting body formation in Myxococcus xanthus," *Molecular Microbiology,* vol. 48, no. 1, pp. 1-8, 2003.

[44] D. Kaiser and R. Welch, "Dynamics of fruiting body morphogenesis," *Journal of Bacteriology,* vol. 186, pp. 919-927, 2004.

[45] L. Jelsbak and L. Søgaard-Andersen, "Pattern formation: fruiting body morphogenesis in Myxococcus xanthus," *Current Opinion in Microbiology,* vol. 3, no. 6, pp. 637-642, 2000.

[46] O. Sozinova, Y. Jiang, D. Kaiser and M. Alber, "A three-dimensional model of myxobacterial aggregation by contact-mediated interactions," *Proceedings of the National Academy of Sciences,* vol. 102, no. 32, pp. 11308-11312, 2005.

[47] D. B. Kearns, "A field guide to bacterial swarming motility," *Nature Reviews Microbiology,* vol. 8, no. 9, pp. 634-644, 2010.

[48] S. T. Islam, I. Vergara Alvarez, F. Saidi, A. Guiseppi, E. Vinogradov, G. Sharma, L. Espinosa, C. Morrone, G. Brasseur, J.-F. Guillemot, A. Benarouche, J.-L. Bridot and G. Ravicoularamin, "Modulation of bacterial multicellularity via spatio-specific polysaccharide secretion," *PLoS Biology,* vol. 46, pp. 143-154, 2020.

[49] T. E. Angelini, M. Roper, R. Kolter, D. A. Weitz and M. P. Brenner, "Bacillus subtilis spreads by surfing on waves of surfactant," *Proceedings of the National Academy of Sciences,* vol. 106, no. 43, pp. 18109-18113, 2009.

[50] M. R. Warren, H. Sun, Y. Yan, J. Cremer, B. Li and T. Hwa, "Spatiotemporal establishment of dense bacterial colonies growing on hard agar," *Elife,* vol. 8, p. e41093, 2019.

[51] Y. Wu and H. C. Berg, "Water reservoir maintained by cell growth fuels the spreading of a bacterial swarm," *Proceedings of the National Academy of Sciences,* vol. 109, no. 11, pp. 4128-4133, 2012.

[52] J. R. Mireles, A. Toguchi and R. M. Harshey, "Salmonella enterica Serovar Typhimurium Swarming Mutants with Altered Biofilm-Forming Abilities: Surfactin Inhibits Biofilm Formation," *Journal of bacteriology,* vol. 183, no. 20, pp. 5848-5854, 2001.

[53] B. G. Chen, L. Turner and H. C. Berg, "The Wetting Agent Required for Swarming in Salmonella enterica Serovar Typhimurium Is Not a Surfactant," *Journal of Bacteriology,* vol. 189, no. 23, pp. 8750-8753, 2007.
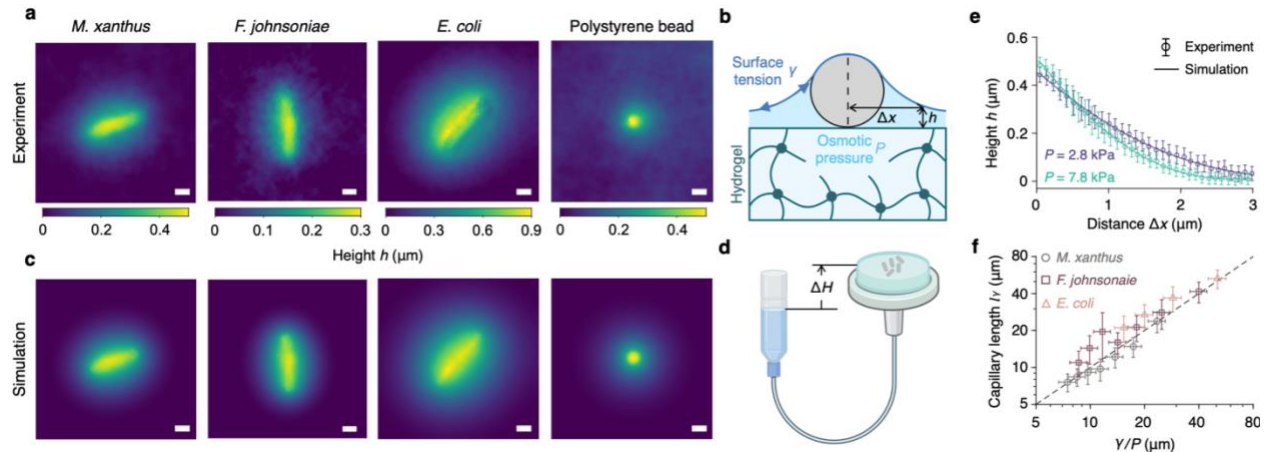
*Figure 1 Water menisci around bacteria and colloidal particles on a hydrated substrate. **(a)** Measured height profiles around single bacteria and a 1 µm diameter polystyrene bead deposited onto an agarose hydrogel. **(b)** Schematic of the modeled water meniscus that results from the balance of surface tension (dark blue) and osmotic pressure difference P (light blue) across the surface of the hydrogel substrate (green meshwork). The model assumes that the water meniscus completely wets the surface of a cell (gray circle with black outline; cross-section). h denotes the height of water above the substrate surface and Δx denotes the horizontal distance to the midline of the cell. See Methods for details. **(c)** Height profiles of the modeled wetting menisci. Scale bar is the same for all images and represents 1 µm. **(d)** Illustration of the experimental device used to control water pressure. A reservoir of media is connected to the bottom of the agarose hydrogel onto which cells are deposited for each experiment. The height difference, ΔH, between the top of the hydrogel surface and the top of the reservoir sets an additional hydrostatic pressure that tunes the net osmotic pressure difference across the hydrogel. **(e)** Water height profiles h(Δx), measured in the midplane perpendicular to the long axis of the cell, at the designated values of P. The reference value P_0 is estimated by fitting the height profile at ΔH = 0, and P is inferred using the relation $P = P_0 + \rho_w g \Delta H$. Circles with error bars indicate experimental measurements (mean +/- standard deviation). Solid curves indicate fits of simulation results. **(f)** Capillary lengths inferred from the height profiles of the wetting menisci at varying values of P. Different colors and symbols represent the designated bacterial species. Black dashed line indicates the line y = x. See Supplementary Information for details.*
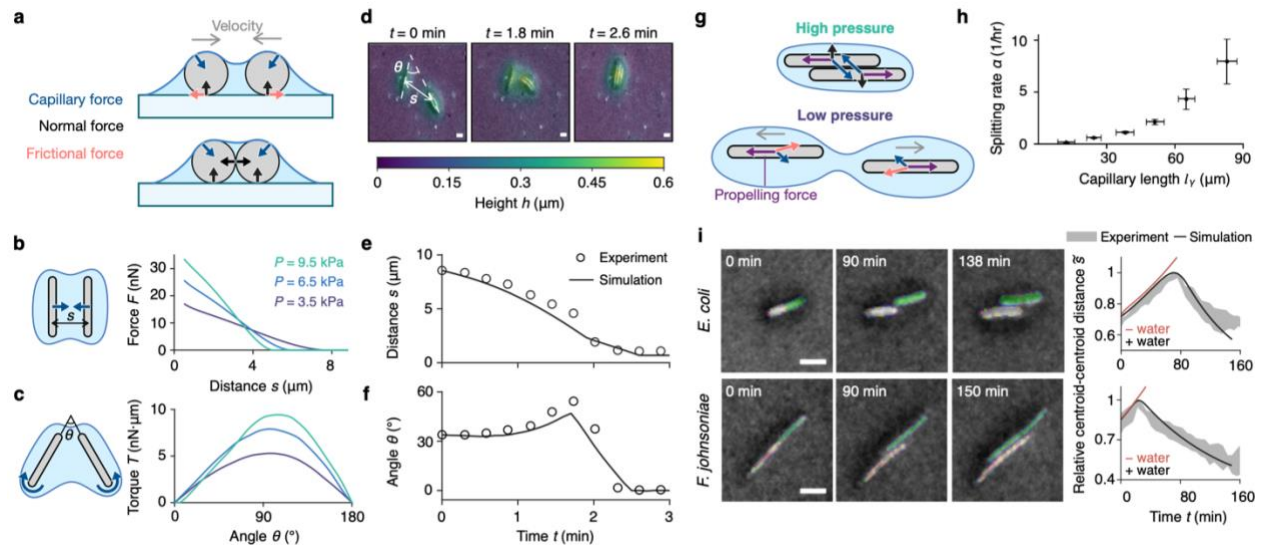
16

*Figure 2 Capillary forces promote mergers and hinder separation of bacterial cells. (**a**) Cross-section schematics showing the forces exerted on non-motile cells. Color code as in Fig. 1**b**. Black arrows indicate normal forces exerted by either the substrate or the cells. Dark blue arrows indicate capillary forces. Pink arrows indicate frictional forces between the cells and the substrate. Gray arrows indicate cell velocities. Two nearby cells with overlapping water menisci (top) are pulled together by capillary forces (bottom). (**b**) Left: top-view schematic of two parallel cells at a distance, s. Color code as in **a**. Right: magnitude of the capillary force at varying distances s for the designated values of osmotic pressure, P. (**c**) Left: top-view schematic of two nearby cells with a fixed center-to-center distance and forming an angle θ. Color code as in **a**. Right: magnitude of the capillary torque at varying angles for the designated values of osmotic pressure, P. The center-to-center distance is set to 5.1 μm. In **b** and **c**, forces and torques are computed for cells with radius R_c = 0.3 μm and length l_c = 4.5 μm. (**d**) Water height profiles, overlaid on reflectance images of* M. xanthus *cells (ΔcglBΩpilA) that merge due to overlapping wetting menisci. Scale bars: 1 μm. (**e, f**) Time evolution of (**e**) center-to-center distance s and (**f**) relative orientation angle θ, as indicated for the two cell groups in **d**. Circles indicate experimental measurements and solid curves indicate fitted results of agent-based simulations. (**g**) Top-view schematics of the forces exerted on self-propelling cells for high osmotic pressure (top) and low osmotic pressure (bottom). Color code as in **a**. Purple arrows indicate self-propelling forces along cell's long axis. (**h**) The rate at which pairs of motile wild-type* M. xanthus *cells separate apart, as schematized in **g**, increases with the measured capillary length. Error bars are boot-strapped standard deviation. (**i**) Stereotyped dynamics of non-motile, dividing cells due to capillary forces. Post-division pairs of cells undergo a period of end-on-end growth before capillary forces trigger them to slide relative to each other. Example pairs are shown for* E. coli *and* F. johnsonaie *(ΔsprB) and the centroid-to-centroid distance, relative to its largest value for each pair, is plotted on the right. Relative centroid-centroid distances are averaged over N=5 and N=10 cells for* E. coli *and* F. johnsonaie*, respectively. Shaded regions represent mean +/- the standard*

17

*deviation of the experimental data. Black curves and red curves represent the simulation results with and without water, respectively. Model parameters used in **e**, **f**, and **i** are listed in Supplementary Table 2.*
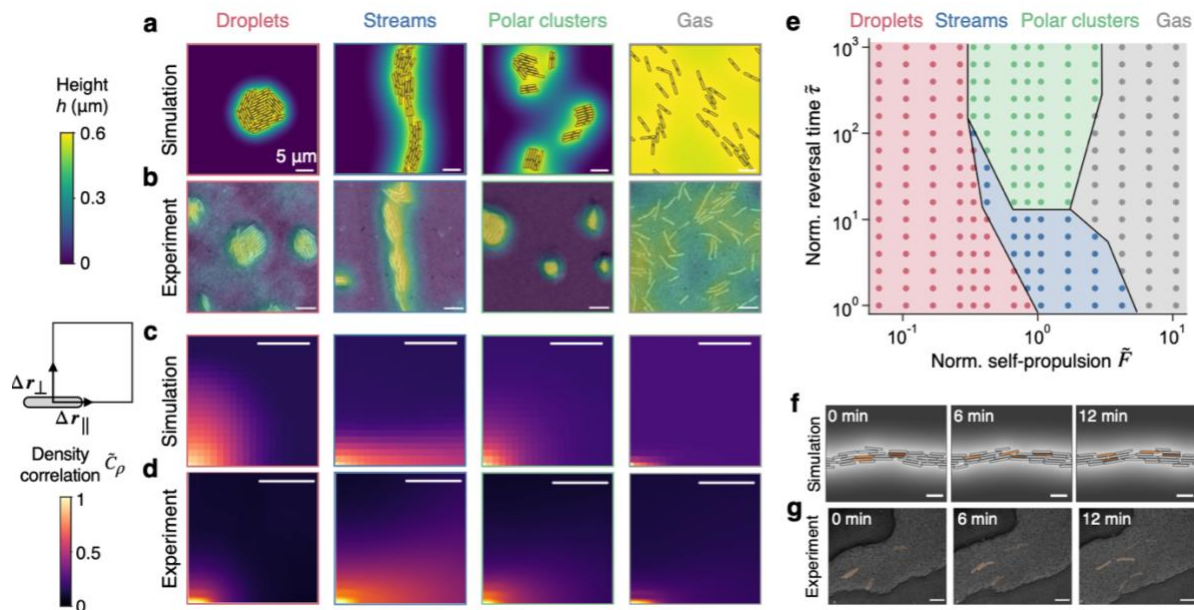
*Figure 3 Capillary attraction and cell motility lead to multiple phases of collective cell organization. (**a**,**b**) Heatmaps of water-height profiles for typical snapshots of the designated phases in (**a**) simulations and (**b**) experiments. (**c**, **d**) Time-averaged steady-state spatial density autocorrelation for the designated phases in (**c**) simulations and (**d**) experiments. (**e**) A section of the phase diagram for simulated cells. Color code as in **a**. Phase boundaries are drawn by hand. The normalized capillary length is set to $\tilde{\ell} = 100$. (**f**, **g**) Snapshots of (**f**) simulation and (**g**) experimental timelapse in which adjacent cells in a stream are labeled by the same color at time t=0 min and tracked over time. Adjacent cells can slide relative to one another, allowing the pairs to move apart and exchange neighbors. In all panels, scale bars: 5 µm.*
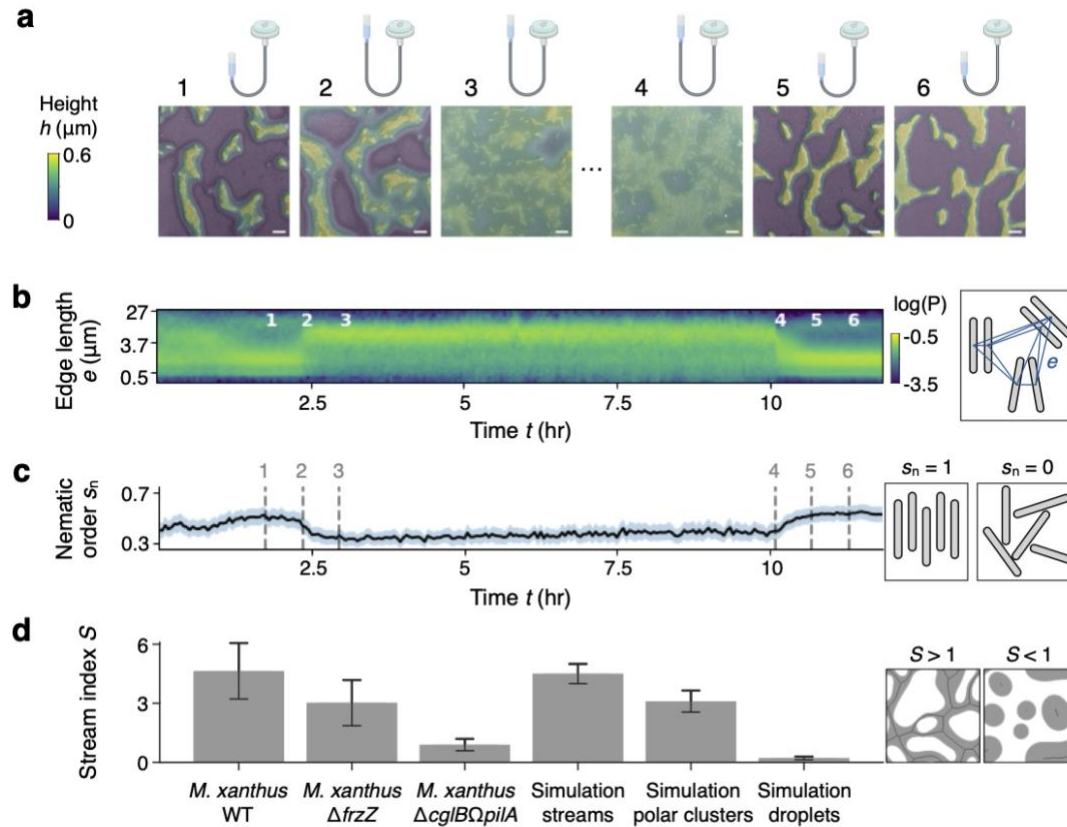
Figure 4 Water availability controls the organization of colonies of M. xanthus*. (a) Experimental time series in which cells initially deposited onto an agarose pad are subjected first to increasing water availability until a gas phase is reached (snapshots 1-3). Hours later, water is drained from the substrate, and the colony goes back to forming tightly packed, nematically ordered groups and streams (snapshots 4-6). Time points (1-6) are marked in **b** and **c**. (**b**) Kymograph of the probability density function (P) of Delaunay edge lengths over time for the time series shown in **a**. Schematic on the right demonstrates how edges (blue lines) are calculated for a small population of cells (gray). At early times and late times in the experiment, low water availability at the surface constrains cells into tightly packed, well-separated groups. At intermediate times, high water availability permits cells to be further spaced apart and homogenizes the population. (**c**) Strength of nematic order s_n, calculated within groups of 8 nearest-neighboring cells and averaged over the population at each time. Shaded regions indicate variance. Right: schematics showing a small population of cells with perfect nematic order ($s_n=1$) or no order ($s_n=0$). (**d**) Stream indices for wild-type M. xanthus, a hypo-reversing mutant of M. xanthus (ΔfrzZ), a non-motile mutant of M. xanthus (ΔcglBΩpilA), and simulated streams, polar clusters, and droplets. Error bars for the experimental results are standard deviations for each corresponding mean calculated over N=25, 9, and 7 frames taken from 2, 3, and 1 separate timelapses, respectively. Error bars for the simulation results are*

20

standard deviations for each corresponding mean calculated over N=50 frames during the steady state of

1 timelapse. Right: schematics showing populations that form streams (S > 1) or droplets (S < 1).

Supplementary Information for

Capillary interactions drive the self-organization of bacterial colonies

Matthew E. Black[a,†], Chenyi Fei[a,†], Ricard Alert[b,c,d], Ned S. Wingreen[a,e,*], Joshua W. Shaevitz[a,f,*]

[a] Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA
[b] Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany
[c] Center for Systems Biology Dresden, 01307 Dresden, Germany
[d] Cluster of Excellence Physics of Life, 01062 Dresden, Germany
[e] Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA
[f] Joseph Henry Laboratories of Physics, Princeton University, Princeton, NJ 08544, USA

[†]These authors contributed equally.
[*]To whom correspondence may be addressed. Email: wingreen@princeton.edu or shaevitz@princeton.edu

TABLE OF CONTENTS

LIST OF SUPPLEMENTARY FIGURES

LIST OF SUPPLEMENTARY TABLES

I. MODELING WATER MENISCI AROUND CELLS

A. A free-energy model for the water meniscus

Laser-scanning microscopy reveals that there are intensity halos around *M. xanthus* cells grown on agar and these halos display continuously decreasing height from the centerlines of the cells toward the agar surface far from the cells (main Fig. 1). Our main hypothesis is that these halos correspond to water menisci formed around the cells. To account for the shape of the water meniscus, we propose a minimal model based on a free energy that takes into account the surface energy of the air-water interface and the energy penalty of extracting water from the agar substrate. Denoting by $h(x, y)$ the height of water above the surface of the agar substrate, the total free energy of the water is given by

$$\mathcal{F}[h] = \gamma \int \sqrt{1 + |\nabla h|^2}\, \mathrm{d}x\mathrm{d}y + P \int h\, \mathrm{d}x\mathrm{d}y, \tag{S1}$$

where $\nabla = (\partial_x, \partial_y)$, γ denotes the surface tension of water, and P denotes the per-volume energy penalty of extracting water out of the substrate, i.e., the osmotic pressure difference across the substrate surface (see Sec. I E for details). Note that γ and P together yield a capillary length $l_\gamma = \gamma/P$ that controls the range of capillary interactions (see Sec. I G). We assume that the equilibrium shape of the water meniscus $h^*(x, y)$ minimizes the free-energy functional Eq. (S1), subject to the constraint that the height of water cannot be lower than the upper cell surfaces z_{upper} or the substrate surface $z = 0$. To express this constraint mathematically, we next specify the geometry of the modeled cells.

We treat each cell as a spherocylinder with cap radius R_c and cylindrical length l_c. The position of cell i is determined by its center-of-mass coordinates $(x_c^{(i)}, y_c^{(i)}, z_c^{(i)})$ and its orientation $(n_x^{(i)}, n_y^{(i)}, n_z^{(i)})$. For simplicity, we assumed that cells are parallel to the substrate surface and the surface is tangent to the cylindrical part of the cell, i.e., $z_c^{(i)} \equiv R_c$ and $n_z^{(i)} \equiv 0$. Thus, the position of each cell i is determined by its center-of-mass coordinates $x_c^{(i)}$ and $y_c^{(i)}$, and its orientation angle $\theta^{(i)} = \arccos n_x^{(i)} = \arcsin n_y^{(i)}$. The height profile of cell i's upper surface is given by

$$z^{(i)}_{\mathrm{upper}}(x, y) = \begin{cases} R_c + \left[R_c^2 - s_\parallel^2(x, y)\right]^{1/2}, & \text{if } |s_\perp| < l_c/2 \text{ and } |s_\parallel| \leqslant R_c, \\ R_c + \left[R_c^2 - d_+^2(x, y)\right]^{1/2}, & \text{if } s_\perp \geqslant l_c/2 \text{ and } d_+ \leqslant R_c, \\ R_c + \left[R_c^2 - d_-^2(x, y)\right]^{1/2}, & \text{if } s_\perp \leqslant -l_c/2 \text{ and } d_- \leqslant R_c, \\ 0, & \text{otherwise}, \end{cases} \tag{S2}$$

where $s_\parallel(x, y) = (x - x_c)\cos\theta^{(i)} + (y - y_c)\sin\theta^{(i)}$ and $s_\perp(x, y) = -(x - x_c)\sin\theta^{(i)} + (y - y_c)\cos\theta^{(i)}$ denote, respectively, the coordinates of (x, y) along or perpendicular to the cell's long axis, and $d_\pm(x, y) = \left[(x - x_c \mp l_c\cos\theta^{(i)}/2)^2 + (y - y_c \mp l_c\sin\theta^{(i)}/2)^2\right]^{1/2}$ denotes the distance from (x, y) to the $\pm$ ends of the cell. Therefore, the constraint on h can be expressed as

$$h(x, y) \geqslant \max(\{z^{(i)}_{\mathrm{upper}}(x, y)\}, 0). \tag{S3}$$

B. Examples of equilibrium height profiles

We perform the constrained energy minimization described in Sec. I A numerically to solve for the equilibrium height profiles. In the simulations, we vary the normalized capillary length $\tilde{\ell} = l_\gamma/R_c$ to change the width of the water meniscus. To verify our simulation results, we consider two special cases below where analytical solutions can be obtained.

Water meniscus around a 2D circle. We start by considering a case of reduced dimensionality in the plane $y = 0$. We consider a 2D circular "cell" of radius R_c, whose center is located at $x = 0$, $z = R_c$. In this case, the total free energy in Eq. (S1) is given by $\mathcal{F}[h(x)] = \gamma \int (\sqrt{1 + (h')^2})dx + P \int h\, dx$ where we define $f' \equiv df/dx$ for an arbitrary

function $f(x)$. The free-energy minimum is achieved when the functional derivative $\delta\mathcal{F}/\delta h = 0$, which yields

$$\left(\frac{h'}{\sqrt{1+(h')^2}}\right)' = \frac{h''}{\left[1+(h')^2\right]^{3/2}} = \frac{P}{\gamma} = l_\gamma^{-1}. \tag{S4}$$

The solution of Eq. (S4) is a circular arc of radius l_γ. Taking into account that the height profile of water must be continuous, we obtain[1]

$$h^*(x) = \begin{cases} R_c + \sqrt{R_c^2 - x^2} & 0 \leqslant |x| \leqslant R_c \sin\theta^* \\ l_\gamma - \sqrt{l_\gamma^2 - [x - (R_c + l_\gamma)\sin\theta^*]^2} & R_c\sin\theta^* \leqslant |x| \leqslant (R_c + l_\gamma)\sin\theta^* \\ 0 & |x| > (R_c + l_\gamma)\sin\theta^* \end{cases} \tag{S5}$$

where $\theta^* = \arccos\frac{l_\gamma - R_c}{l_\gamma + R_c}$. Our simulation results match closely with the analytical solution (Supplementary Fig. 1).

Water meniscus around a 3D sphere. A similar procedure as above can be performed to compute the shape of the water meniscus around a 3D spherical "cell" of radius R_c whose center is located at $x = 0, y = 0, z = R_c$. Due to polar symmetry, the water height profile $h = h(r)$ is only a function of $r = (x^2 + y^2)^{1/2}$. The free-energy functional is given by $\mathcal{F}[h(r)] = \gamma \int 2\pi r\sqrt{1+(h')^2}\,dr + P\int 2\pi r h\,dr$, and the optimization $\delta\mathcal{F}/\delta h = 0$ yields

$$\frac{1}{r}\left(\frac{rh'}{\sqrt{1+(h')^2}}\right)' = \frac{1}{l_\gamma} \Rightarrow \frac{rh'}{\sqrt{1+(h')^2}} = \frac{r^2}{2l_\gamma} + c_1, \tag{S6}$$

where c_1 is an integration constant. The explicit expression for the solution of Eq. (S6) is cumbersome and contains multiple elliptical functions. Nevertheless, c_1 can be determined from $h' = 0$ at $r = r_{\max} \equiv \mathrm{argmin}[h(r) = 0]$, and Eq. (S6) can thus be rewritten as

$$u(r) \equiv -\frac{h'}{\sqrt{1+(h')^2}} = \frac{1}{2l_\gamma}(r_{\max}^2/r - r). \tag{S7}$$

Indeed, the simulation results are in good agreement with this analytical expression (Supplementary Fig. 1).

C. Estimating the capillary length from the equilibrium height profiles

In this section, we discuss how to estimate the capillary length l_γ from the measured height profile of the water meniscus in experiment (Fig. 1). We measured the height profile $h^*(d)$ in the midplane perpendicular to the long axis of the cell, where d denotes the horizontal distance to the long axis. Note that for a spherocylindrical cell that is located at $x_c = 0, y_c = 0$ and whose orientation angle is $\theta = \pi/2$, we have $h^*(d) \equiv h^*(x = \pm d, y = 0)$.

As discussed in Sec. I B **Water meniscus around a 2D circle**, for an infinitely long cell, $h^*(d)$ corresponds a circular arc of radius l_γ, and thus l_γ can in theory be estimated from the radius of curvature R_κ of $h^*(d)$, i.e., $\hat{l}_\gamma = R_\kappa \equiv \left[1+(h')^2\right]^{3/2}/h''|_{h=h^*}$. In practice, there are two problems: (1) calculating derivatives of a noisy measurement $h^*(d)$ can be unreliable, and (2) the experimental cells are not infinitely long. To address the first problem, we measure the distance $d_{1/2}$ at which h^* drops below $R_c/2$ and estimate the radius of curvature R_κ to be $R_\kappa = d_{1/2}^2/R_c$ assuming that $R_\kappa \gg R_c$ (see Supplementary Fig. 2a for the geometry). We further verified that, for all the bacterial cells tested, the condition $R_\kappa \gg R_c$ is satisfied (main Fig. 1). To address the second problem, we simulate the equilibrium height profiles for varying values of cell aspect ratio $l_c/(2R_c)$ and capillary strength $(PR_c)/\gamma$, and compute $R_\kappa = d_{1/2}^2/R_c$ for the simulated profiles (Supplementary Fig. 2b). We find that, indeed, $R_\kappa = l_\gamma$ when $l_c/(2R_c) = \infty$ and $R_\kappa = l_\gamma/\alpha_{\mathrm{corr}}$ ($\alpha_{\mathrm{corr}} > 1$) for finite $l_c/(2R_c)$. Thus, we infer the capillary length from the experimental measurements using $\hat{l}_\gamma = \alpha_{\mathrm{corr}}R_\kappa$ (with α_{corr} obtained from Supplementary Fig. 2b) and ignore the uncertainty of the correction factor α_{corr} for simplicity.

---

[1] This expression only works for $l_\gamma \geqslant R_c$. When $l_\gamma < R_c$, i.e., the central angle θ^* of the two circular arcs in Eq. (S5) become larger than $\pi/2$, and both $\pm$ signs should be kept in front of the square-root terms.

D. Model parameters

We estimate the model parameters in our model of water meniscus directly from experiment:

Cell length and radius: We estimate the cap radius R_c by half the maximal height measured in experiment. The cylindrical length l_c is determined from the bright-field images of bacterial cells. We choose $R_c = 0.3$ μm and $l_c = 4.5$ μm for *M. xanthus* cells, $R_c = 0.4$ μm and $l_c = 2.4$ μm for *E. coli* cells, and $R_c = 0.15$ μm and $l_c = 2.7$ μm for *F. johnsoniae* cells. For the full simulation of *M. xanthus* cells, we set $R_c = 0.3$ μm and we draw l_c from a uniform distribution $\mathcal{U}[3.6\ \mathrm{μm}, 5.4\ \mathrm{μm}]$.

Water surface tension: The surface tension γ of the liquid media is measured experimentally. See Sec. VI C for details. The values of γ are reported in Table 3.

Osmotic pressure: To estimate the values of P under different conditions, we assume that $P = P_0 + \rho_w g \Delta H$, where ρ_w denotes the mass density of water, g denotes the gravitational acceleration constant, ΔH denotes the height difference between the gel surface and the water level of the syringe connected to the gel (Fig. 1d), and P_0 denotes the osmotic pressure difference across the gel surface at $\Delta H = 0$. The value of P_0 is determined by fitting the height profiles of the modeled water menisci to those measured in experiments with $\Delta H = 0$. We consider two fitting procedures:

(1) We compare the 3D height profiles $h(x, y)$ of the water menisci around N individual cells between experiment and theory. For each individual cell i, we denote by $h_{\exp}^{(i)}(x, y)$ and $h_{\sim}^{(i)}(x, y)$ the experimental measurements and the simulation results, respectively. Since the experimental cells are not perfect spherocylinders, we minimize the free energy Eq. (S1) using the constraint

$$h_{\sim}^{(i)}(x, y) \geqslant \begin{cases} h_{\exp}^{(i)}(x, y), & \text{if } h_{\exp}^{(i)}(x, y) \geqslant \frac{3}{4} \max[h_{\exp}^{(i)}(x, y)], \\ 0, & \text{if } h_{\exp}^{(i)}(x, y) < \frac{3}{4} \max[h_{\exp}^{(i)}(x, y)] \end{cases} \tag{S8}$$

to obtain the simulated equilibrium water meniscus $h_{\sim}^{(*,i)}$. The fitting error is measured by the root-mean-square height difference $\epsilon_h = \sqrt{\langle (h_{\sim}^{(*,i)} - h_{\exp}^{(i)})_{\Omega_\epsilon}^2 \rangle}$ on the region $\Omega_\epsilon = \{(x, y) | h_{\exp}^{(i)}(x, y) < \frac{3}{4} \max[h_{\exp}^{(i)}(x, y)]\}$. We vary the value of P_0 and obtain the optimal fitting value $P_0^{(*,i)}$ that minimizes ϵ_h. Finally, the value of P_0 is determined by $P_0^* = \frac{1}{N} \sum_{i=1}^{N} P_0^{(*,i)}$ and the uncertainty of P_0 is given by $\sigma^{P_0} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left[P_0^{(*,i)} - P_0^* \right]^2}$.

(2) Alternatively, we compare the modeled and measured water height profiles in the midplane perpendicular to the long axis of the cell (Fig. 1e). We average the measured height profiles over multiple cells at various distances d from the cell midline to obtain $\bar{h}_{\exp}(d) \pm \sigma_{\exp}^h(d)$. We set $R_c = \max[\bar{h}_{\exp}(d)]/2$, and simulate the equilibrium height profile $h_{\sim}^*(d; P_0)$ with varying values of P_0. To quantify the fitting error, we compute the negative log-likelihood function $\mathcal{L}(P_0) = \frac{1}{2} \sum_d \left[\frac{h_{\sim}^*(d; P_0) - \bar{h}_{\exp}(d)}{\sigma_{\exp}^h(d)} \right]^2$. Subsequently, the value of P_0 is determined by $P_0^* = \operatorname{argmin} \mathcal{L}(P_0)$ and the uncertainty σ^{P_0} satisfies $\mathcal{L}(P_0^* \pm \sigma^{P_0}) = \min \mathcal{L}(P_0) + 1/2$.

As shown in Supplementary Fig. 5, the two fitting procedures lead to similar estimations of P_0. Specifically, fitting the full 3D height profile yields $P_0 = 2.69 \pm 0.61$ kPa and fitting the midplane 2D height profile yields $P_0 = 2.81 \pm 0.29$ kPa.

E. Additional notes on the osmotic pressure difference across the gel surface

In Sec. I D, we find that there is a nonzero osmotic pressure difference P_0 across the gel surface even when $\Delta H = 0$. We hypothesize that there is a vertical osmotic pressure gradient across the gel due to evaporation — the top of the gel has a higher polymer volume fraction, and hence a higher osmotic pressure, than the bottom of the gel (Supplementary Fig. 6a). In this subsection, we explore whether this gradient can account for a nonzero P_0. Since the gel is able to swell or shrink, we consider two configurations of the gel, one that represents the reference configuration and one that corresponds to the current configuration. We denote by Z_0 and z the vertical coordinates in the reference configuration and the current configuration, respectively.

Governing equations. Consider an infinitesimal horizontal slice of the gel that is located at $[Z_0, Z_0 + \mathrm{d}Z_0]$ in the reference configuration and at $[z, z + \mathrm{d}z]$ in the current configuration. Since the volume of polymers in this slice of the gel is conserved, we obtain $\phi_0 \mathrm{d}Z_0 = \phi \mathrm{d}z$, where ϕ_0 and ϕ denote, respectively, the volume fraction of polymers in the reference and current configurations, and we assume ϕ_0 to be independent of Z_0 for simplicity. In addition, the mass conservation of water yields $\partial(\Delta V_{\mathrm{w}})/\partial t = (J|_{Z_0} - J|_{Z_0 + \mathrm{d}Z_0})S$, where the volume of water in this gel slice can be expressed as $\Delta V_{\mathrm{w}} = S \mathrm{d}Z_0(\phi_0/\phi - \phi_0)$, S denotes the surface area of the gel, and J denotes the flux of water. In the current configuration, J is given by $J = \kappa \frac{\partial \Pi}{\partial z}$ where κ denotes the permeability of the gel and Π denotes the osmotic pressure of the gel. Introducing $\mathrm{d}z = \phi_0 \mathrm{d}Z_0/\phi$ into the expression for J, we obtain $J = \kappa \frac{\partial \Pi}{\partial Z_0} \frac{\phi}{\phi_0}$. Using this result, the continuity equation for the aqueous phase of the gel is given by

$$\frac{\partial(\phi_0/\phi)}{\partial t} = -\kappa \frac{\partial}{\partial Z_0}\left(\frac{\phi}{\phi_0}\frac{\partial \Pi}{\partial Z_0}\right). \tag{S9}$$

To close the equation, we use the Flory-Rehner theory to model the osmotic pressure of the hydrogel [1, 2], i.e., $\Pi(\phi) = -\frac{kT}{v_c}\left[\ln(1-\phi) + \phi + \chi\phi^2\right] - G_0\phi^{1/3}$, where the first term results from the free energy of mixing and the second term results from the elastic energy of stretching the polymers. Here, $\frac{kT}{v_c}$ denotes the thermal free energy density, χ denotes the solvent-polymer interaction parameter, and G_0 is the shear modulus of the gel in the reference state. For agarose gels, we assume $\chi \approx 1/2$ and obtain $\Pi(\phi) = \frac{kT}{3v_c}(\phi^3 - \phi_{\mathrm{eq}}^{8/3}\phi^{1/3})$ where $\phi_{\mathrm{eq}} = (\frac{3G_0 v_c}{kT})^{3/8}$ denotes the polymer volume fraction of a freely swollen gel. Finally, we set the reference state $\phi_0 = \phi_{\mathrm{eq}}$ for simplicity.

Boundary conditions. In experiment, the bottom of the gel is connected with a water reservoir, and evaporation of water occurs at the top surface of the gel. To describe this experimental set up, we choose the following boundary conditions: $\Pi = C = \rho_{\mathrm{w}}g\Delta H$ at the bottom of the gel $Z_0 = 0$, and $J = J_{\mathrm{ev}}$ at the top of the gel $Z_0 = L_0$. Here, C can be changed by tuning ΔH in experiment, and J_{ev} has been measured previously to be $J_{\mathrm{ev}} = 20$ nm/s [3].

Nondimensionalization. We define dimensionless variables $\tilde{\phi} = \phi/\phi_0$, $\tilde{\Pi} = \Pi/\Pi_0$, $\tilde{t} = t/T_0$, and $\tilde{Z} = Z_0/L_0$, where $\Pi_0 = 1$ kPa, $T_0 = 1$ min, and $L_0 = 6$ mm are the relevant scales for the experimental system. Upon non-dimensionalize Eq. (S9), we derive

$$\frac{\partial(\tilde{\phi}^{-1})}{\partial \tilde{t}} = -\tilde{K}\frac{\partial}{\partial \tilde{Z}}\left[\tilde{\phi}\frac{\tilde{Q}\partial(\tilde{\phi}^3 - \tilde{\phi}^{1/3})}{\partial \tilde{Z}}\right] \Rightarrow \frac{\partial \tilde{\phi}}{\partial \tilde{t}} = \tilde{K}\tilde{\phi}^2 \frac{\partial}{\partial \tilde{Z}}\left[\tilde{Q}(3\tilde{\phi}^3 - \tilde{\phi}^{1/3}/3)\frac{\partial \tilde{\phi}}{\partial \tilde{Z}}\right], \tag{S10}$$

where $\tilde{K} = \frac{\kappa \Pi_0 T_0}{L_0^2}$, $\tilde{Q} = (\frac{kT}{3v_c}\phi_0^3)/\Pi_0$, and we have used $\tilde{\Pi} = \tilde{Q}\left[(\phi/\phi_0)^3 - (\phi/\phi_0)^{1/3}\right]$. Similarly, the boundary condition for $\tilde{\phi}$ can be derived to be: (1) $\tilde{Q}(\tilde{\phi}^3 - \tilde{\phi}^{1/3}) = C/\Pi_0 \equiv \tilde{C}$ at $\tilde{Z} = 0$, and (2) $\tilde{K}\tilde{Q}(3\tilde{\phi}^3 - \tilde{\phi}^{1/3}/3)\frac{\partial \tilde{\phi}}{\partial \tilde{Z}} = \frac{J_{\mathrm{ev}}}{(L_0/T_0)} = 2 \times 10^{-4}$ at $\tilde{Z} = 1$.

Comparing with the experiment. In an experiment where the water reservoir is lowered by $\Delta H = 0.51$ m, corresponding to a change in C from $C = 0$ kPa to $C = 5$ kPa (similar to the procedure in main Fig. 4), the system reaches the new steady state in 10–15 min and the gel shrinks by about 1%. To mimic this experimental procedure, we started $\tilde{\phi}$ at $\tilde{t} = 0$ with the steady state profile $\tilde{\phi}(\tilde{Z})$ at $\tilde{C} = 0$, and simulated Eq. (S10) with $\tilde{C} = 5$ to track the system's dynamics of approaching the new steady state. We define $\tilde{P}(\tilde{t}) = \tilde{Q}(\tilde{\phi}^3(\tilde{t}) - \tilde{\phi}^{1/3}(\tilde{t}))|_{\tilde{Z}=1}$ as the osmotic pressure difference across the gel surface, and we fit $\tilde{P}(\tilde{t})$ with $\tilde{P}(\tilde{t}) = \tilde{P}_0 + \Delta\tilde{P}(1 - e^{-\tilde{t}/\tilde{\tau}_{\mathrm{ss}}})$ to obtain the relaxation time $\tilde{\tau}_{\mathrm{ss}}$. In addition, we also compute the steady-state gel thickness at $\tilde{C} = 0$ and $\tilde{C} = 5$ using the expression $L_{\infty,\tilde{C}} = L_0 \int_0^1 \tilde{\phi}^{-1}d\tilde{Z}$. We use the relative gel thickness $\alpha = L_{\infty,5}/L_{\infty,0}$ and the relaxation time $\tau_{\mathrm{ss}} = \tilde{\tau}_{\mathrm{ss}}T_0$ to compare the model to the experiment, yielding $\tilde{K} = 0.7$–2.1×10^{-4} and $\tilde{Q} \approx 182$ (Supplementary Fig. 6b and c). Finally, we use the fitted values of $\tilde{K}$ and $\tilde{Q}$ to estimate the steady state $\tilde{P}$ (or P) for varying values of $\tilde{C}$ (or C), corresponding to different values of ΔH in experiments. Our analysis leads to a nonzero $P_0 \approx 2$ kPa, which is consistent with the value of P_0 obtained from fitting the height profile. We also verify that P varies (approximately) linearly with the height difference ΔH between the water level of the reservoir and the gel surface (Supplementary Fig. 6d).

F. Calculating the capillary force exerted on cells

We next consider the force exerted by the water meniscus on each cell, which we term the "capillary force". We consider a water meniscus that is in mechanical equilibrium with the cells and has a height profile of $h^*(x,y)$. For an infinitesimal surface element of the water meniscus that is in contact with a cell, we assume that the per-area force exerted on the cell by this element is $\boldsymbol{f}(x,y) \equiv -f(x,y)\boldsymbol{N}(x,y)$, where f denotes the magnitude of the force and $\boldsymbol{N} = (-\frac{\partial_x h^*}{\sqrt{1+|\nabla h^*|^2}}, -\frac{\partial_y h^*}{\sqrt{1+|\nabla h^*|^2}}, \frac{1}{\sqrt{1+|\nabla h^*|^2}})$ defines the local normal.

To compute the magnitude $f(x,y)$, we consider a perturbation to the water meniscus $\boldsymbol{u} = \delta(x,y)\boldsymbol{N}(x,y)$ and compute the free-energy change $\Delta\mathcal{F}$ associated with this perturbation. First, according to the virtual work principle, $\Delta\mathcal{F} = -\int \boldsymbol{f} \cdot \boldsymbol{u} \mathrm{d}S = \int f(x,y)\delta(x,y)\mathrm{d}S = \int f(x,y)\delta(x,y)\sqrt{1+|\nabla h^*|^2}\mathrm{d}x\mathrm{d}y$, where $\mathrm{d}S$ denotes the infinitesimal area in 3D and the integral runs over the surface where the water meniscus is in contact with the upper surface of the cell. Additionally, the free-energy change can also be calculated directly using the free energy functional $\mathcal{F}[h] = \gamma \int \sqrt{1+|\nabla h|^2}\mathrm{d}x\mathrm{d}y + P\int h\mathrm{d}x\mathrm{d}y$. The perturbed water height profile is given by $h^{\mathrm{p}}(x,y) = h^*(x - u_x, y - u_y) + u_z \approx h^*(x,y) - u_x(\partial_x h^*) - u_y(\partial_y h^*) + u_z = h^*(x,y) + \delta(x,y)\sqrt{1+|\nabla h^*|^2}$. Thus, $\Delta\mathcal{F} = \mathcal{F}[h^{\mathrm{p}}] - \mathcal{F}[h^*] \approx \int \frac{\delta\mathcal{F}}{\delta h}(h^{\mathrm{p}} - h^*)\mathrm{d}x\mathrm{d}y = \int \frac{\delta\mathcal{F}}{\delta h}\delta(x,y)\sqrt{1+|\nabla h^*|^2}\mathrm{d}x\mathrm{d}y$. Comparing the two expression for $\Delta\mathcal{F}$, we obtain

$$f(x,y) = \frac{\delta\mathcal{F}}{\delta h} = -\gamma\nabla \cdot \left(\frac{\nabla h^*}{\sqrt{1+|\nabla h^*|^2}}\right) + P. \tag{S11}$$

Introducing Eq. (S11) into the expression for $\boldsymbol{f}$, we obtain the capillary force acting on the cell per 3D surface area. If we define $\boldsymbol{f}_{\mathrm{cap}}$ to be the force per projected area $\mathrm{d}x\mathrm{d}y$, i.e., $\boldsymbol{f}_{\mathrm{cap}}\mathrm{d}x\mathrm{d}y \equiv \boldsymbol{f}\mathrm{d}S = \boldsymbol{f}\sqrt{1+|\nabla h^*|^2}\mathrm{d}x\mathrm{d}y$, then $\boldsymbol{f}_{\mathrm{cap}}$ is given by

$$\boldsymbol{f}_{\mathrm{cap}} = \sqrt{1+|\nabla h^*|^2}\boldsymbol{f} = \left[-\gamma\nabla \cdot \left(\frac{\nabla h^*}{\sqrt{1+|\nabla h^*|^2}}\right) + P\right](\partial_x h^*, \partial_y h^*, -1). \tag{S12}$$

The z component of Eq. (S12) indicates that the water meniscus exerts a downward force on the cells. However, in the present work, we ignore the z component, assuming this is balanced by an upward force from the substrate, and focus on the horizontal capillary force and torque acting on each cell.

Thus, the total horizontal capillary force and torque exerted on cell i can be calculated as

$$\boldsymbol{F}_{\mathrm{cap}}^{(i)} = \int_{\Omega^{(i)}} \left[-\gamma\nabla \cdot \left(\frac{\nabla h^*}{\sqrt{1+|\nabla h^*|^2}}\right) + P\right]\nabla h^*\mathrm{d}x\mathrm{d}y, \tag{S13}$$

$$T_{\mathrm{cap}}^{(i)} = \int_{\Omega^{(i)}} \left[-\gamma\nabla \cdot \left(\frac{\nabla h^*}{\sqrt{1+|\nabla h^*|^2}}\right) + P\right][(x - x_{\mathrm{c}})\partial_y h^* - (y - y_{\mathrm{c}})\partial_x h^*]\mathrm{d}x\mathrm{d}y, \tag{S14}$$

where $\Omega^{(i)}$ denotes the region where the water meniscus is in contact with cell i, i.e., $\Omega^{(i)} = \{(x,y)|h^*(x,y) > R_{\mathrm{c}}$ and $h^*(x,y) = z_{\mathrm{upper}}^{(i)}(x,y)\}$ (see Eq. (S2)). In the simulations, we performed numerical integrals of Eqs. (S13) and (S14) to compute the capillary force and torque. As a cross-check, we also perturbed the position and orientation of the cell and calculated the force and torque from the change in the free energy of water, i.e.,

$$F_{\mathrm{cap},x}^{(i)} = \frac{\mathcal{F}[h^*] - \mathcal{F}[h^*|_{x_{\mathrm{c}}^{(i)} \to x_{\mathrm{c}}^{(i)} + \Delta x^{(i)}}]}{\Delta x^{(i)}}, \quad F_{\mathrm{cap},y}^{(i)} = \frac{\mathcal{F}[h^*] - \mathcal{F}[h^*|_{y_{\mathrm{c}}^{(i)} \to y_{\mathrm{c}}^{(i)} + \Delta y^{(i)}}]}{\Delta y^{(i)}}, \quad T_{\mathrm{cap}}^{(i)} = \frac{\mathcal{F}[h^*] - \mathcal{F}[h^*|_{\theta^{(i)} \to \theta^{(i)} + \Delta\theta^{(i)}}]}{\Delta\theta^{(i)}}, \tag{S15}$$

where $h^*|_{\mathcal{C}}$ denotes the new equilibrium height profile upon a change $\mathcal{C}$ in the position of the cell. The two methods yield the same computational results (Supplementary Fig. 7). For the rest of our simulations, we used the first integral-based method to calculate the capillary force and torque because it is more time efficient than the second method.

G. Additional notes on the magnitude and range of capillary force

To gain more insights into the magnitude and range of capillary forces, we consider a group of N cells and calculate its per-cell capillary attraction $\langle F_{\text{cap}} \rangle_N$ by an infinitely large colony located a distance Δ away (see Supplementary Fig. 8). For a fixed number N of cells, we find that maximum $\langle F_{\text{cap}} \rangle_N$ occurs at $\Delta = 0$, and that the magnitude of capillary attraction increases with the osmotic pressure of the gel P while the range of attraction decreases with P as $\sim P^{-1/2}$. This scaling relation can be understood by considering the simple case of a water meniscus around a 2D circle (Sec. I B). Equation (S5) shows that the half width of the water meniscus is $d_{\text{wm}} = (R_{\text{c}} + l_\gamma) \sin \theta^*$ where $\theta^* = \arccos \frac{l_\gamma - R_{\text{c}}}{l_\gamma + R_{\text{c}}}$. When $l_\gamma \gg R_{\text{c}}$, we derive that $d_{\text{wm}} \approx l_\gamma \theta^* \approx 2\sqrt{R_{\text{c}} l_\gamma}$, and thus $d_{\text{wm}} \sim P^{-1/2}$ since $l_\gamma = \gamma/P$. Next, we set $\Delta = 0$ and look at how the maximum capillary attraction $\langle F_{\text{cap}} \rangle_N$ varies with the number of cells N, finding that it decreases with N as $\langle F_{\text{cap}} \rangle_N = \langle F_{\text{cap}} \rangle_\infty + \frac{\langle F_{\text{cap}} \rangle_1 - \langle F_{\text{cap}} \rangle_\infty}{N}$. In particular, we find that $\langle F_{\text{cap}} \rangle_\infty$ scales linearly with P. To gain some physical intuition for this relationship, we again turn to the simple case of the water meniscus around a circular cell. We consider a chain of closely packed circular cells extending from $x = -\infty$ to $x = 0$. The water height profile follows Eq. (S5) for $x \geqslant 0$ and $h \approx 2R_{\text{c}}$ for $x < 0$ where cells are filled up. The capillary force acting on the cell at $x = 0$ can thus be calculated from Eq. (S13) to be $F_{\text{cap}} = 2\gamma R_{\text{c}}(1 + \frac{R_{\text{c}}}{l_\gamma})\frac{R_{\text{c}}}{R_{\text{c}} + l_\gamma} \approx 2\gamma R_{\text{c}}^2/l_\gamma = 2P R_{\text{c}}^2$ when $l_\gamma \gg R_{\text{c}}$.

II. AGENT-BASED MODEL

A. Equations of motion

In this subsection, we summarize the different forces that we consider in the agent-based model, and list the equations of motion for each cell.

Cell shape. As noted in Sec. I A, we consider spherocylindrical cells oriented horizontally. For simplicity, we only consider the in-plane motion of the cells. Thus, unless otherwise specified, the vectors discussed below will be two dimensional in the $x - y$ plane. We denote by $\boldsymbol{r}_{\mathrm{c}} = (x_{\mathrm{c}}^{(i)}, y_{\mathrm{c}}^{(i)})$ the center-of-mass position of cell i, and we define $\boldsymbol{n}_{\parallel}^{(i)} = (\cos\theta^{(i)}, \sin\theta^{(i)})$ and $\boldsymbol{n}_{\perp}^{(i)} = (-\sin\theta^{(i)}, \cos\theta^{(i)})$ to denote, respectively, the unit vectors parallel to and perpendicular to the long axis of cell i.

Steric interaction between cells. For each pair of cells i and j, we compute the smallest distance d_{ij} between the centerlines of the two cell cylinders. Denote by $\boldsymbol{r}^{(i,j)} = \boldsymbol{r}_{\mathrm{c}}^{(i)} + s^{(i,j)}\boldsymbol{n}_{\parallel}^{(i)}$ the coordinates of the point on cell i's centerline that is the closest to cell j's centerline. Thus, the distance d_{ij} can be calculated by $d_{ij} = |\boldsymbol{r}^{(i,j)} - \boldsymbol{r}^{(j,i)}|$. The magnitude of the repulsive force between cells i and j is modeled as $F_{ij} = \frac{B}{2}\left[1 - \tanh(\frac{d_{ij} - 2R_{\mathrm{c}}}{d_B})\right]$, where B and d_B denote the magnitude and range of steric interaction, respectively. $F_{ij} \approx 0$ if $d_{ij} \gg 2R_{\mathrm{c}} + d_B$, i.e., when the cells are not in contact with each other. The total steric force and torque acting on cell i are given by

$$\boldsymbol{F}_{\mathrm{steric}}^{(i)} = \sum_{j \neq i} F_{ij} \frac{\boldsymbol{r}^{(i,j)} - \boldsymbol{r}^{(j,i)}}{d_{ij}}, \tag{S16}$$

$$T_{\mathrm{steric}}^{(i)} = \sum_{j \neq i} s^{(i,j)} F_{ij} \boldsymbol{n}_{\parallel}^{(i)} \times \frac{\boldsymbol{r}^{(i,j)} - \boldsymbol{r}^{(j,i)}}{d_{ij}}, \tag{S17}$$

where the cross product of two-dimensional vectors $\boldsymbol{f} = (f_x, f_y)$ and $\boldsymbol{g} = (g_x, g_y)$ is defined by $\boldsymbol{f} \times \boldsymbol{g} = f_x g_y - f_y g_x$.

Friction between cells and substrate. Cell motion is opposed by friction from the substrate surface. We model the friction as a viscous drag that opposes the motion of each infinitesimal segment of the cell cylinder's centerline. Denote by $\boldsymbol{v}_{\mathrm{c}}^{(i)} = d\boldsymbol{r}_{\mathrm{c}}/dt$ and $\omega_{\mathrm{c}}^{(i)} = d\theta^{(i)}/dt$ the translational and rotational velocity of cell i. The velocity of the point $\boldsymbol{r}^{(i)}(s) = \boldsymbol{r}_{\mathrm{c}}^{(i)} + s\boldsymbol{n}_{\parallel}^{(i)}$ on cell i is $\boldsymbol{v}^{(i)}(s) = \boldsymbol{v}_{\mathrm{c}}^{(i)} + \omega_{\mathrm{c}}^{(i)} s\boldsymbol{n}_{\perp}^{(i)}$. Assuming anisotropic friction coefficients in the directions parallel to and perpendicular to the long axis of cell i, we derive the infinitesimal frictional force $d\boldsymbol{F}_{\mathrm{fric}}$ acting on $\boldsymbol{r}^{(i)}(s)$ to be given by $d\boldsymbol{F}_{\mathrm{fric}}(s) = -\boldsymbol{\xi} \cdot \boldsymbol{v}^{(i)}(s)ds = -(\xi_{\parallel}\boldsymbol{n}_{\parallel}^{(i)}\boldsymbol{n}_{\parallel}^{(i)} + \xi_{\perp}\boldsymbol{n}_{\perp}^{(i)}\boldsymbol{n}_{\perp}^{(i)}) \cdot (\boldsymbol{v}_{\mathrm{c}}^{(i)} + \omega_{\mathrm{c}}^{(i)} s\boldsymbol{n}_{\perp}^{(i)})ds = -\left[\xi_{\parallel}(\boldsymbol{v}_{\mathrm{c}}^{(i)} \cdot \boldsymbol{n}_{\parallel}^{(i)})\boldsymbol{n}_{\parallel}^{(i)} + \xi_{\perp}(\boldsymbol{v}_{\mathrm{c}}^{(i)} \cdot \boldsymbol{n}_{\perp}^{(i)})\boldsymbol{n}_{\perp}^{(i)} + \xi_{\perp}\omega_{\mathrm{c}}^{(i)} s\boldsymbol{n}_{\perp}^{(i)}\right]ds$. Thus, the total frictional force and torque acting on cell i are given by

$$\boldsymbol{F}_{\mathrm{fric}}^{(i)} = \int_{-l_{\mathrm{c}}/2}^{l_{\mathrm{c}}/2} d\boldsymbol{F}_{\mathrm{fric}} = -\xi_{\parallel} l_{\mathrm{c}}(\boldsymbol{v}_{\mathrm{c}}^{(i)} \cdot \boldsymbol{n}_{\parallel}^{(i)})\boldsymbol{n}_{\parallel}^{(i)} - \xi_{\perp} l_{\mathrm{c}}(\boldsymbol{v}_{\mathrm{c}}^{(i)} \cdot \boldsymbol{n}_{\perp}^{(i)})\boldsymbol{n}_{\perp}^{(i)}, \tag{S18}$$

$$T_{\mathrm{fric}}^{(i)} = \int_{-l_{\mathrm{c}}/2}^{l_{\mathrm{c}}/2} s\boldsymbol{n}_{\parallel}^{(i)} \times d\boldsymbol{F}_{\mathrm{fric}} = -\int_{-l_{\mathrm{c}}/2}^{l_{\mathrm{c}}/2} \xi_{\perp}\omega_{\mathrm{c}}^{(i)} s^2 ds = -\frac{1}{12}\xi_{\perp} l_{\mathrm{c}}^3 \omega_{\mathrm{c}}^{(i)}. \tag{S19}$$

Self-propelling force of each cell. The self-propelling force is assumed to be along the long axis of each cell, i.e.,

$$\boldsymbol{F}_{\mathrm{prop}}^{(i)} = F_{\mathrm{p0}}^{(i)} \chi^{(i)} \boldsymbol{n}_{\parallel}^{(i)}, \tag{S20}$$

where F_{p0} denotes the magnitude of self propulsion and $\chi = \pm 1$ follows a dichotomous Poisson process with rate τ_{r}^{-1}. The cells do not generate self-propelling torques, i.e., $T_{\mathrm{prop}} = 0$.

Capillary force acting on the cells. The capillary force and torque exerted by the water meniscus are discussed in Sec. I F and are given by Eqs. (S13) and (S14).

Equations of motion. Taken together, the balance of the above forces determine the equations of motion for the modeled cells,

$$\frac{\mathrm{d}\boldsymbol{r}_{\mathrm{c}}^{(i)}}{\mathrm{d}t} = \Big(\frac{\boldsymbol{n}_{\parallel}^{(i)}\boldsymbol{n}_{\parallel}^{(i)}}{\xi_{\parallel}l_{\mathrm{c}}} + \frac{\boldsymbol{n}_{\perp}^{(i)}\boldsymbol{n}_{\perp}^{(i)}}{\xi_{\perp}l_{\mathrm{c}}}\Big) \cdot (\boldsymbol{F}_{\mathrm{steric}}^{(i)} + \boldsymbol{F}_{\mathrm{prop}}^{(i)} + \boldsymbol{F}_{\mathrm{cap}}^{(i)}), \tag{S21}$$

$$\frac{\mathrm{d}\theta^{(i)}}{\mathrm{d}t} = \frac{12}{\xi_{\perp}l_{\mathrm{c}}^3}(T_{\mathrm{steric}}^{(i)} + T_{\mathrm{cap}}^{(i)}). \tag{S22}$$

Note that the dichotomous noise for reversal is the only source of noise in the system.

Growth of non-motile cells. To study how capillary forces influences the development of growing colonies, we follow previous work [4] and model bacterial growth as elongation of modeled spherocylinderical cells, whose projected area is given by $A_{\mathrm{c}} = \pi R_{\mathrm{c}}^2 + 2R_{\mathrm{c}}l_{\mathrm{c}}$. Denoting by α_{g} the areal growth rate, and introducing the expression for A_{c} into $\mathrm{d}A_{\mathrm{c}}/\mathrm{d}t = \alpha_{\mathrm{g}}S$, we obtain

$$\frac{\mathrm{d}l_{\mathrm{c}}}{\mathrm{d}t} = \alpha_{\mathrm{g}}(l_{\mathrm{c}} + \frac{\pi}{2}R_{\mathrm{c}}). \tag{S23}$$

Each modeled cell has an initial length of l_0, elongates according to Eq. (S23) until it reaches a final length of $l_{\mathrm{m}} = 2l_0 + 2R_{\mathrm{c}}$, and divides into two daughter cells. Cell division is modeled by replacing the mother cell at $\boldsymbol{r}_{\mathrm{c}}^{(m)}$ with two spherocylinders of length l_0 at positions $\boldsymbol{r}_{\mathrm{c}}^{(m)} \pm (l_0/2 + R_{\mathrm{c}})\boldsymbol{n}_{\parallel}^{(m)}$ with the same orientation $\boldsymbol{n}_{\parallel}^{(m)}$ as the mother cell.

B.   Simulations and parameters

The cell positions and the equilibrium water height profile are updated iteratively. The equations of cell motion are simulated using a forward Euler method with an adaptive time increment such that for every time step the maximum displacement of any cell segment does not exceed $0.1R_{\mathrm{c}}$. The simulation parameters are chosen or determined as follows:

Capillary force: As discussed in Sec. I D, the surface tension γ is measured to be $\gamma = 66$ mN/m. The osmotic pressure P is varied between $P = 2.8$ kPa and $P = 8.8$ kPa.

Steric interaction: To capture the steric interactions between cells, we choose $B = 8\gamma R_{\mathrm{c}}$, which is much larger than typical capillary force and self-propelling force, to ensure that the modeled cells cannot penetrate into each other. We set $d_B = 0.1R_{\mathrm{c}}$ to ensure that the interaction is very short-ranged.

Friction coefficients: We estimated the coefficients of the anisotropic friction between the cells and the substrate from an experiment in which two nearby non-motile cells/cell clusters are pulled together by capillary attraction (see main Fig. 2). We fit the modeled merger dynamics to the experiment. The fitting error was calculated by comparing the center-to-center distance $d_{12} = |\boldsymbol{r}_{\mathrm{c}}^{(1)} - \boldsymbol{r}_{\mathrm{c}}^{(2)}|$ and the angle $\phi_{12} = \theta^{(1)} - \theta^{(2)}$ of the two cell clusters between the model and the experiment. For each time point t, the fitting error is given by the root-mean-square-deviation $e_t = \big[(\frac{d_{12,t}^{\mathrm{exp}} - d_{12,t}^{\mathrm{sim}}}{L})^2 + (\frac{\phi_{12,t}^{\mathrm{exp}} - \phi_{12,t}^{\mathrm{sim}}}{\Phi})^2\big]^{1/2}$ where the superscripts exp and sim denote the quantities in experiment and simulation, respectively, and we set $L = 2.5$ µm and $\Phi = 1$ rad to be the characteristic scales for length and angle. The total fitting error was calculated as $\sum_t e_t$. Finally, we varied the friction coefficient $\xi_{\perp}$ and friction anisotropy $\frac{\xi_{\perp} - \xi_{\parallel}}{\xi_{\perp} + \xi_{\parallel}}$ to determine the optimal fitting parameters that minimize $\sum_t e_t$ (Supplementary Fig. 9), which yielded $\xi_{\parallel} = 80$ Pa · min and $\frac{\xi_{\perp} - \xi_{\parallel}}{\xi_{\perp} + \xi_{\parallel}} = 0.85$.

Phase diagrams: To explore the collective dynamics in the model, we varied three dimensionless model parameters: a normalized reversal period $\tilde{\tau} = \tau_{\mathrm{r}}/\tau_{\mathrm{c}}$ where $\tau_{\mathrm{c}} = \xi_{\parallel}l_{\mathrm{c}}^2/|\boldsymbol{F}_{\mathrm{prop}}|$ is the time it takes for a non-reversing free-running cell to travel its cell body length, a normalized self-propulsion strength $\tilde{F} = |\boldsymbol{F}_{\mathrm{prop}}|/PR_{\mathrm{c}}l_{\mathrm{c}}$ where $PR_{\mathrm{c}}l_{\mathrm{c}}$ is the scale of the longitudinal capillary force acting on isolated cells, and a normalized capillary length $\tilde{\ell} = l_{\gamma}/R_{\mathrm{c}} = \gamma/(PR_{\mathrm{c}})$.

Cell growth: The areal growth rate α_{g} of bacteria was determined by tracking the total area covered by the cells over time (Supplementary Fig. 13).

C. A toy model for the splitting of cell groups

In the presence of a water meniscus, a group of cells can split into two subgroups if they propel sufficiently strong in opposite directions to overcome the opposing capillary force. We wondered how this capillary attraction depends on the partition of the subgroups, which describes the number of cells in each subgroup. To explore this question, we consider two subgroups with N_1 and N_2 cells that are aligned in the same direction but propel in opposite directions (Supplementary Fig. 10). Since the equilibrium water meniscus experiences zero net force exerted by the cells, the total capillary force exerted on the two groups must have the same magnitude $F_{\mathrm{cap}}^{(\mathrm{g})}$ but opposite directions. Assuming that each cell has the same self-propelling force F_{p0} and the cells in the same subgroup move at the same speeds v_1 and v_2, we derive the force-balance conditions for the two groups

$$F_{\mathrm{cap}}^{(\mathrm{g})} - N_1 F_{\mathrm{p0}} = N_1 \xi_{\parallel} v_1, \tag{S24a}$$

$$N_2 F_{\mathrm{p0}} - F_{\mathrm{cap}}^{(\mathrm{g})} = N_2 \xi_{\parallel} v_2. \tag{S24b}$$

Splitting occurs when $v_2 > v_1$, i.e., $F_{\mathrm{p0}} > F_{\mathrm{cap}}^{(\mathrm{g})}(1/N_1 + 1/N_2)/2$. As shown in Supplementary Fig. 10, for a fixed number of total cells $N_1 + N_2$, the effective per-cell capillary attraction $2F_{\mathrm{cap}}^{(\mathrm{g})}(1/N_1 + 1/N_2)$ is the smallest when $N_1 \approx N_2$, suggesting that symmetric splitting of cell groups is favored. In addition, for a fixed number N_2, the effective per-cell capillary attraction on cells in group 1 decreases with increasing N_1, as noted in Sec. I G.

D. Quantifying the elongation and splitting of cell groups

As illustrated in Supplementary Fig. 10 and discussed in Sec. II C, cells within a group can propel in the opposite directions, which creates a force dipole that tends to either elongate the cell group when the capillary force is strong or split apart the group when the capillary force is small.

To quantify the elongation of the cell group, we first consider a group of $N = 10$ cells with relatively small values of $\tilde{F} < 1$. The contour of the cell group is determined by the contour line of the water height profile $h_{\mathrm{w}}^* = R_{\mathrm{c}}$. We follow previous literature [5] to analyze the deformation modes of the contour. Briefly, we first compute the center of all points on the contour, defined as their average position weighted by the local curvilinear distance, and calculate their polar coordinates (r_i, θ_i) with respect to the center. Subsequently, we decompose the contour into Fourier modes, i.e., $r(\theta) = R\left[1 + \sum_{m \geqslant 1} a_n \cos(m\theta) + \sum_{m \geqslant 1} b_m \sin(m\theta)\right]$ where $R = \frac{1}{2\pi} \sum_i \frac{R_i + R_{i-1}}{2} \times (\theta_i - \theta_{i-1})$ denotes the zero mode component, and compute $c_m = \sqrt{a_m^2 + b_m^2}$ to characterize the mode-m deformation of the contour (Supplementary Fig. 11). Mode $m = 1$ corresponds to the translation mode of the contour and thus $a_1 = b_1 \equiv 0$ by our definition. In our simulations, the dominant deformation mode of cell groups is mode $m = 2$, which corresponds to the elliptical elongation of the contour (Supplementary Fig. 11a). To quantify the shape fluctuation of the cell groups, we compute the fluctuation of mode $m = 2$, i.e., $\langle c_2(t)^2 \rangle_t - \langle c_2(t) \rangle_t^2$, finding that it increases with increasing reversal time $\tilde{\tau}$ of the cells and increasing self-propelling force $\tilde{F}$ (Supplementary Fig. 11b).

To quantify the splitting of the cell group, we consider a group of $N = 10$ cells with $\tilde{F} \sim 1$. We use the criterion $h_{\mathrm{w}}^* \geqslant h_{\mathrm{contour}}$ to determine the cell regions $\mathcal{G}_{\mathrm{cell}}$ and track the number n_{cc} of connected components in $\mathcal{G}_{\mathrm{cell}}$. The splitting events are identified by changes in n_{cc} from $n_{\mathrm{cc}} = 1$ to $n_{\mathrm{cc}} > 1$. The splitting rate of cell groups can be determined from the inverse of the average waiting times in the $n_{\mathrm{cc}} = 1$ state. As shown in Supplementary Fig. 11c, our simulations show that the splitting rate also increases with increasing reversal time $\tilde{\tau}$ of the cells and increasing self-propelling force $\tilde{F}$.

E.   Capillary forces facilitate merger of growing colonies

Since water menisci should generally exist around bacterial cells grown on hydrated substrates, we wondered how capillary forces might affect the development of growing but non-motile colonies. To explore this question, we imaged *E. coli* colonies, which are known to be poorly motile on hard agar [6], and found that two adjacent colonies elongate and expand preferentially toward each other (Supplementary Fig. 14). We hypothesized that the elongation of growing colonies is due to the asymmetric capillary force acting on them. Specifically, the water meniscus between the two colonies has a smaller slope than that in other regions, and thus the in-plane capillary force is smaller in the middle, allowing the colonies to expand more easily toward each other. To test this hypothesis, we also simulated colony growth in the presence of a water meniscus, and our simulation results are in quantitative agreement with the experiment. To further verify our hypothesis, we simulated colony growth in the absence of capillary forces, finding that in this case the adjacent colonies stay roughly circular and do not expand toward each other (Supplementary Fig. 14).

III.   CLASSIFYING AND QUANTIFYING DIFFERENT PHASES OF COLLECTIVE CELL DYNAMICS

A.   Phase classification

Simulations of the agent-based model described in Sec. II yielded a range of collective dynamics when the normalized parameters of water and cell motility are varied. Qualitatively, the simulated dynamics can be categorized into four phases:

1. a "gas" phase in which cells are spatially distributed rather than packed together by the capillary force,

2. a "polar clusters" phase in which cells are clustered into small groups that move persistently over long distances and constantly undergo merging and splitting events,

3. a "streams" phase in which cells self-organize into a quasi-one-dimensional structure, and

4. a "droplets" phase in which cells are packed together and trapped in cell groups that do not move persistently.

To quantify the collective dynamics of N cells, we compute the equal-time correlation function of cell density ρ

$$C_\rho(\Delta r_\parallel, \Delta r_\perp, t) = \frac{1}{N} \sum_{i=1}^{N} \rho(\boldsymbol{r}_c^{(i)}) \rho(\boldsymbol{r}_c^{(i)} + \Delta r_\parallel \boldsymbol{n}_\parallel^{(i)} + \Delta r_\perp \boldsymbol{n}_\perp^{(i)}) \tag{S25}$$

and normalize C_ρ by its maximum $C_\rho(0, 0, t)$ to obtain $\tilde{C}_\rho(\Delta r_\parallel, \Delta r_\perp, t)$. The average and standard deviation of $\tilde{C}_\rho(\Delta r_\parallel, \Delta r_\perp, t)$ with respect to time, denoted by $\mathcal{A} = \langle \tilde{C}_\rho \rangle_t$ and $\mathcal{B} = \sqrt{\langle \tilde{C}_\rho^2 \rangle_t - \langle \tilde{C}_\rho \rangle_t^2}$, characterize the morphological and dynamical features of $\tilde{C}_\rho$, respectively (main Fig. 3 and Supplementary Fig. 15a).

To classify the simulation results with different parameters in an unbiased way, we develop a measure of distance between any pair of simulated dynamics, and use a standard clustering method to group different simulations into four clusters (Supplementary Fig. 15b). We follow previous computer-vision literature [7, 8] and measure the pairwise "distance" between the simulated dynamics using the distances of $\mathcal{A}$ and $\mathcal{B}$ in the Sobolev space. For arbitrary matrices $\mathcal{X}_i$ and $\mathcal{X}_j$, the Sobolev norm of order $1/2$ is given by

$$s_{ij}(\mathcal{X}) \equiv \left[\frac{1}{n(\mathbf{K})} \sum_{\boldsymbol{k} \in \mathbf{K}} (1 + |\boldsymbol{k}|)^{1/2} |\mathfrak{x}_i - \mathfrak{x}_j|^2 \right]^{1/2}, \tag{S26}$$

where $\mathfrak{x}_i$ and $\mathfrak{x}_j$ denote, respectively, the discrete Fourier transforms of $\mathcal{X}_i$ and $\mathcal{X}_j$, and $n(\mathbf{K})$ denotes the number of discrete frequency vectors in the frequency domain $\mathbf{K}$. We define the distance D_{ij} between simulations i and j as $D_{ij} = \left[s_{ij}^2(\mathcal{A}) + (5s_{ij}(\mathcal{B}))^2 \right]^{1/2}$ where the factor 5 is chosen such that the terms containing $s_{ij}(\mathcal{A})$ and $s_{ij}^2(\mathcal{B})$ have similar magnitudes. The clustering is performed using the WPGMA (Weighted Pair Group Method with Arithmetic Mean) method and implemented in MATLAB. Finally, we handpick one representative simulation for each phase of collective dynamics to verify and annotate the clustering results.

B.   Phase quantification

As shown in main Fig. 3, for a fixed self-propelling force, the collective cell dynamics transition from the gas phase when the cell-cell capillary attraction is low, to the streams and polar clusters phases when the cell-cell capillary attraction is intermediate, and to the droplets phase when the cell-cell capillary attraction is high. The effect of the capillary attraction on the collective cell dynamics manifests itself in a trade-off between cell mobility and cell-cell adjacency. To quantitatively describe this trade-off, we compute the instantaneous cell speeds, normalized by the average free-running speed $|\mathbf{F}_{\text{prop}}|/(\xi_\parallel l_c)$, and the cell-cell distances[2], normalized by the average cell body length l_c, for the different phases. Our simulations show that both the cell speed and the cell-cell distance decreases with

---

[2] Following previous work [4], we define cell-cell distance as the smallest distance between the centerlines of the cell cylinders.

increasing magnitude of capillary attraction (Supplementary Fig. 17a,b). The streams and polar clusters phases show similar distribution of cell-cell distances, but their collective dynamics are different – cells moving in quasi-1D streams frequently slip past each other while cells moving in small polar clusters stay together for a relatively long time. To capture this difference, we compute the cell-cell adjacency matrix $A_{ij}(t)$, where A_{ij} is 1 if the distance between cells i and j are smaller than $4R_c$ and 0 otherwise. The average autocorrelation function of this matrix $C_A(\Delta t) \equiv \sum_{i \neq j} \langle A_{ij}(t) A_{ij}(t+\Delta t) \rangle_t / \sum_{i \neq j} \langle A_{ij}(t) A_{ij}(t) \rangle_t$ describes the probability that a pair of adjacent cells remain adjacent after time Δt. The life time of adjacent cell pairs is the longest in the droplets phase, second longest in the polar clusters phase, third longest in the streams phase, and the shortest in the gas phase (Supplementary Fig. 17c).

In the droplets phase where the self-propelling force of each cell is much smaller than the capillary force, entire cell groups undergo collective translational motion. Intuitively, over time scales longer than the mean reversal time τ_r of the cell, such group motion is diffusive, and the group diffusivity D_g should decrease with increasing number N of cells in the group because the overall propelling force of the group $\sum_{i=1}^{N} \mathbf{F}_{\text{prop}}^{(i)}$ scales with $\sqrt{N}$ while the overall friction of the group scales (approximately) with N. To verify our intuition, we simulate the dynamics of cell groups with varying cell numbers N in the droplet phase. We use the same thresholding method as in Sec. II D to determine the cell regions and track the center-of-mass trajectories $\mathbf{r}_g(t)$ of the cell groups. The mean squared displacement (MSD) is subsequently computed as $\text{MSD}(\Delta t) = \langle |\mathbf{r}_g(t+\Delta t) - \mathbf{r}_g(t)|^2 \rangle_t$. We obtain the group diffusivity D_g by fitting a linear function $D_g \Delta t$ to the MSD with a lag time $\Delta t > 3\tau_r$. Our analyses show that, when friction is isotropic, i.e., $\xi_\perp = \xi_\parallel$, both the mean squared speed $\langle V_g \rangle^2$ and the diffusivity D_g of the group scale inversely with N and they are related by $D_g = \langle V_g \rangle^2 \tau_r$ (Supplementary Fig. 18). For anisotropic friction $\xi_\perp > \xi_\parallel$, the relationship $D_g = \langle V_g \rangle^2 \tau_r$ remains true. In this case, D_g and $\langle V_g \rangle^2$ decrease with N more rapidly than N^{-1} (Supplementary Fig. 18) because cells can align in different orientations in groups with large N and get pushed sideways, thus increasing the overall friction and decreasing the group diffusivity and mobility compared to the case of isotropic friction.

IV. CONTINUUM MODEL

A. Phase-field formulation

To describe the large-scale organization of cell groups in the presence of water, we study a minimal phase-field model. Briefly, we consider groups of cells that self-propel in M pairs of opposite directions $\boldsymbol{n}_\pm^{(i)}$ with densities $\rho_\pm^{(i)}$ and velocities $\boldsymbol{v}_\pm^{(i)}$. Since we are interested in the large-scale morphology of *M. xanthus* cell groups, we introduce a phase field ϕ that tracks the region of cells. $\phi = 1$ denotes regions filled with cells and $\phi = 0$ denotes cell-free voids. Thus, the mass conservation equation can be expressed as

$$\partial_t(\phi\rho_s^{(i)}) + \nabla \cdot (\phi\rho_s^{(i)}\boldsymbol{v}_s^{(i)}) = \phi\tau_\rho^{-1} \sum_{r=\pm} \sum_{j=0}^{N-1} W_{r,s}^{(j,i)}\rho_r^{(j)} + q_{\text{align}}(\phi, \{\rho_s^{(i)}\}) \qquad \text{for } s = \pm, \text{and } i = 0, ..., M-1. \quad \text{(S27)}$$

Here, τ_ρ denotes the rate of changing propulsion directions and $W_{r,s}^{(j,i)}$ denotes the transition probability matrix. The alignment term q_{align} describes the dynamics of cell alignment due to cell-cell and cell-meniscus interactions. To model the nematic cell-cell alignment, we consider the local nematic tensor order parameter $\mathbf{Q} = \sum_{i,s} \rho_s^{(i)}\mathbf{Q}_s^{(i)} = \sum_{i,s} \rho_s^{(i)}(2\boldsymbol{n}_s^{(i)}\boldsymbol{n}_s^{(i)} - \mathbf{I})$, and an associated free energy,

$$\mathcal{F}_{\text{nem}}[\mathbf{Q}, \phi] = \int \frac{K_1}{2}(2 - \mathbf{Q} : \mathbf{Q})^2 \phi d\boldsymbol{x} + \int \frac{K_2}{2}|\nabla\mathbf{Q}|^2\phi d\boldsymbol{x}, \quad \text{(S28)}$$

where the first term favors local nematic ordering with a coefficient K_1, and the second term penalizes gradient of nematic order parameter with a coefficient K_2. To capture the alignment with the boundary of water meniscus, we consider an additional surface free-energy term with nematic symmetry $\mathcal{F}_{\text{srf}} = K_3 \int \sum_{i,s} \rho_s^{(i)}(\boldsymbol{n}_s^{(i)} \cdot \nabla\phi)^2 d\boldsymbol{x}$. By introducing the expression for $\mathbf{Q}$ into Eq. (S28) and taking the functional derivative with respect to $\rho_s^{(i)}$, we derive that

$$q_{\text{align}} = -\frac{\delta\mathcal{F}_{\text{nem}}}{\delta\rho_s^{(i)}} - \frac{\delta\mathcal{F}_{\text{srf}}}{\delta\rho_s^{(i)}} = 2K_1\phi(2 - \mathbf{Q} : \mathbf{Q})(\mathbf{Q} : \mathbf{Q}_s^{(i)}) + K_2\nabla \cdot (\phi\nabla\mathbf{Q}) : \mathbf{Q}_s^{(i)} - K_3(\boldsymbol{n}^{(i)} \cdot \nabla\phi)^2. \quad \text{(S29)}$$

Using the phase field ϕ, we derive the force-balance condition

$$\phi\xi\boldsymbol{v}_s^{(i)} = \phi s\beta\boldsymbol{n}^{(i)} - \nabla(\phi p) + \boldsymbol{f}_{\text{water}}(\phi), \quad \text{(S30)}$$

where the effect of the water meniscus is described by $\boldsymbol{f}_{\text{water}}$ and is nonzero only at the boundary of cell regions. In our agent-based model, we found that the water meniscus exerts an in-plane pressure on the periphery of a cell group and the pressure is proportional to the curvature of the boundary, similar to the Laplacian pressure due to surface tension (Supplementary Fig. 21). Thus, in the continuum model, we treat the cell population as 2D droplets, and approximate the effect of water menisci by an equivalent liquid surface tension σ, which will be described below. Finally, as described in previous works [9–11], the time evolution of ϕ follows

$$\partial_t\phi + \bar{\boldsymbol{v}} \cdot \nabla\phi = \Gamma\left[\epsilon\nabla^2\phi - \epsilon^{-1}G'(\phi) + \epsilon\kappa_\phi|\nabla\phi|\right], \quad \text{(S31)}$$

where the phase field ϕ is advected by the mean velocity $\bar{\boldsymbol{v}} \equiv \sum_{s=\pm} \sum_{i=0}^{N-1} \rho_s^{(i)}\boldsymbol{v}_s^{(i)} / \sum_{s=\pm} \sum_{i=0}^{N-1} \rho_s^{(i)}$. We choose the standard double-well potential $G(\phi) = 18\phi^2(1 - \phi^2)$ such that the bulk regions have either $\phi = 1$ or $\phi = 0$. Here, Γ denotes the relaxation rate of the phase field, ϵ denotes the width of the boundary layer, and $\kappa \equiv -\nabla \cdot (\frac{\nabla\phi}{|\nabla\phi|})$ denotes the local curvature. Using the phase field ϕ, we express the surface force $\boldsymbol{f}_{\text{water}}$ as $\boldsymbol{f}_{\text{water}} = \sigma(\epsilon\nabla^2\phi - G'(\phi)/\epsilon)\nabla\phi$ [11].

Choice of parameters: We simulate the above equations on a square domain $[-10, 10] \times [-10, 10]$. We set the phase-field parameters $\epsilon = 0.5$ and $\Gamma = 1$. We set the scale of time, force, and density by setting $\xi = 1$, $\beta = 1$, and $\sum_{s=\pm} \sum_{i=0}^{M-1} \rho_s^{(i)} = 1$, respectively. For simplicity, we study the case $M = 3$. We set the propensity of reversing self-propelling directions to be the highest $W_{-s,s}^{(i,i)} = 0.6$, and we set the propensity of changing to other directions

to be the same $W_{r,s}^{(j,i)}|_{j\neq i} = 0.1$. To reproduce the polar clusters, streams, and droplets observed in our agent-based simulations, we set the alignment parameters to be $K_1 = 1$, $K_2 = 0.01$, and $K_3 = 1$. The remaining parameters σ and τ_ρ are varied to explore how the strength of capillary force (relative to the self-propelling force $\beta = 1$) and the cell reversal frequency influence the large-scale dynamics.

B. Numerical scheme

The numerical scheme for solving Eqs. (S27-S31) was described previously in detail [11]. Briefly, we employed an isotropic finite-difference method to compute spatial derivatives on a discretized grid of 256×256 points. Equations (S27) and (S31) were solved using an Euler forward scheme with a fixed time increment $\Delta t = 2 \times 10^{-3}$. To ensure numerical convergence, a small diffusion term were introduced to Eq. (S27).

To solve the force-balance equation Eq. (S30), we separate the velocities $\boldsymbol{v}_s^{(i)}$ into two terms $\boldsymbol{v}_s^{(i)} \equiv \boldsymbol{u}_s^{(i)} + \boldsymbol{w}_s^{(i)}$, where $\boldsymbol{u}_s^{(i)} = s\beta\boldsymbol{n}^{(i)}/\xi$. Introducing this expression into Eq. (S30) and the incompressibility condition, we obtain

$$\phi\xi\boldsymbol{w}_s^{(i)} = -\nabla(\phi p) + \boldsymbol{f}_{\text{water}}(\phi), \tag{S32a}$$

$$\nabla \cdot \Big[\sum_{s=\pm}\sum_{i=0}^{N-1}\rho_s^{(i)}\boldsymbol{w}_s^{(i)}\Big] = -\nabla \cdot \Big[\sum_{s=\pm}\sum_{i=0}^{N-1}\rho_s^{(i)}\boldsymbol{u}_s^{(i)}\Big]. \tag{S32b}$$

Since Eq. (S32a) is independent of s and i, we look for solutions $\boldsymbol{w}_s^{(i)} \equiv \boldsymbol{w}$ ($\forall s, i$). Thus, $\boldsymbol{w}$ follows

$$\phi\xi\boldsymbol{w} = -\nabla(\phi p) + \boldsymbol{f}_{\text{water}}(\phi), \text{ and } \nabla \cdot \boldsymbol{w} = g \equiv -\nabla \cdot \Big[\sum_{s=\pm}\sum_{i=0}^{N-1}\rho_s^{(i)}\boldsymbol{u}_s^{(i)}\Big]. \tag{S33}$$

Note that in deriving Eq. (S33), we have used the normalization $\sum_{s=\pm}\sum_{i=0}^{N-1}\rho_s^{(i)} = 1$.

Equation (S33) was solved iteratively using a semi-implicit Fourier-spectral method. In particular, we formally rewrite Eq. (S33) as

$$\tilde{\xi}\boldsymbol{w} - \tilde{\nu}\nabla^2\boldsymbol{w} = -\nabla(\phi p) + \boldsymbol{f}_{\text{water}} + (\tilde{\xi} - \xi\phi)\boldsymbol{w} - \tilde{\nu}\nabla^2\boldsymbol{w} \equiv -\nabla(\phi p) + \boldsymbol{F}, \tag{S34}$$

where we have introduced two positive constants $\tilde{\xi}$ and $\tilde{\nu}$ to stabilize the numerical scheme, and we have grouped all the non-pressure terms into $\boldsymbol{F}$. Taking the divergence of Eq. (S34) and using $\nabla \cdot \boldsymbol{w} = g$, we obtain $-\nabla^2(\phi p) + \nabla \cdot \boldsymbol{F} = (\tilde{\xi} - \tilde{\nu}\nabla^2)g$. We denote by $[f]_{\boldsymbol{k}} = \int f(\boldsymbol{x})e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}d^2\boldsymbol{x}$ the Fourier transform of an arbitrary function $f(\boldsymbol{x})$. The Fourier transform of the simplified equation leads to $|\boldsymbol{k}|^2[\phi p]_{\boldsymbol{k}} + i\boldsymbol{k} \cdot [\boldsymbol{F}]_{\boldsymbol{k}} = (\tilde{\xi} + \tilde{\nu}|\boldsymbol{k}|^2)[g]_{\boldsymbol{k}}$. Introducing this relation into the Fourier transform of Eq. (S34), we obtain that $(\tilde{\xi} + \tilde{\nu}|\boldsymbol{k}|^2)[\boldsymbol{w}]_{\boldsymbol{k}} = [\boldsymbol{F}]_{\boldsymbol{k}} - \frac{\boldsymbol{k}\boldsymbol{k}}{|\boldsymbol{k}|^2} \cdot [\boldsymbol{F}]_{\boldsymbol{k}} - \frac{i\boldsymbol{k}}{|\boldsymbol{k}|^2}(\tilde{\xi} + \tilde{\nu}|\boldsymbol{k}|^2)[g]_{\boldsymbol{k}}$. We solve for $\boldsymbol{w}$ iteratively using the following recursion relation

$$\{[\boldsymbol{w}]_{\boldsymbol{k}}\}^{m+1} = (\tilde{\xi} + \tilde{\nu}|\boldsymbol{k}|^2)^{-1}\Big\{[\boldsymbol{F}]_{\boldsymbol{k}} - \frac{\boldsymbol{k}\boldsymbol{k}}{|\boldsymbol{k}|^2} \cdot [\boldsymbol{F}]_{\boldsymbol{k}}\Big\}^m - \frac{i\boldsymbol{k}}{|\boldsymbol{k}|^2}[g]_{\boldsymbol{k}}, \tag{S35}$$

where $\{\}^m$ denotes the expression evaluated at the mth iteration step (for a particular time step). We obtain $\{\boldsymbol{w}\}^{m+1}$ from the inverse Fourier transform of $\{[\boldsymbol{w}]_{\boldsymbol{k}}\}^{m+1}$, which is then used to update $\{\boldsymbol{F}\}^{m+1}$ in the next iteration step. The iteration ended when $\max|\{\boldsymbol{w}\}^{m+1} - \{\boldsymbol{w}\}^m| < 10^{-3}\max\{\boldsymbol{w}\}^{m+1}$. Finally, the pressure can be obtained by applying inverse Fourier transform to $[\phi p]_{\boldsymbol{k}} = -\frac{i\boldsymbol{k}}{|\boldsymbol{k}|^2} \cdot [\boldsymbol{F}]_{\boldsymbol{k}} + (\tilde{\xi}/|\boldsymbol{k}|^2 + \tilde{\nu})[g]_{\boldsymbol{k}}(\boldsymbol{k} \neq \boldsymbol{0})$.

C. Stream index

Simulations of the continuum model yielded two distinct morphologies of the cell populations: one consisting of stream-like aggregates of cells, and one consisting of droplet-like aggregates of cells. Thus, we sought to develop a metric to quantify the "stream-ness" of the overall morphology, which we term the "stream index" S. For simplicity,

we define the cell regions as $\mathcal{R} \equiv \{(x,y) | \phi(x,y) > 0.5\}$ using the phase field ϕ. To highlight the topological features of $\mathcal{R}$, we extract the skeletons $sk(\mathcal{R}_i)$ of all the connected components $\mathcal{R}_i$ in $\mathcal{R}$ (Supplementary Fig. 20). For an arbitrary shape $\mathcal{E}$, the topological skeleton $sk(\mathcal{E})$ is defined as the ridges (i.e., curvature singularities) of the distance transformation of $\mathcal{E}$, which calculates the minimal distance of any point $E \in \mathcal{E}$ to the shape boundary $\partial\mathcal{E}$. Note that the length of skeleton $|sk(\mathcal{E})|$ is zero for a circular shape, and $|sk(\mathcal{E})| > 0$ for an elongated shape. Thus, a dimensionless quantity $|sk(\mathcal{E})|/\sqrt{|\mathcal{E}|}$ can be used to quantify the "stream-ness" of shape $\mathcal{E}$, where $|\mathcal{E}|$ denotes the area of $\mathcal{E}$. Finally, for a structure $\mathcal{R}$ that contains n disconnected shapes $\mathcal{R}_i$, we define the stream index S as $S = \frac{1}{\sqrt{n}} \frac{\sum_{i=1}^{n} |sk(\mathcal{R}_i)|}{(\sum_{i=1}^{n} |\mathcal{R}_i|)^{1/2}}$ where the factor $1/\sqrt{n}$ is introduced such that n identical shapes have the same S as each individual shape. Our analyses show that the S index defined as such is able to clearly distinguish between stream-like ($S \geqslant 1$) and droplet-like ($S < 1$) morphologies (Supplementary Fig. 20).

V. CELL CULTURE AND STRAINS

A. Cell Culture and Growth

All strains used in this study are listed in Supplementary Table 1.

M. xanthus was grown in 1% CTT media consisting of 10 g/L Peptone (Ovia), 1% v/v Tris-HCl buffer, and 0.1% v/v 1 M KH2PO4 manually adjusted to pH=7.6. All cells were struck from frozen stocks onto 1.5% v/v agar/CTT plates. The day before each experiment, single colonies were picked from plates and incubated in 20 mL of CTT, shaking at 32°C overnight. Cells were collected at an optical density, OD600 ~0.4. We then concentrated or diluted cells to the desired cell density by centrifugation at 8000rpm for 5 minutes in a tabletop centrifuge (Eppendorf 5415D) and subsequently resuspending the pellet in fresh CTT of the desired volume.

F. johnsoniae cells were grown in 1% CYE media, first from frozen stocks on 1.5% agarose/CYE plates. Single colonies were selected and grown overnight in liquid CYE, shaking, at 26°C. Prior to each experiment, overnight cultures were back-diluted 1:100 in fresh motility media and grown shaking for 2 hours before the cells were spun and resubmerged in fresh motility media. CYE media consists of 10 g/L Peptone (Ovia), 5 g/L yeast extract (BD), 8 mM $MgSO_4$, and 10 mM Tris (pH=7.5). Motility media consists of 1 part CYE without $MgSO_4$ (see above) and 2 parts water.

E. coli was grown first as single colonies on 1.5% v/v agarose/LB plates. Single colonies were selected and grown overnight in liquid LB, shaking at 37°C. Prior to each experiment, overnight cultures were back-diluted 1:100 in fresh LB and grown shaking at 37°C for 2 hours before the cells were transferred to an agarose pad for imaging.

VI.   EXPERIMENTAL DEVICE AND SETUP

A.   Microscopy

All images presented were obtained on a Keyence VK-X1000 confocal microscope. A custom MatLab program was used for microscope control and image acquisition. At each time, the microscope captures an image of (1) the reflectance of the sample at each pixel and (2) the height of each pixel relative to the image's minimum. This is done by scanning through a finite axial range and determining, at each pixel, the optimal reflectance and its corresponding height using a proprietary algorithm. Height images were corrected for surface tilt, see VII A. For *F. johnsoniae* and *E. coli*, the cells were not sufficiently discernable in the reflectance images to be labeled and segmented by our neural network. Therefore, we also captured a color image of the cells at each time. This image by a similar scanning process as used to acquire the reflectance and height images, except the light is passed to a CMOS camera chip instead of a photodiode. A final image is constructed by determining, for each pixel, which point in the stack of collected images is the most in focus (as determined by an undisclosed algorithm).

All timelapses were captured at a rate of 1-3 minutes per frame. At each point in time, the microscope was automatically refocused and the appropriate scan range determined by either determining the range used in the previous frame and adding a small "buffer" to ensure a complete scan or by an automatic range-determination algorithm internal to the microscope. Lateral sample drift was corrected by post-processing, see Section VII B for details.

B.   Device Description

We designed a custom device to control the osmotic pressure across the hydrogel substrate. The device works by applying a hydrostatic head to the bottom of the gel. This hydrostatic head then serves as an additional pressure on top of the typical osmotic pressure (gradient) that exists in any hydrogel exposed to air. To do this, a reservoir of liquid media (or water) is connected directly to the gel. By moving the reservoir up or down relative to the gel, a hydrostatic head is applied to the gel. This hydrostatic pressure acts in addition to the usual osmotic pressure at the gel/air interface and thus directly modifies the water availability, and thus the cell wetting, at the surface. See Supplementary Figs. 3 and 4.

All custom parts were made by laser cutting acrylic sheets of varying thickness (specified below, per part). Acrylic sheets were purchased from McMaster-Carr and laser cut on a Universal Laser Systems VLC 3.60 laser cutter.

The base of the device consists of an elevated plate with a hole in the middle (Supplementary Fig. 3a). This is cut from a single piece of 0.5"-thick acrylic. This middle hole is cut such that it does not penetrate through the plate, but instead fits around and supports a male Luer lock (McMaster Cat. #) that will penetrate completely through the plate. The base is elevated by four 0.5^3" legs (0.5^2" squares cut from 0.5"-thick acrylic) that are glued to the bottom of the base with acrylic glue (SciGrip IPS Weld-On # 3). Halfway along each outer edge, a through hole is cut. Heat-set inserts (McMaster Cat. # 94459A320) for 8-32 screws were then fit into these through holes with a soldering iron.

Placed atop the base is a support plate, consisting of a 0.25"-thick circular piece of acrylic with a through-hole cut such that a male Luer lock can be press-fit into it. This plate should sit flush to the base when constructed properly. A spacer ring that outlines this plate is then placed on top of it. The thickness of this spacer ring will determine the thickness of the gel. For all experiments in this work, a 3 mm thick spacer was used. Finally, a thin plate (cut from 3 mm-thick acrylic) is placed on top. Crucially, this plate contains a 15mm opening in the middle, as well as four holes near its outer edge that are to be lined up with the threaded holes in the base. The middle, 15 mm-diameter opening in the middle of the plate is left exposed to air and is where the sample will be deposited and imaged. As such, its area is a key determinant of how rapidly the gel loses water due to evaporation. We found that 15 mm-diameter was wide enough to comfortably fit the nose of the objective close enough to sample such that it could be imaged while not being so wide that the the additional hydrostatic head from the reservoir would be insufficient to keep the gel hydrated over the course of long (multi-hour) times.

The device is assembled by stacking the components, as previously described (Supplementary Fig. 3). Finally, the pieces are fastened together by pushing four 8-32 screws through the top plate through holes and into the inserts in the base. Relevant dimensions for the device are shown in Supplementary Fig. 3c,d.

After the pieces are fastened together, a syringe filled with media and connected to a fluid line (Masterkleer PVC plastic tubing, McMaster Cat. # 5233K51) is connected to the Luer lock on the plate and filled completely such that no air bubbles are in the line. Molten agarose gel is then poured into the top opening until it is completely filled. The gel is then allowed to cool and thus, solidify. We note that while a small amount of agarose does go into the line, this can be minimized (¡1mm depth into the gel) by pouring the gel at a temperature just above it's melting temperature. Once the gel has solidified, cells are pipetted onto the gel, the spot allowed to dry, and the device is transferred to the microscope. Here, the syringe is disconnected from the fluid line and the line is reconnected to the reservoir. As a reservoir, we used plungerless 10 mL syringes with female Luer locks screwed onto the end. The reservoir was mounted to a rail connector that then ran along a piece of optical dovetail rail mounted perpendicular to the optical table. After setting the reservoir height, we allowed the system to stabilize for 15 minutes before starting to image.

C.   Measurement of Liquid Surface Tension

Surface tension for all media used was measured by the pendant drop tensiometry using a custom-built tensiometer. The tensiometer consists of a Basler Ace acA3088-57uc camera equipped with a Computar MLH-10X close-up manual zoom lens focused on a backlit, blunt-end dispensing syringe. Pendant drops were manually loaded into the syringe and imaged once before being pushed out of the syringe and replaced with a new drop. Each trial consists of a single, imaged drop. Images were processed using OpenDrop [12].

Measured surface tensions for all media used in this study is given in Supplementary Table 3.

VII. IMAGE AND DATA ANALYSIS

All analysis was done in Python 3.9 [13] using the NumPy [14], SciPy libraries [15], and scikit-image [16] libraries. Plots were generated using Matplotlib [17].

A. Image Height Correction

Height maps for all analyzed images were corrected for surface tilt. To do this, we first automatically identified pixels corresponding to the underlying substrate and then fit them to a first or second order polynomial.

To identify the substrate, the reflectance image was converted to an 8-bit image and a local entropy filter was applied using a diameter=30 pixel disk as a structuring element (Fig. 26c). The resultant image was then flattened and the values histogrammed to generate a distribution of entropy values (Fig. 26d). Because (groups of) cells exhibit greater local entropy due to their exhibiting an edge around the surface, the distribution of entropy values was typically bimodal with one mode representing the typical (lower) entropy value of the surface versus the other being higher and typically corresponding to the cells. We fit the resulting distribution to a 2 component Gaussian mixture model and determined the threshold for calling a pixel part of the surface by multiplying the mean of the smaller component by 1.1.

Heights for the list of resulting (x, y) pixel coordinates that are considered part of the underlying substrate were then fit to a three-dimensional surface. We found that for all experiments, either a 1st ($h(x,y) = ax + by + c$) or 2nd ($h(x,y) = ax^2 + by^2 + cxy + d$) were sufficient models of the underlying surface ($R^2 > 0.9$, Fig. 26e). This fitted function was then applied to all (x,y) coordinates to generate a representation of the substrate and subtracted off from the original height map to generate a "corrected" map. This corrected map was used for all analyses.

B. Timelapse Drift Correction

To account for global sample drift over the course of long time series, images were aligned into a global coordinate system for each time series. Subsequent analyses were conducted in this coordinate system. To do this, we first calculated a relative shift between all pairs of adjacent frames by calculating the phase difference between the Fourier transforms of these frames and finding the maximum value of the resulting cross power spectrum. See [18] for details. These shifts are then cumulatively summed to yield a "global" shift for each frame, assuming that the first frame is unshifted. The largest global shift also yields a sufficient amount of padding that is applied to all images such that none of them are truncated upon shifting. Shifts are applied to each image by Fourier transforming each image, multiplying the result by a the shift operator $e^{i\vec{\alpha}\omega}$ where $\vec{\alpha}$ is the desired shift and ω is spatial frequency, and then inverse Fourier transforming the result.

C. Convolutional neural network-based cell segmentation

1. Network Architecture

We use a standard U-net architecture [19]. The contracting and expanding paths of the net consist of 5 convolutional blocks, each corresponding to a single spatial scale. Each convolutional block consists of 2 convolutional layers with ReLU activations, 3x3 pixel kernels, and implicit padding such that the output maintains the shape of the input. Spatial downsampling in the contracting path is done by 2-dimensional maximum pooling such that the spatial dimensions of the input are reduced by a factor of 2 after each pooling operation. In the expanding path, we use bilinear interpolation to expand the spatial dimensions by a factor of two between spatial scales. All network weights are initialized with Kaiming initialization [20].

For *M. xanthus*, the network takes as input both the laser intensity image and its corresponding height map, z-scored with respect to the mean and standard deviation of the training set data. For *F. johnsoniae* and *E. coli*, the RGB image collected at each time point is concatenated onto the two-channel laser intensity image and height maps. For tracking single *M. xanthus* cells in groups, an additional dataset was generated similar to that for *F. johnsoniae*

and *E. coli* wherein an RGB image of the cells was also collected. This 5 channel image is similarly z-scored with respect to the mean and standard deviation of the training set data. The network outputs a multi-channel confidence map that predicts, independently, the probability that each pixel does/not belong to either a cell or border pixel.

All code was written in PyTorch [21].

2. Training set generation

A ground truth dataset of cell masks was generated by manual annotation of 191 images for *M. xanthus*. For *F. johnsoniae*, a ground truth dataset of 46 images was used. For *E. coli*, 19 images were included in the ground truth dataset. For the dataset of *M. xanthus* cells with included RGB images, $52 images$ were included. Cell masks were manually annotated using either a custom Matlab GUI, or by manually drawing the masks on an iPad using the ibis Paint app. First, the GUI (or iPad) was used to create an initial dataset consisting of matched pairs of cell masks and their corresponding laser intensity and height images. This dataset was then used to train a preliminary version of the cell segmentation network. The trained network was then used to predict segmentations on new images, which were then manually corrected using a custom Matlab GUI before being added to the ground truth dataset. This "human in the loop" procedure was repeated until satisfactory network performance was reached.

A "cell border" class was added to the ground truth cell masks following the procedure in [22]. Briefly, cell masks were dilated with a 2 pixel-radius disk and all pixels in the resulting mask that are not labelled as "cell" in the ground truth are marked as "border." We note that in our labeling, cells by definition must have at least a 1 pixel gap and thus a border region always exists between 2 cells.

During training, this ground truth dataset was split 80/20 into training and validation sets.

3. Training details

Segmentation networks were trained using the Adam optimizer with initial learning rate 0.0005 and moment estimate decay rates, β_1, β_2 set to 0.9, 0.999, respectively. Over the course of training, the learning rate was reduced by a factor of 0.25 after the validation loss had reached a plateau for 100 epochs. To prevent overfitting, the network was trained until the validation loss reached a plateau for 250 epochs (early stopping). We use dice-weighted cross entropy (Eqn. S36) as a loss function where the cross entropy term was annealed (τ in the Eqn. S36) over the course of training such that the final network was trained solely on the dice loss.

$$\text{Loss} = \underbrace{\frac{2|\sigma(X) \cap Y|}{|\sigma(X)| + |Y|}}_{\text{Dice}} + \tau \underbrace{[Y\log(\sigma(X)) + (1 - Y)\log(1 - \sigma(X))]}_{\text{Cross Entropy}} \tag{S36}$$

To do this, we used a linear annealing schedule such that at each epoch, e the cross-entropy term was multiplied by $1 - \min(e/(\text{max. epochs}), 1)$.

Networks were trained on batches of 4 512x512 pixel crops, drawn randomly from the training set. During training, the training set was augmented by random horizontal and vertical flips ($p_{\text{flip}} = 0.5$), random rotations ($\pm 15°$), and random affine shears ($\pm 15°$ along both x- and y-axes).

Separate networks were trained for each species.

D. Tracking and Counting Cell Groups

1. Cell Group Detection

Groups of cells were detected by thresholding the (corrected) height images (see VII A) such that all pixels above a set threshold height were called as part of a group while all pixels below were marked as background. Groups were then determined by labeling connected components in the thresholded image. Connected components smaller than the average area of a single cell (100 pixels for *M. xanthus*) were discarded. Ideally, the height threshold used for this determination would be a very small value (*e.g.* 1 nm) above the detected surface, which is by definition set as

reference 0 nm. We found that doing so was not robust to experimental noise and surface heterogeneities while also causing groups that split apart to briefly, apparently "re-merge" due to small height variations just after splitting. To balance faithful group segmentation and robustness to such noise, we set the threshold to 1/2 the average cell height (for *M. xanthus*, 250 nm).

2.   Group Tracking

Groups were tracked over time by on a frame-by-frame basis, linking adjacent pairs of frames until a full time series of tracks had been accumulated. To do so, we used the pixel masks generated by our group detection algorithm (see VII D 1) and for each frame pair, constructed a bipartite graph where vertices represent labeled group masks and edges connect the vertices of masks in adjacent frames that have a non-zero number of overlapping pixels. We chose an optimal frame-to-frame matching of such masks by computing the maximum weight matching of the just-constructed bipartite graph. Sometimes, such groups either split into two, smaller groups or merged with adjacent groups to create a single, larger group. We maintained a "lineage tree," that kept track of which groups were the product of such events. Merge events are detected by allowing unmatched groups in the forward direction to their best match, regardless of whether or not this matched group has already been matched to another group in the original frame. Splitting events are detected by running the matching in reverse and repeating the procedure done to detect merges. This tree was used when inferring the number of cells in each group (see VII D 3). To detect individual tracks within these tree-structured graphs, we enumerated all straight paths in the graph by depth-first search. Starting from the tree root, a consecutive string of vertices with only a single in- and out-neighbor were labeled as a track. When a node with more than 1 out-neighbors was detected (corresponding to a split event), that node was called as being a part of the current track while a new track was initiated at each out-neighbor. When a node with more than 1 in-neighbors was detected (corresponding to a merge event), the track was terminated at the current node. The search was continued until all vertices were labeled as being part of a track.

3.   Cell Counting in Groups

We found that using laser and height images, alone, was not of sufficiently high resolution to accurately segment and track cells over time within groups[3]. To remedy this, we instead opted to simply "count" cells in tracked groups. To do this, predicted cell masks were combined with the detected cell groups, yielding a list of tuples for each group containing the frame number and the number of cells detected within the group at that frame.

We first calibrated the fidelity of these counts by detecting cell groups in our ground truth, manually labeled dataset (see VII D 1). We then segmented the images using our convolutional neural network and compared the detected number of cells in each group to the ground truth (Fig. 25a). We found that, for a given group size N, the average predicted size over all predicted groups was a good match to the ground truth (Fig. 25b, top) while the variance of these predictions was underdispersed relative to a Poisson counting model (Fig. 25b, bottom). We found that the distribution of predicted sizes P was well-described by a generalized, underdispersed Poisson distribution[23]:

$$p_{\lambda,w}(C) = \frac{\lambda(\lambda + wC)^{C-1}}{C!} e^{-\lambda - wC} \tag{S37}$$

which has mean $\mu = \frac{\lambda}{1-w}$ and variance $\sigma^2 = \frac{\lambda}{(1-w)^3}$. We calibrated the dispersion parameter, w, by finding the optimal w that would allow us to accurately predict the ground truth by computing $P = \lceil \mu \rceil$ of the fitted distribution. We found that $w = -0.94$ allowed us to perfectly predict the number of cells in groups of all sizes $N \leq 16$ except for $N = 15$ (Fig. 25c).

For cell groups that never merged or split with other groups, the number of cells in that group was inferred by fitting the distribution of number of detected cells in each frame to Eqn. S37 with $w = -0.94$ and computing $P = \lceil \mu \rceil$ as was done in the calibration procedure, above. For cell groups that merged of split with other groups, we leveraged the dependency structure implied by these events to better inform our estimation. At each merge/split event, we

---

[3] Note that the data shown in Fig. 3e was taken using a high resolution scan as well as an accompanying RGB image (see VI A, VII C)

performed joint inference on the larger and two smaller groups by asserting that the count for the largest group (the merged group or the group that splits apart), must be the sum of the two smaller ones. As was shown by Consul and Jain [24], $p_{\lambda_0,w}(C) = p_{\lambda_1,w}(C) + p_{\lambda_2,w}(C) = p_{\lambda_1+\lambda_2,w}(C)$ and thus we can obtain all three λ_x parameters for the 3 groups (and then compute $P = \lceil \mu \rceil = \lceil \frac{\lambda}{1-w} \rceil$) by simultaneous fitting of all the data for all 3 groups.

E. Single cell tracking

1. Candidate cell enumeration

We developed a simple enumeration scheme in which labeled regions found by watershed segmentation of a segmentation neural network's output probability "image" that were in direct contact with other labeled regions are grouped together from which a list of mutually exclusive candidate solutions is generated per-such-group. Direct contact was established between potentially-adjacent regions by computing each region's exterior boundary pixels and checking for overlapping pixels between pairs of region borders. Exterior boundary pixels were calculated by morphological dilation of each region followed by subtraction of the region's original pixels. Groups of continuously touching regions[4] were labeled as such. For each of these groups, a list of mutually exclusive candidate solutions is generated. To generate the list, we assume that the regions generated by the initial watershed segmentation cannot be further split to yield valid regions. We further assume that each pair of touching regions could have come about through the aberrant splitting of a single cell. We enumerate all such individual cases within a group along with all possible combinations within which such a merge can exist alongside the other cells in said group. The above-described enumeration is then done recursively until the group is merged into a single region. This scheme yields a list of all possible cells that could have arisen from within the group, as well as a list of mutually exclusive combinations of cells that could have arisen from the network output and watershed segmentation.

2. Non-motile pairs of E. coli & F. johnsoniae

To track post-division daughter cells in non-motile *E. coli* and *F. johnsoniae*, we first generated a series of predicted cell probabilities using our convolutional neural network (see VII C). At each frame, these probabilities were used to generate a list of candidate cells using the algorithm described, above (see VII E 1). Simultaneously, candidate cell groups were tracked (see VII D). Cell groups were then linked to candidate cell segmentations by detecting pixel overlaps between the two sets of masks. At each frame and for each group, we checked whether that group contained a candidate segmentation with exactly two cells and if so, included it in a running list of such frames for that group. All such frames were then compiled and cells from adjacent frames in this sequence were linked to one another by maximizing the total intersection over union of cell masks between the masks of these two frames. This maximization was accomplished by treating the adjacent frame pairs as a bipartite graph with nodes representing cell masks and edges representing the (negative) intersection over union value between the two masks and using minimum weight matching.

Calculation of the relative centroid-to-centroid distance for these pairs (shown in Fig. 2i) was done by first extracting the centroid of each cell mask at each frame and computing the Euclidean distance between them. This yields an absolute distance between the two centroids. The sequence of such distances was then smoothed by a 3-frame rolling mean filter for each pair of daughter cells (Fig. 27a). For each pair, the time at which the centroids were maximally separated was found and set as a reference 0 (Fig. 27b). Finally, we divided all the absolute centroid-to-centroid distances by the value of maximum separation to yield a relative centroid-to-centroid distance (Fig. 27c). Note that in Fig. 2i, time was recalculated relative to that shown in Fig. 27c by setting the minimum value to be 0 instead of the reference time to which all of the sequences were aligned.

---

[4] Note that touching is not necessarily an all-to-all relation. If region **A** touches **B** and **C** but **B** and **C** do not touch, the resultant group would include **A**, **B**, and **C**.

3. Motile M. xanthus

Motile pairs of *M. xanthus* cells within a phase (as shown in Fig. 3e) were first manually annotated in a single frame at the start of the time series. For each subsequent frame, all potential cell segmentations were enumerated using the algorithm described, above (see VII E 1). These were then linked in time using a novel algorithm that solves a constrained linear sum optimization problem across a series of frames.

F. Wetting Profile Measurement

To measure the capillary length of cells in experiment, we first isolated single cells (or pairs, as in Fig. 2) using our group tracking and counting algorithm (see VII D 2). The contour of the "group mask" found by thresholding of the height map was extracted and the major axis of an ellipse fitted to these contour pixel coordinates was extracted by singular value decomposition of the contour pixel coordinates. To do this, the n contour pixel coordinates were mean-centered and concatenated into a $2xn$ matrix and the singular value decomposition of this matrix was computed. From this, we constructed a transformation that would take scalar values in the interval [-1,1] and map them onto the major axis of the cell by taking the dot product of the matrix of left singular vectors, U, with the matrix Is where I is the (appropriately-shaped) identity matrix and s is a vector of singular values and scaling the resulting matrix by $\sqrt{2/n}$ (Eqn. S38).

$$f = \sqrt{\frac{2}{n}}\underline{\underline{U}} \cdot (\underline{s}\,\underline{\underline{I}}) \tag{S38}$$

We used the resulting transform to generate points along the middle half — $\pm 1/4$ of the major axis length from the contour center, $(\bar{x}, \bar{y})$ — by computing $f([-0.5, 0.5]) + (\bar{x}, \bar{y})$. For each generated point along the major axis, we generated a linearly-spaced sampling line along the direction perpendicular to the major axis at that point. We sampled each point of this line by linearly-interpolating from the height map at the specified pixel coordinates. These values were then accumulated and averaged by distance from the major axis to yield a single curve of height vs. distance from major axis. This curve was then fit to a circular arc, as discussed in I C.

We note that fitting of a circular was unstable and often yielded invalid fits on our experimental data. For some analyses (Fig. 22, 12), we fit the measured meniscus profile to an exponential decay (Eqn. S39 and extracted the fitted decay length, ℓ_c.

$$h(x) = Ae^{-x/\ell_c} + c \tag{S39}$$

This measurement shows the same qualitative relationship with osmotic pressure as that shown in Fig. 1f (Fig. 23).

SUPPLEMENTARY TABLES

Species	Strain	Genotype	Reference
M. xanthus	DK1622	Wild-type	[25]
M. xanthus	DK10410	$DK1622\Delta pilA$	[26]
M. xanthus	SA3985	$DK1622\Delta frzZ$	[27]
M. xanthus	TM913	$DZ2\Delta cglB\Omega pilA$ (TetR cassette insertion at mxan_5783)	[28]
F. johnsoniae	CJ1827	Wild-type[a]	[29]
F. johnsoniae	CJ1922	$\Delta sprB$	[29]
E. coli	MG1655	Wild-type	[30]

[a] This strain is a rpsL mutant of the usual *F. johnsoniae* wild-type strain UW101.

Supplementary Table 1: Strains used in this study.

Symbols	Descriptions	Values	Notes
P_0	Reference osmotic pressure difference across the gel surface	2.8 ± 0.3 kPa	Fig. 1; *M. xanthus*
P_0	Reference osmotic pressure difference across the gel surface	1.7 ± 0.3 kPa	Fig. 1; *F. johnsoniae*
P_0	Reference osmotic pressure difference across the gel surface	1.3 ± 0.2 kPa	Fig. 1; *E. coli*
P_0	Reference osmotic pressure difference across the gel surface	1.6 ± 0.1 kPa	Fig. 1; polystyrene bead
$\xi_\parallel$	Longitudinal friction coefficient	280 Pa·min	Fig. 2e,f; *M. xanthus*
ϵ	Friction anisotropy	0.85	Fig. 2e,f; *M. xanthus*
$\xi_\parallel$	Longitudinal friction coefficient	393 kPa·min	Fig. 2i; *E. coli*
ϵ	Friction anisotropy	0.62	Fig. 2i; *E. coli*
$\xi_\parallel$	Longitudinal friction coefficient	305 kPa·min	Fig. 2e,f; *F. johnsoniae*
ϵ	Friction anisotropy	0.09	Fig. 2i; *F. johnsoniae*

Supplementary Table 2: Model parameters obtained by fitting to experimental data.

Liquid	Surface Tension (mean $\pm$ std. dev.), mN/mm	No. Trials
Water	71.7 ± 1.9	10
CTT	64.2 ± 0.9	10
CTT + *M. xanthus* wild-type	66.6 ± 1.2	10
CTT + *M. xanthus* $\Delta frzZ$	66.1 ± 0.9	10
LB	60.2 ± 1.1	6
Motility Media	66.1 ± 1.7	8

Supplementary Table 3: Measured surface tensions for all media used in this study.
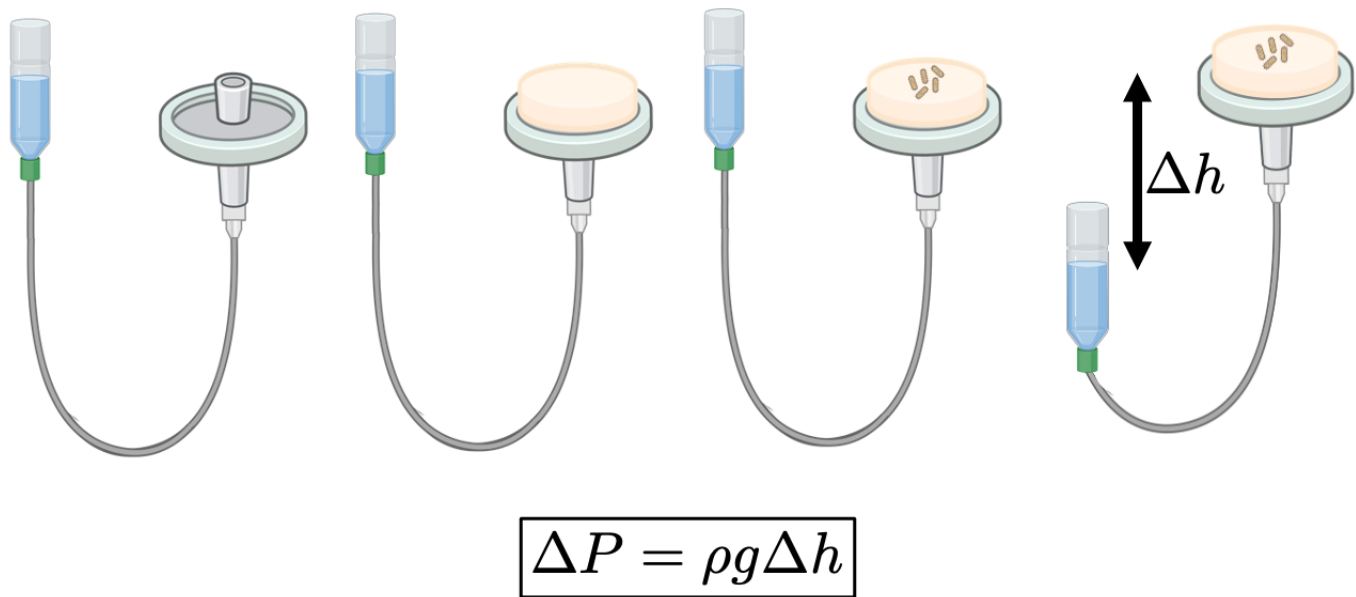
SUPPLEMENTARY FIGURES
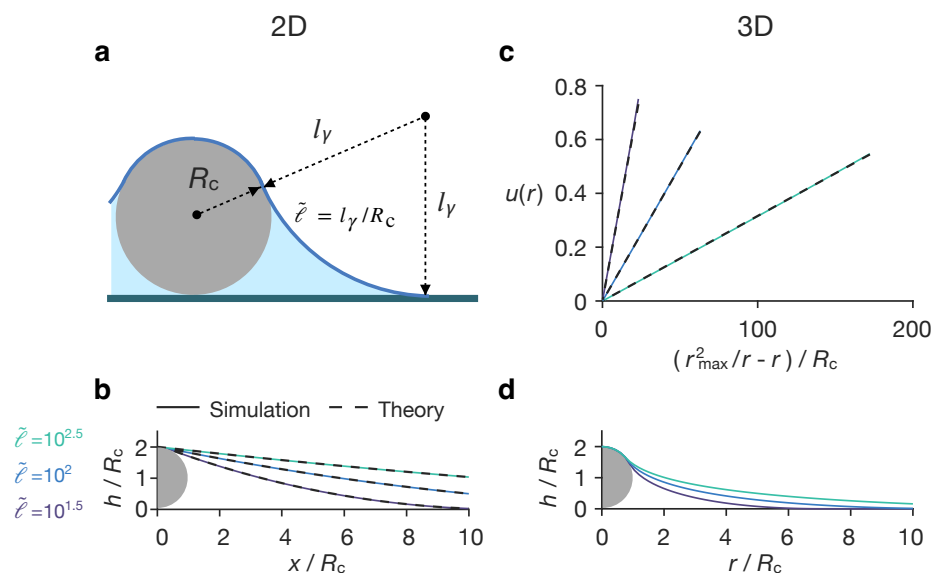
Supplementary Fig. 1. The shape of the water meniscus around a circular or spherical cell. (**a**) and (**b**) show results for a 2D circular "cell", and (**c**) and (**d**) show results for a 3D spherical "cell". (**a**) Schematic of the 2D height profile given by Eq. (S5). Color code as in main Fig. 1. R_c denotes the radius of the 2D circular "cell" and l_γ denotes the capillary length of the meniscus. $\tilde{\ell} = l_\gamma/R_c$ denotes the normalized capillary length. (**b**) Simulation (solid) and analytical (dashed) results for the equilibrium height profile of the water meniscus around a 2D circular "cell". (**c**) Height profiles obtained from the simulations of a spherical "cell" satisfy Eq. S7. Solid and dashed lines indicate the simulation and analytical results, respectively. See Sec. I B for details. (**d**) Simulation results for the equilibrium height profile of water meniscus around a 3D spherical "cell". In panels **b**−**d**, colors indicate the designated values of $\tilde{\ell}$.
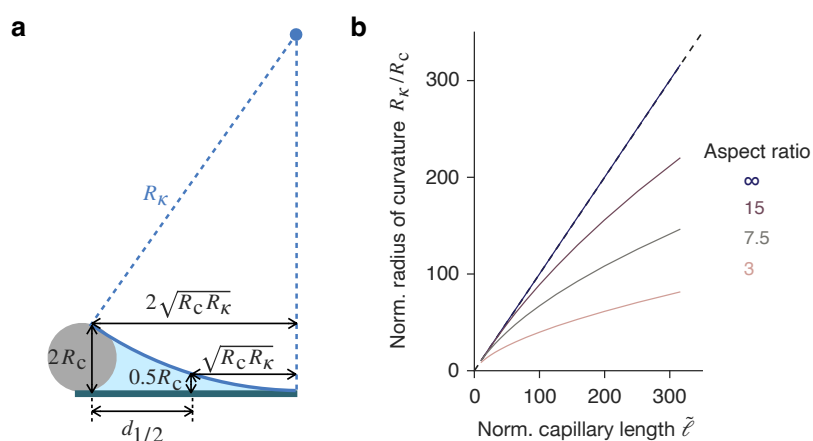
Supplementary Fig. 2. The midplane height profile varies with cell aspect ratio. (**a**) Schematic of the geometry used to estimate the radius of curvature R_κ of the midplane water height profile. Color code as in main Fig. 1. See Sec. I C for details. (**b**) The radius of curvature R_κ, normalized by cell radius R_c, plotted against normalized capillary length $\tilde{\ell} = l_\gamma / R_c = \gamma / (P R_c)$ at the designated cell aspect ratios. Black dashed line indicates $R_\kappa = l_\gamma$.
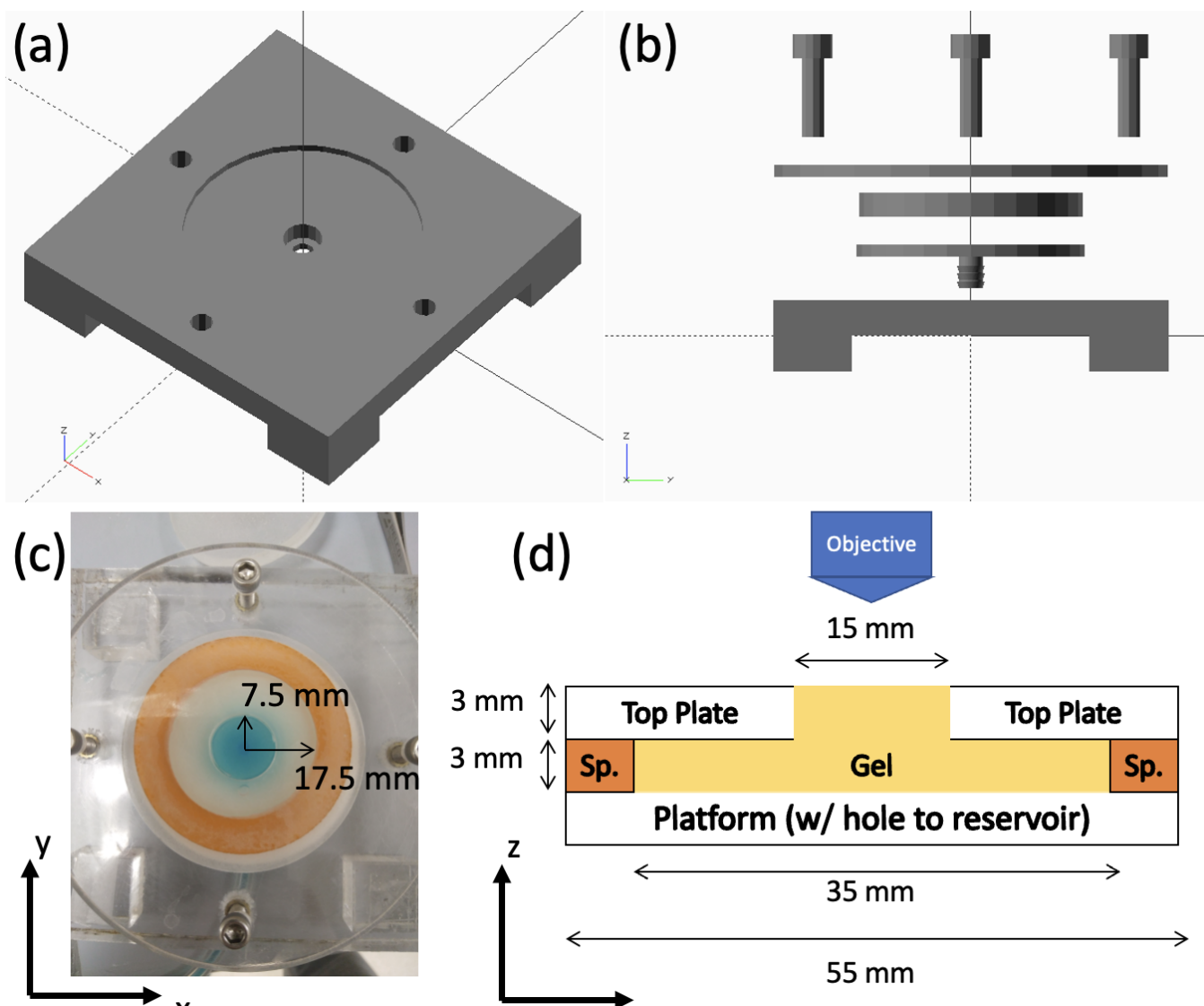
Supplementary Fig. 3. **(a)** 3d view of the base plate. Note the middle cutout has a lower lip that is used to support the outer edge of the Luer lock that will connect to the reservoir. **(b)** Exploded side view of the device, with all pieces shown in the order of their assembly. **(c)** Top view of an assembled device with a gel, with outer radii for the inner through hole and the top plate, itself, shown. **(d)** Side view schematic of the top part of the device, with relevant dimensions for all pieces shown.

$$\Delta P = \rho g \Delta h$$

Supplementary Fig. 4. Cartoon showing the hydrostatic pressure control device and its basic operating principle. A plate with a hole is connected to a reservoir of water (or liquid media). The gel is placed on top of the plate, directly coupling it to the reservoir. By moving the reservoir up or down relative to the top of the gel surface, a hydrostatic head is applied to the bottom of the gel. This head then acts on top of the osmotic pressure inherent to the gel, either "pushing liquid into" or "pulling liquid out of" the gel.
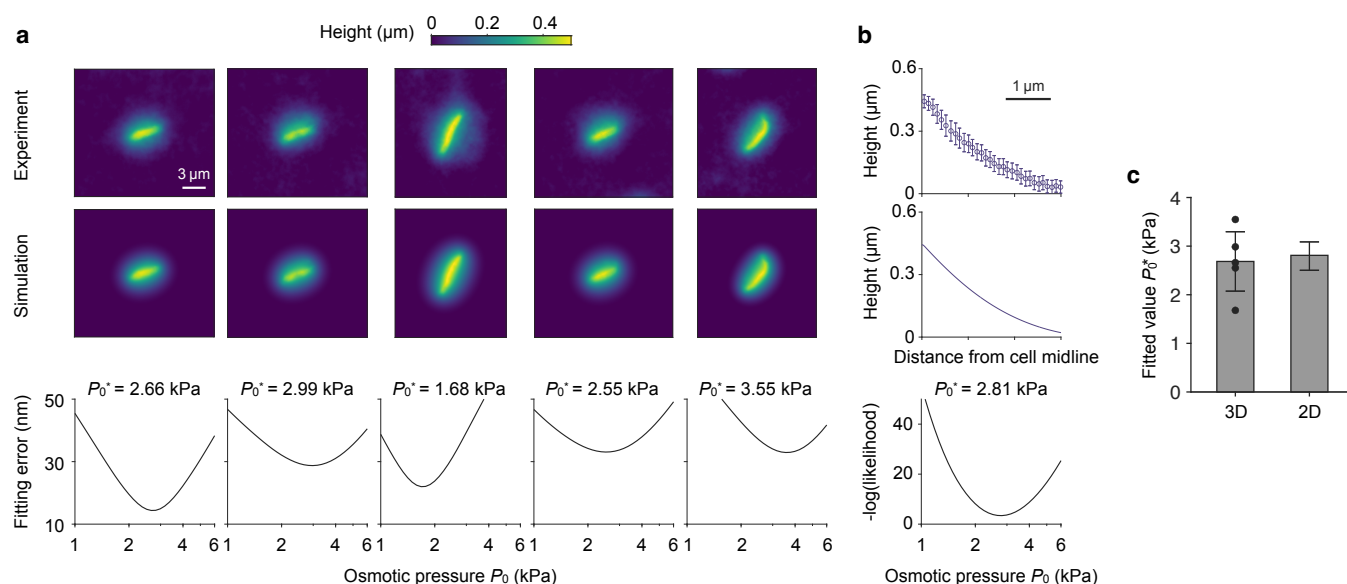
Supplementary Fig. 5. Fitting the 3D and 2D height profiles of the water menisci yield similar results. (**a**, **b**) The osmotic pressure difference P_0 is determined by fitting the 3D (**a**) or 2D (**b**) height profiles of the modeled water menisci to those measured in experiments. *Top* panels show the experimental measurements and *middle* panels show the simulation results with the optimal fitting parameters P_0^*. *Bottom* panels show the fitting errors at varying values of P_0. In **a**, each column represents an individual cell and the fitting error is measured by the root-mean-square height difference between the model and the experiment. In **b**, the height profile is measured in the mid-plane perpendicular to the cell body axis, and the fitting error is measured by the negative log-likelihood of the fitting. See Sec. I D for details. (**c**) Bar plot for the fitted values of P_0. Fitting the 3D and 2D height profiles yield $P_0 = 2.69 \pm 0.61$ kPa and $P_0 = 2.81 \pm 0.29$ kPa, respectively.
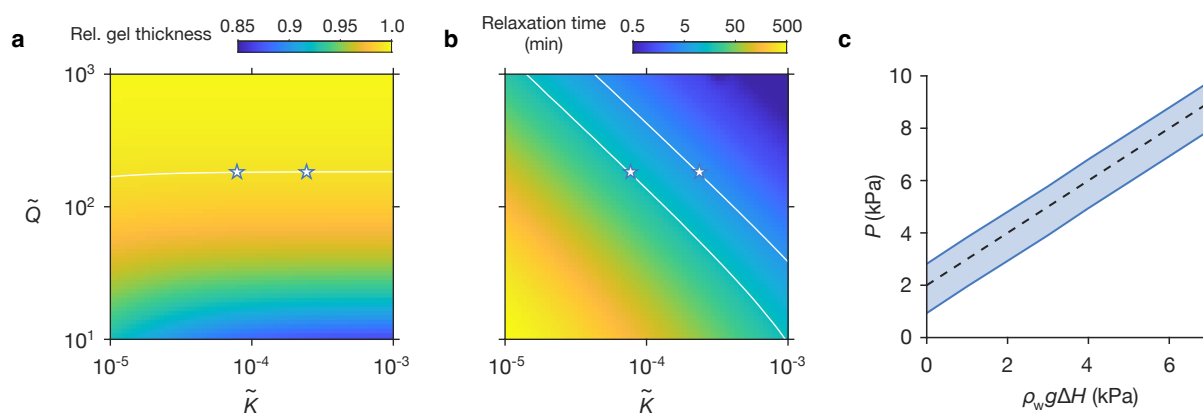
Supplementary Fig. 6. Water evaporation leads to a osmotic pressure gradient across the gel. (**a**) Relative gel thickness α and (**b**) the relaxation time τ_{ss} of the model (Eq. S10) at varying values of $\tilde{K}$ and $\tilde{Q}$. See Sec. I E for details. White curve in **a** denotes $\alpha = 0.99$. White curves in **b** denote $\tau_{ss} = 5$ min and $\tau_{ss} = 15$ min, respectively. White stars show the intersection points between the curve in **a** and the two curves in **b**, which determine the range of values for $\tilde{K}$ and $\tilde{Q}$ in experiment. (**c**) The osmotic pressure difference P across the gel surface at varying height differences ΔH between the surface of the gel and the water level of the reservoir. The top and bottom curves represent, respectively, the results for the parameters indicated by the left and right stars in **a** and **b**. Black dashed line indicates a slope of 1.
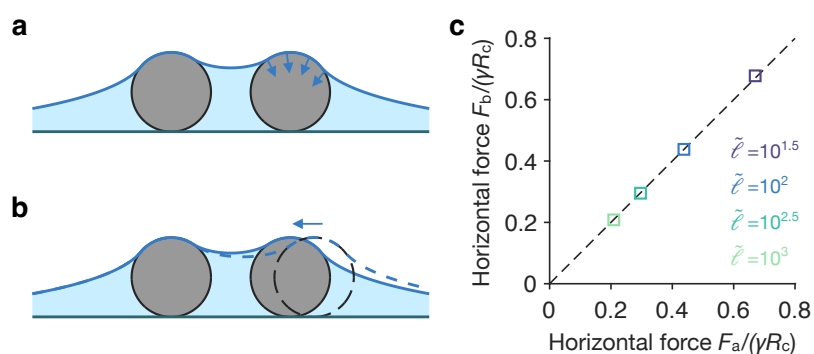
Supplementary Fig. 7. Calculation of the in-plane capillary force. (**a**, **b**) Schematics of two methods for calculating the capillary force exerted on the cell. The force can be computed from (**a**) a force integral over the region where the water meniscus is in contact with a cell, or from (**b**) the free-energy difference upon displacement of a cell (dashed). See Sec. I F for details. (**c**) Comparison between the force calculated using the method shown in **a** (denoted by F_a) and the force calculated using the method shown in **b** (denoted by F_b). The forces are normalized by γR_c. For reference, the dashed line shows $F_b = F_a$. Colors indicate the designated values of the normalized capillary length $\tilde{\ell}$.
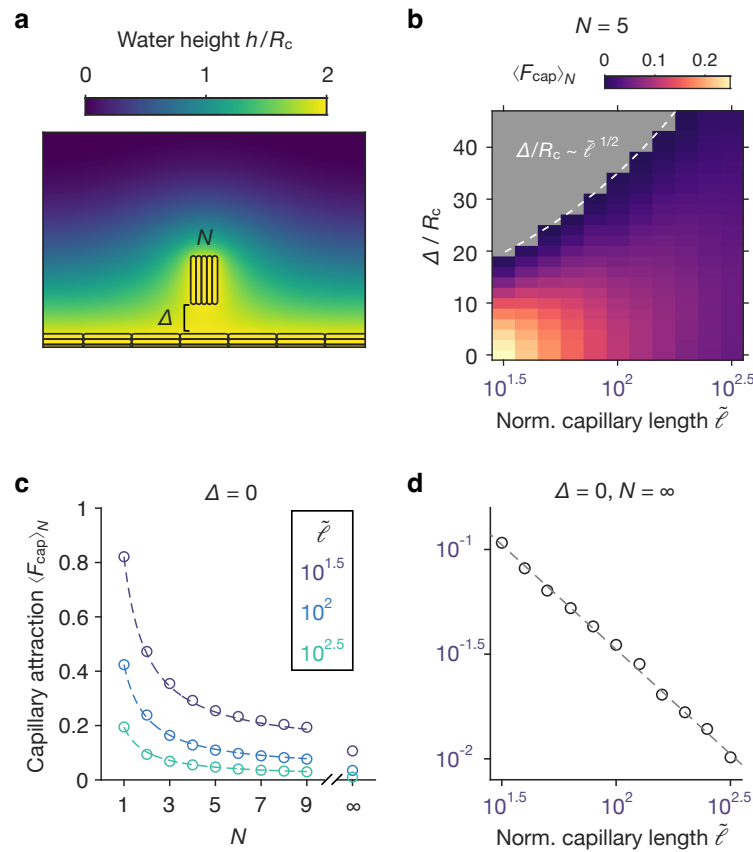
Supplementary Fig. 8. Characterizing the magnitude and range of capillary attraction. (**a**) Illustration of the simulation setup. A group of N cells is positioned a distance Δ away from an infinitely large colony. The average capillary force acting on each cell is denoted by $\langle F_{\text{cap}} \rangle_N$. (**b**) Capillary attraction $\langle F_{\text{cap}} \rangle_N$, normalized by γR_{c}, at varying distance Δ and normalized capillary length $\tilde{\ell} = \gamma/(PR_{\text{c}})$ for $N = 5$. Gray denotes that $\langle F_{\text{cap}} \rangle_N = 0$. White dashed curve indicates $\Delta/R_{\text{c}} \sim (\tilde{\ell})^{1/2}$. (**c**) Capillary attraction $\langle F_{\text{cap}} \rangle_N$, normalized by γR_{c}, at varying cell number N and normalized capillary length $\tilde{\ell}$ (color) for $\Delta = 0$. Circles indicate simulation data, and dashed curves indicate $\langle F_{\text{cap}} \rangle_N = \langle F_{\text{cap}} \rangle_\infty + (\langle F_{\text{cap}} \rangle_1 - \langle F_{\text{cap}} \rangle_\infty)/N$. (**d**) Capillary attraction $\langle F_{\text{cap}} \rangle_\infty$, normalized by γR_{c}, at varying normalized capillary length $\tilde{\ell} = \gamma/(PR_{\text{c}})$ for $\Delta = 0$. Circles indicate simulation data, and the gray dashed line indicates $\langle F_{\text{cap}} \rangle_\infty \sim \tilde{\ell}^{-1}$.
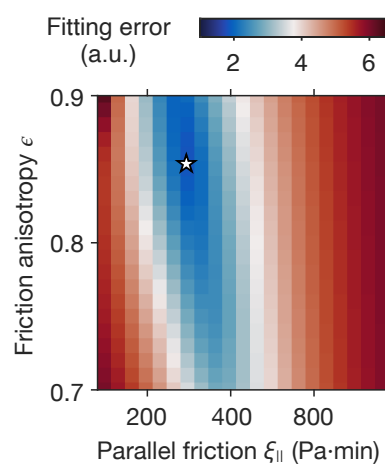
Supplementary Fig. 9. Fitting friction coefficients for *M. xanthus* cells. Fitting error of the merger dynamics between simulation and experiment at varying friction coefficient $\xi_{\parallel}$ and friction anisotropy ϵ (see Sec. II B). Star denotes the optimal fitting values.
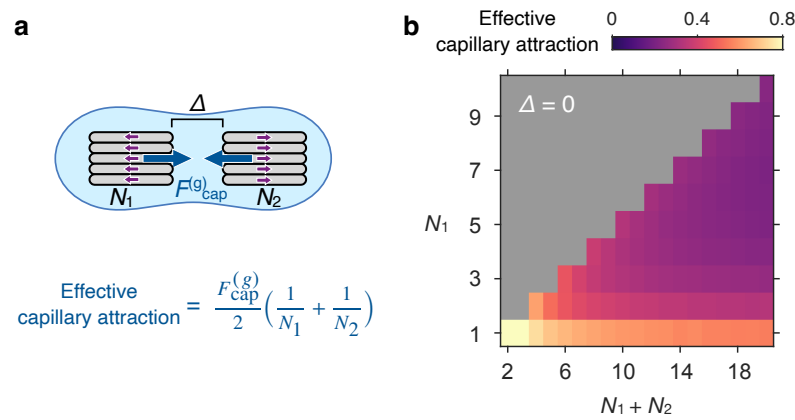
Supplementary Fig. 10. Effective capillary attraction varies with the partition of cell groups. (**a**) Schematic of the capillary attraction $F_{cap}^{(g)}$ between two subgroups of N_1 and N_2 cells that propel in the opposite directions. Δ denotes the distance between the two subgroups. The effective per-cell capillary attraction $F_{cap}^{(g)}(1/N_1 + 1/N_2)/2$ indicates the minimal propelling force for the two subgroups to split up. (**b**) Effective capillary attraction $F_{cap}^{(g)}(1/N_1 + 1/N_2)/2$, normalized by γR_c, at varying cell numbers N_1 and N_2 for $\Delta = 0$. Gray indicates $N_1 > N_2$ which is symmetric to $N_1 < N_2$.
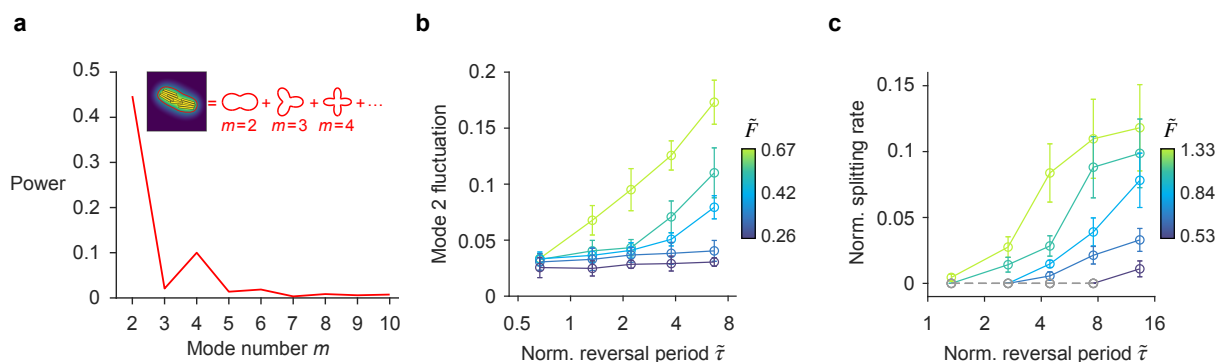
Supplementary Fig. 11. Elongation and splitting of a cell group. (**a**) Power spectrum c_m of contour deformation for the cell group shown in the inset. Deformation modes $m = 2$, 3, and 4 are illustrated in the inset. See Sec. II D for details. (**b**) Mode 2 fluctuation, defined as $\langle c_2^2 \rangle - \langle c_2 \rangle^2$, and (**c**) splitting rate of a 10-cell group at varying normalized reversal period $\tilde{\tau}$ and normalized self-propelling force $\tilde{F}$. In **b** and **c**, error bars represent standard deviations of 5 replicate simulations. In **c**, the dashed lines and circles indicate that the cell group does not split over the course of the simulation.
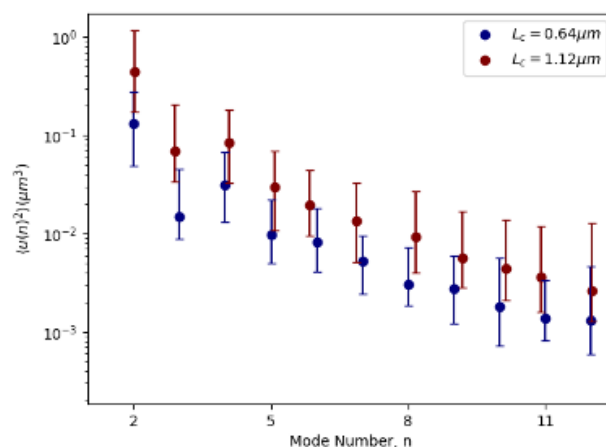
Supplementary Fig. 12. Power spectrum of contour deformations for small groups of cells vary with capillary length. Power spectrum of contour deformations for groups of size $N = 5$ *M. xanthus* cells, measured at different capillary lengths. Decreasing capillary length further confines the group, causing reduced fluctuations.
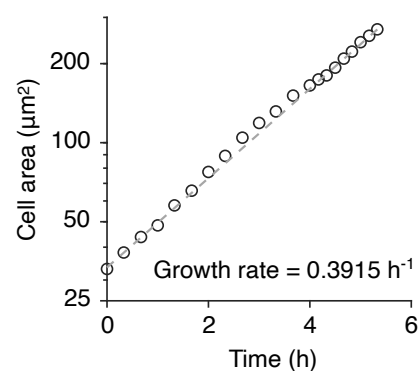
Supplementary Fig. 13. Fitting the areal growth rate of *E. coli* colonies. Area covered by cells as a function of time. Circles indicate experimental measurements and the dashed line indicates an exponential fit with the designated growth rate.
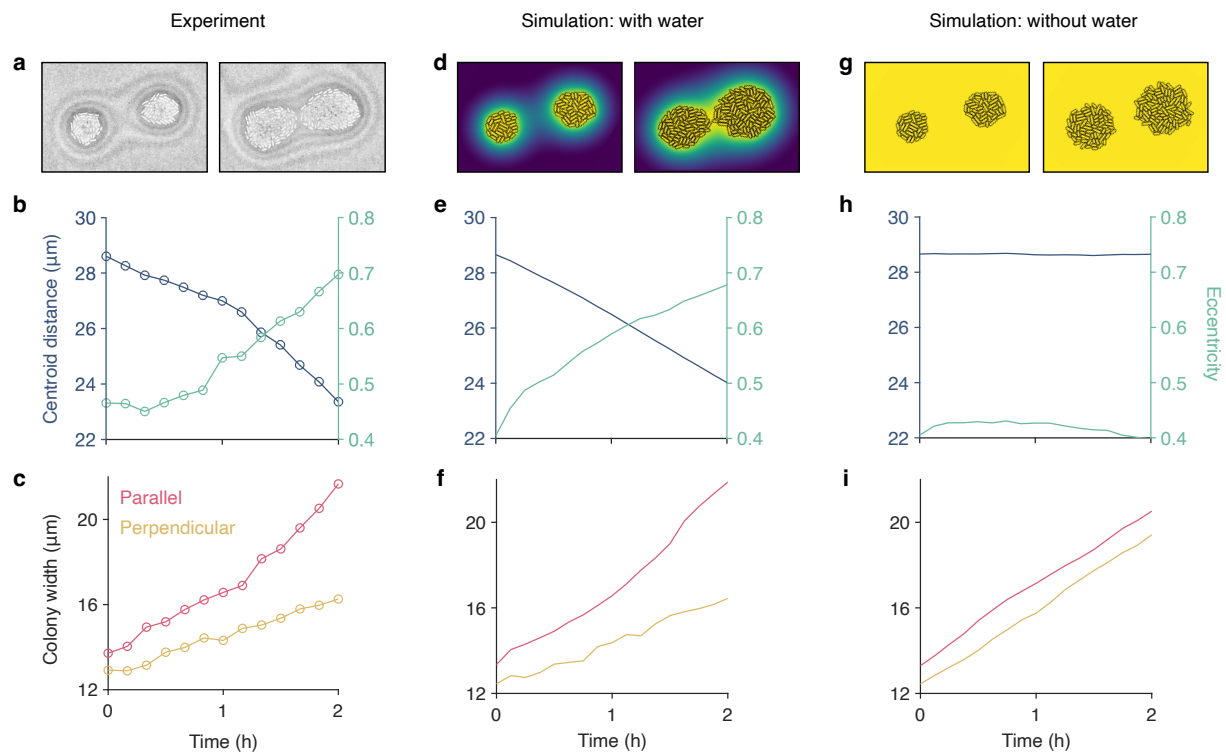
Supplementary Fig. 14. Capillary forces facilitate merger of growing *E. coli* colonies. The growth dynamics of two adjacent colonies are compared between (**a**–**c**) experiment, (**d**–**f**) simulation with water, and (**g**–**i**) simulation without water. (**a**,**d**,**g**) Snapshots of the colonies at $t = 0$ (left) and $t = 2$ h (right). (**b**,**e**,**h**) Time evolution of the centroid distance (blue) and average eccentricity (green) of the colonies. (**c**,**f**,**i**) Time evolution of the average colony widths in the directions parallel (red) and perpendicular (yellow) to the centroid-centroid connecting line.
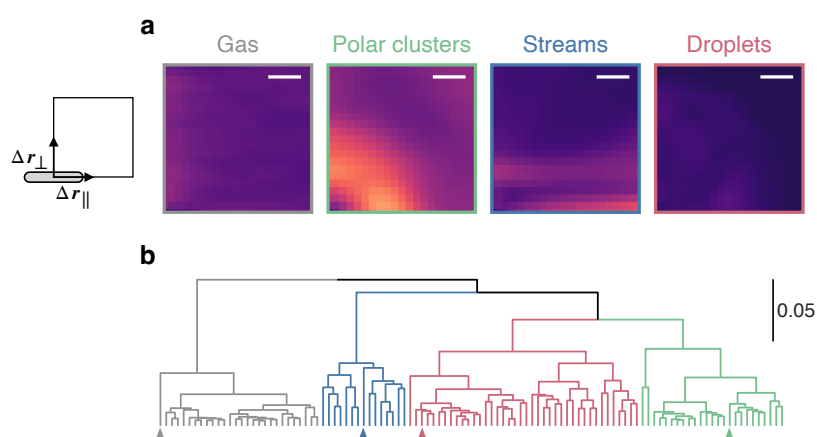
Supplementary Fig. 15. Classification of the simulated collective dynamics. (**a**) Heatmaps showing root-mean-square variation of the normalized density correlation function $\tilde{C}_\rho$ in time for the designated phases of collective dynamics. Scale bars: 3 μm. Plots correspond to main Fig. 3c,d. (**b**) The WPGMA dendrogram of the simulated collective dynamics. See Sec. III A for details. Scale bar indicates a distance of 0.05. Arrowheads indicate the four select examples shown in panel **a**.
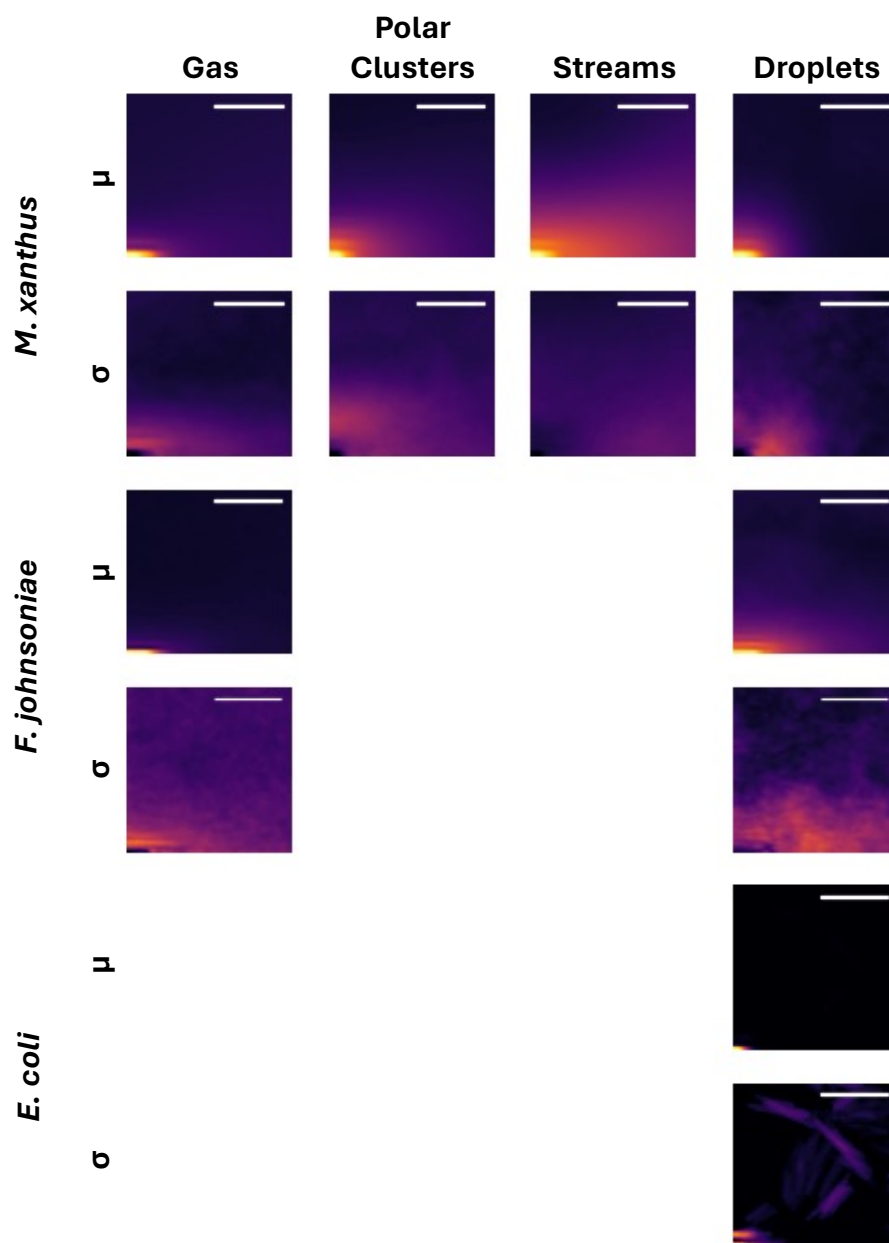
Supplementary Fig. 16. Density correlations for each collective phase found in all species studied. All examples shown were manually selected from representative experiments. Groups of two rows, one corresponding to the mean (μ) and the other the standard deviation (σ) of the density correlation function described in III B are given for each species. Scale bars: 5 μm.
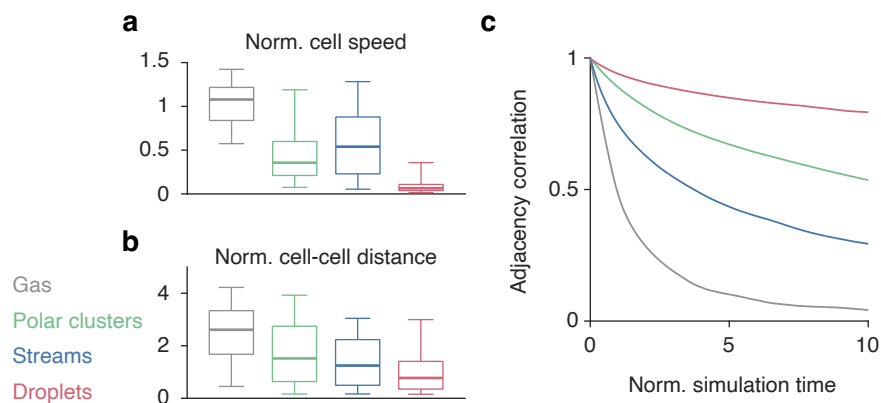
Supplementary Fig. 17. Different phases of collective dynamics display a mobility-adjacency trade-off. Box plots of (**a**) normalized cell speed and (**b**) cell-cell distance for the designated phases. Whiskers range from the 5th percentile to the 95th percentile. Boxes range from the 25th percentile to the 75th percentile. Horizontal bars inside the boxes denote the median values. (**c**) The average autocorrelation function of the cell-cell adjacency matrix (defined in Sec. III B) for the designated phases. The simulation time is normalized by the time $\tau_{\text{c}} = \xi_{\parallel} l_{\text{c}}^2 / |\boldsymbol{F}_{\text{prop}}|$ it takes for a non-reversing free-running cell to move over one cell body length. We varied three dimensionless parameters in our simulations: a normalized reversal period $\tilde{\tau} = \tau_{\text{r}}/\tau_{\text{c}}$, a normalized self-propulsion strength $\tilde{F} = |\boldsymbol{F}_{\text{prop}}|/PR_{\text{c}}l_{\text{c}}$ where $PR_{\text{c}}l_{\text{c}}$ is the scale of the longitudinal capillary force acting on isolated cells, and a normalized capillary length $\tilde{\ell} = l_\gamma/R_{\text{c}} = \gamma/(PR_{\text{c}})$. The simulation parameters are: (1) $\tilde{\ell} = 10^{3.5}$, $\tilde{F} = 10$, and $\tilde{\tau} = 10$ for the gas phase, (2) $\tilde{\ell} = 10^{2.5}$, $\tilde{F} = 1$, and $\tilde{\tau} = 10^2$ for the polar clusters phase, (3) $\tilde{\ell} = 10^{2.5}$, $\tilde{F} = 1$, and $\tilde{\tau} = 1$ for the streams phase, and (4) $\tilde{\ell} = 10^2$, $\tilde{F} = 0.3$, and $\tilde{\tau} = 3$ for the droplets phase.
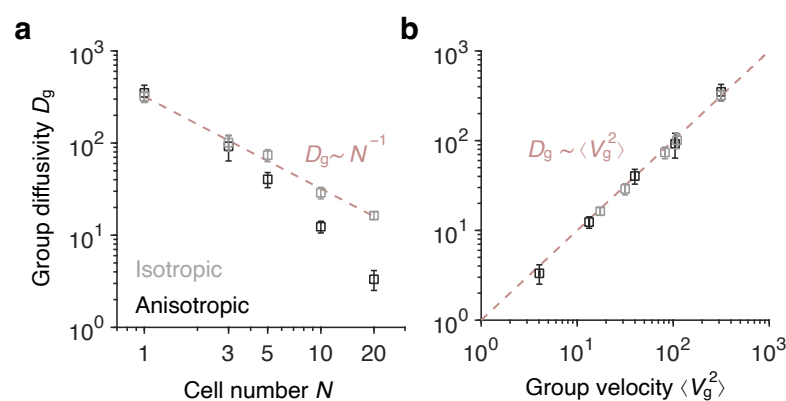
Supplementary Fig. 18. Group diffusivity decreases with increasing group size in the droplet phase. Group diffusivity D_{g} versus (**a**) the number N of cells in the group and (**b**) the mean squared velocity $\langle V_{\mathrm{g}}^2 \rangle$ of the group. Gray and black denote simulations with isotropic and anisotropic friction, respectively. Error bars represent standard deviations of 5 replicate simulations. Red dashed lines denote $D_{\mathrm{g}} \sim N^{-1}$ in **a** and $D_{\mathrm{g}} \sim \langle V_{\mathrm{g}}^2 \rangle$ in **b**.
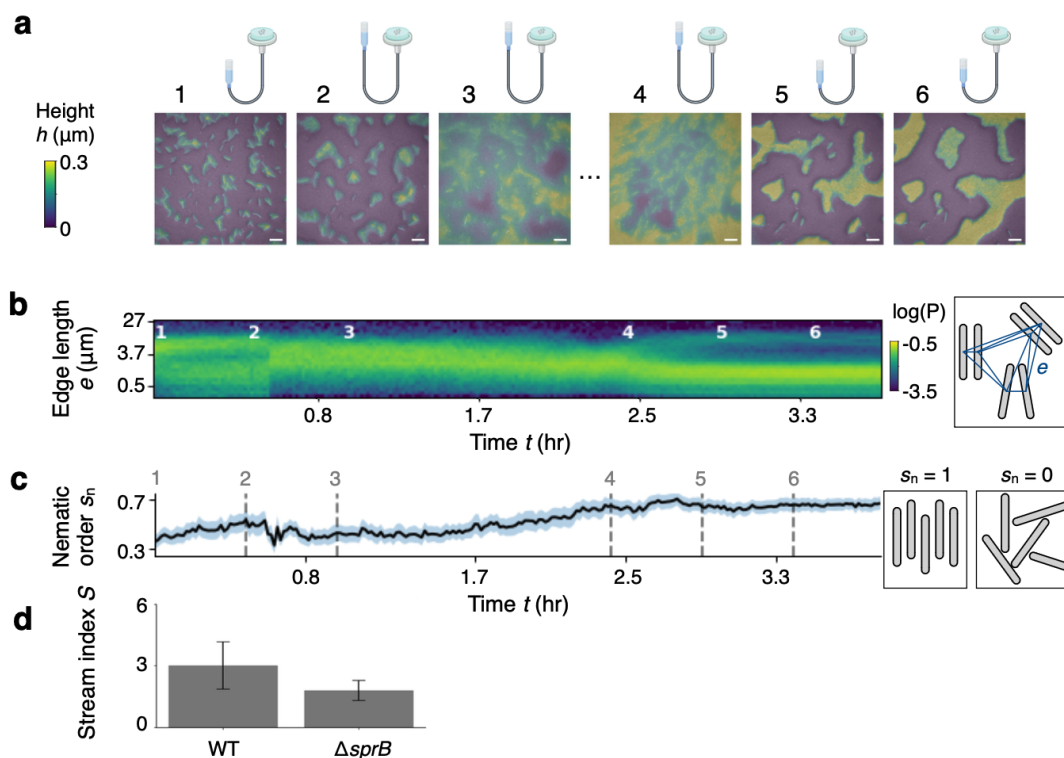
Supplementary Fig. 19. Water availability influences the organization of colonies of *F. johnsoniae*. (**a**) Experimental time series in which cells initially deposited onto an agarose pad are subjected first to increasing water availability until a gas phase is reached (snapshots 1–3). Hours later, water is drained from the substrate, and the colony goes back to forming tightly packed, nematically ordered groups and streams (snapshots 4–6). Time points (1-6) are marked in **b** and **c**. (b) Kymograph of the probability density function (P) of Delaunay edge lengths over time for the time series shown in **a**. Schematic on the right demonstrates how edges (blue lines) are calculated for a small population of cells (gray). At early times and late times in the experiment, low water availability at the surface constrains cells into tightly packed, well-separated groups. At intermediate times, high water availability permits cells to be further spaced apart and homogenizes the population. (**c**) Strength of nematic order s_n, calculated within groups of 8 nearest-neighboring cells and averaged over the population at each time. Shaded regions indicate variance. Right: schematics showing a small population of cells with perfect nematic order ($s_n = 1$) or no order ($s_n = 0$). (**d**) Stream index values for wild-type *F. johnsoniae* and a non-motile mutant ($\Delta sprB$). Error bars are standard deviations for each corresponding mean calculated over N=28 and 5 frames taken from 3 and 1 separate timelapses, respectively. Plots correspond to main Fig. 4.
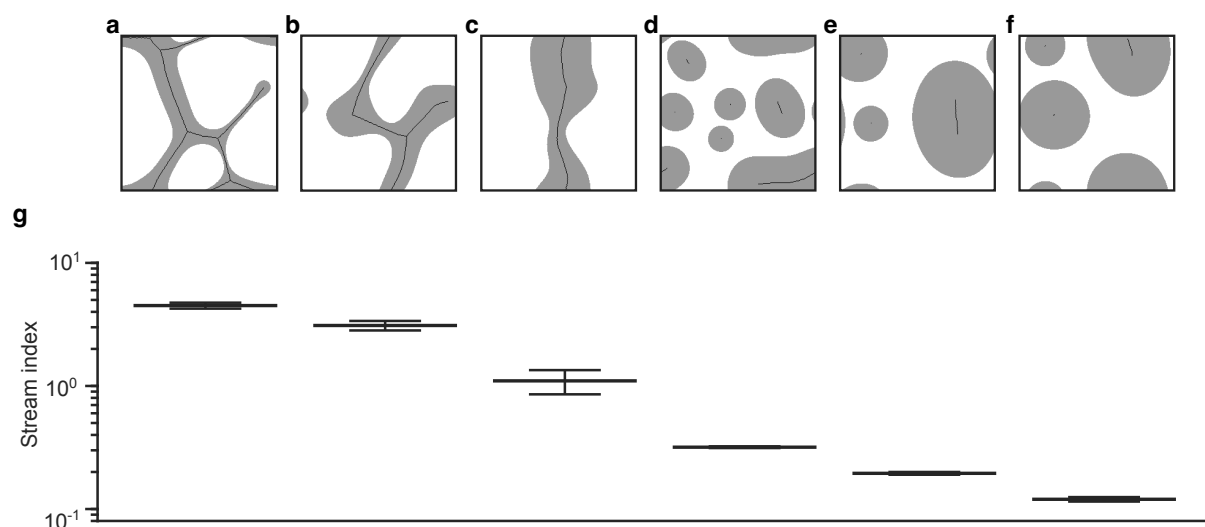
Supplementary Fig. 20. Quantifying the large-scale morphologies of cell populations. (**a**–**f**) Typical snapshots of the cell populations described by the continuum model with varying parameters (Sec. IV). White and gray denote voids and cell regions, respectively. Black curves indicate the skeletons of the cell regions as defined in Sec. IV C. Simulation parameters are: (**a**) $\sigma = 1, \tau_\rho = 10$, (**b**) $\sigma = 1, \tau_\rho = 10^2$, (**c**) $\sigma = 1, \tau_\rho = 10^4$, (**d**) $\sigma = 4, \tau_\rho = 10^1$, (**e**) $\sigma = 20, \tau_\rho = 10^4$, (**f**) $\sigma = 20, \tau_\rho = 10^1$. (**g**) Stream indices for the corresponding simulations in panels **a**–**f**. Data are represented as mean $\pm$ std with respect to time for each simulation. In panels **a**–**c**, we set $K_1 = 1$, $K_2 = 0.01$, and $K_3 = 1$. In panels **d**–**f**, we set $K_1 = K_2 = K_3 = 0$. See Sec. IV for details.
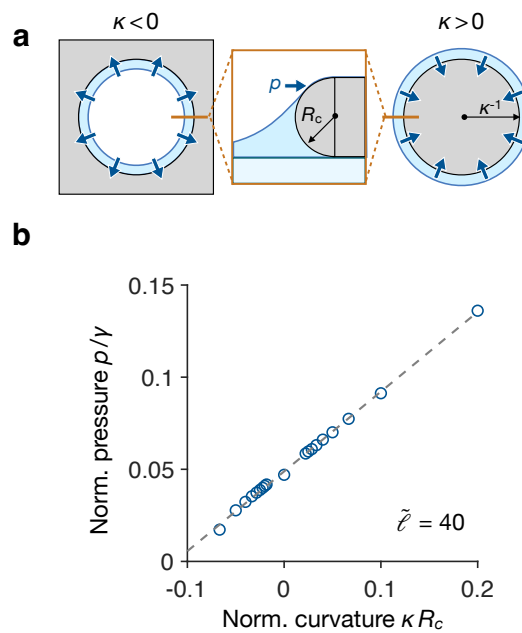
Supplementary Fig. 21. The in-plane capillary force acts as an Laplace pressure at large scales. (**a**) Schematic of the agent-based simulations. Gray denotes cell regions bounded by interfaces of curvature κ. Blue denotes the water. Blue arrows denote the in-plane pressure p exerted on the cells. R_c denotes the cap radius of each cell. The left and right panels show the top view schematics for $\kappa < 0$ and $\kappa > 0$, respectively. The middle panel shows a close-up cross section. (**b**) Normalized pressure p/γ at varying normalized curvatures κR_c for the designated parameter $\tilde{\ell} = 40$. Note that p denotes a 2D pressure, defined as force per unit length, and hence has the same unit as γ. Circles denote simulation results, and the dashed line indicates a linear fit $p/\gamma = \alpha_1 \kappa R_c + \alpha_0$ with fitting parameters $\alpha_1 = 0.433$ and $\alpha_2 = 0.05$.
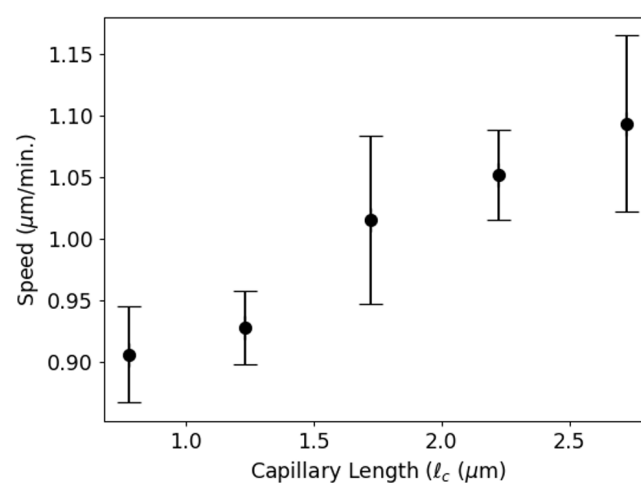
Supplementary Fig. 22. Speed vs. measured capillary length for single, motile *M. xanthus* cells.
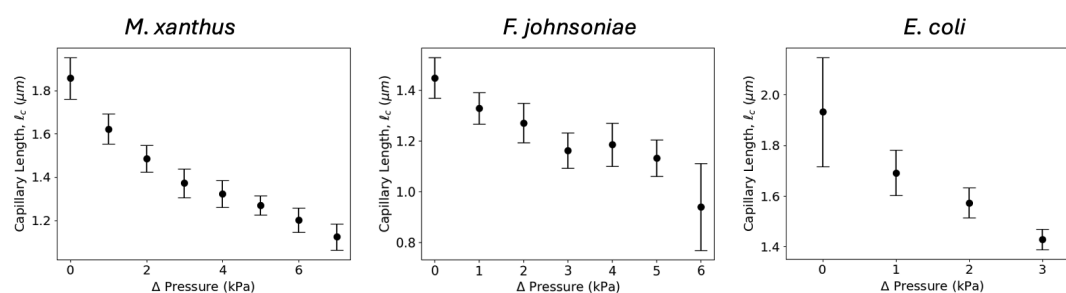
Supplementary Fig. 23. Exponential fits to measured meniscus profiles are anti-correlated with applied hydrostatic pressure. Single cells of the indicated species were deposited onto 1.5% v/v agarose pads and their wetting menisci measured (Capillary Length, ℓ_c, y-axis on all graphs) at varying levels of applied hydrostatic pressure (ΔPressure, x-axis on all graphs).
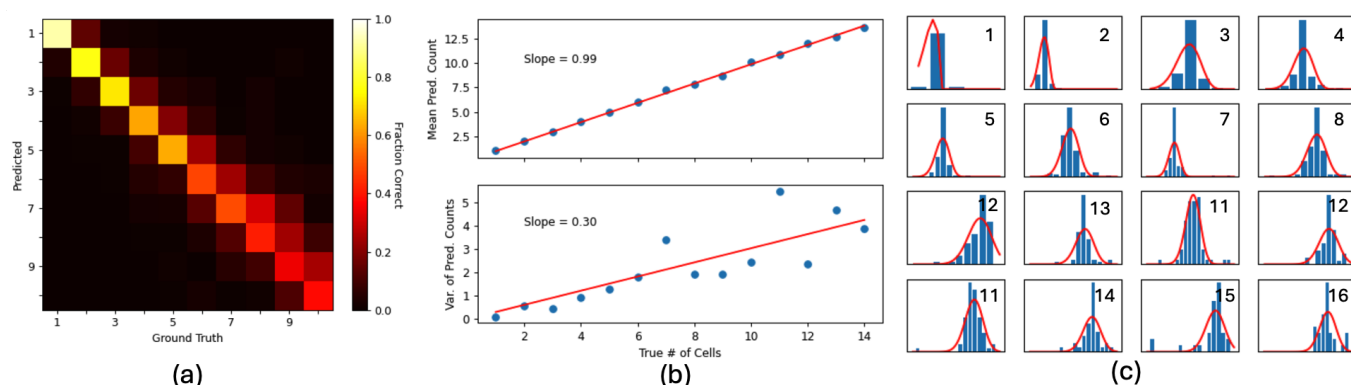
Supplementary Fig. 24. Calibration of the distribution of predicted counts for fixed group size. **(a)** Camera (RGB) image of dense groups of *F. johnsoniae* cells on an agarose pad, overlaid with the corresponding instance segmentation from our neural network. Region shown in (b) and (c) is shown with the white box. **(b,c)** Highlighted region of the original image (b) and the corresponding instance segmentation (c) for the region shown, boxed in (a). **(d)** Laser reflectance image of dense groups of *M. xanthus* cells on an agarose pad, overlaid with the corresponding instance segmentation from our neural network. Region shown in (e) and (f) is shown with the white box. **(e, f)** Highlighted region of the original image (e) and the corresponding instance segmentation (f) for the region shown, boxed in (d).
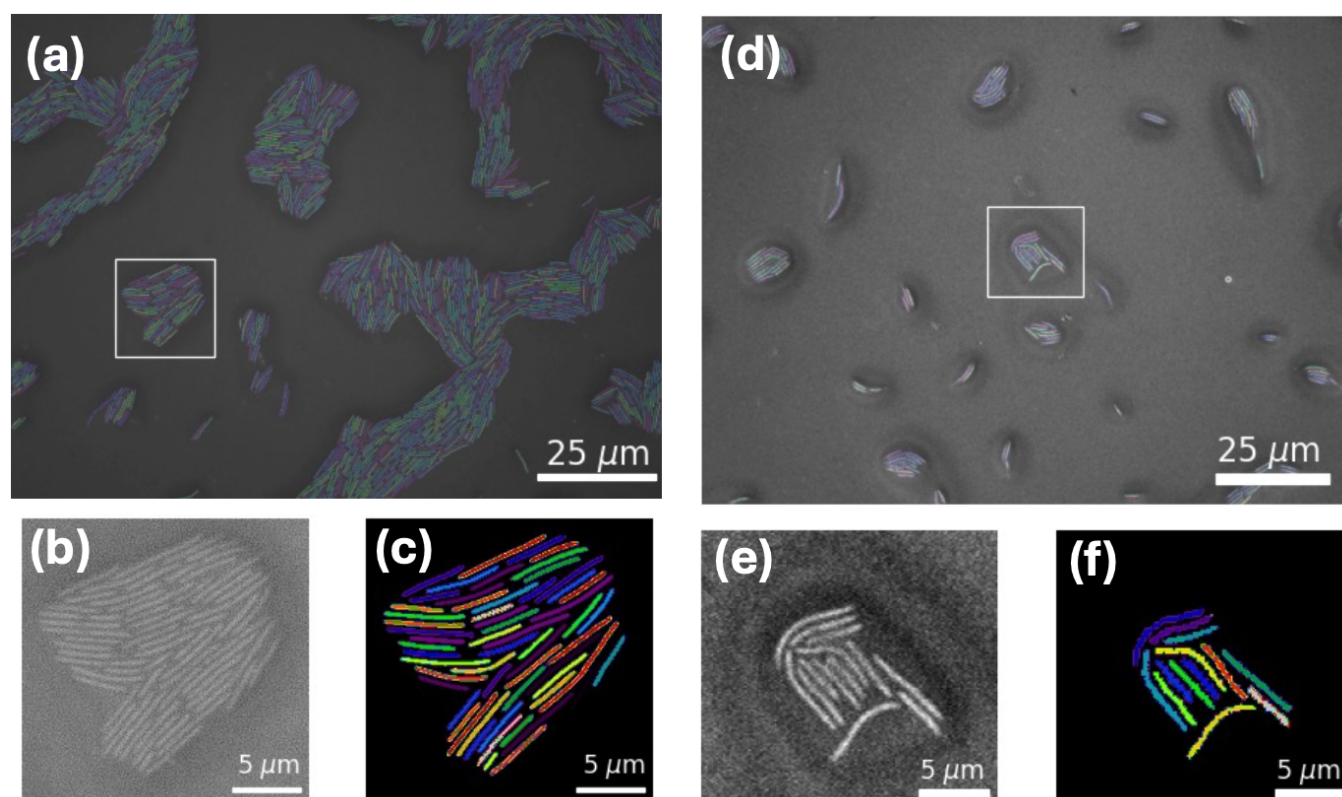
Supplementary Fig. 25. Representative segmentation results for *F. johnsoniae* and *M. xanthus*. **(a)** Heatmap indicated the fraction of all predicted counts for groups of size N (x-axis). **(b)** Linear regressions of the mean (top) and variance (bottom) of predicted counts for groups of size N (x-axis). Note that if the predicted counts were Poisson-distributed, we would expect the slope of the variance vs. ground truth regression to be 1. Thus, we conclude that counts are underdispersed. **(c)** Histograms of predicted cell counts for each collection of ground truth groups of size N (indicated in the top-right corner of each graph) up to groups of size $N = 16$. Red lines are fits to a generalized Poisson distribution [23] with fixed w=-0.94.
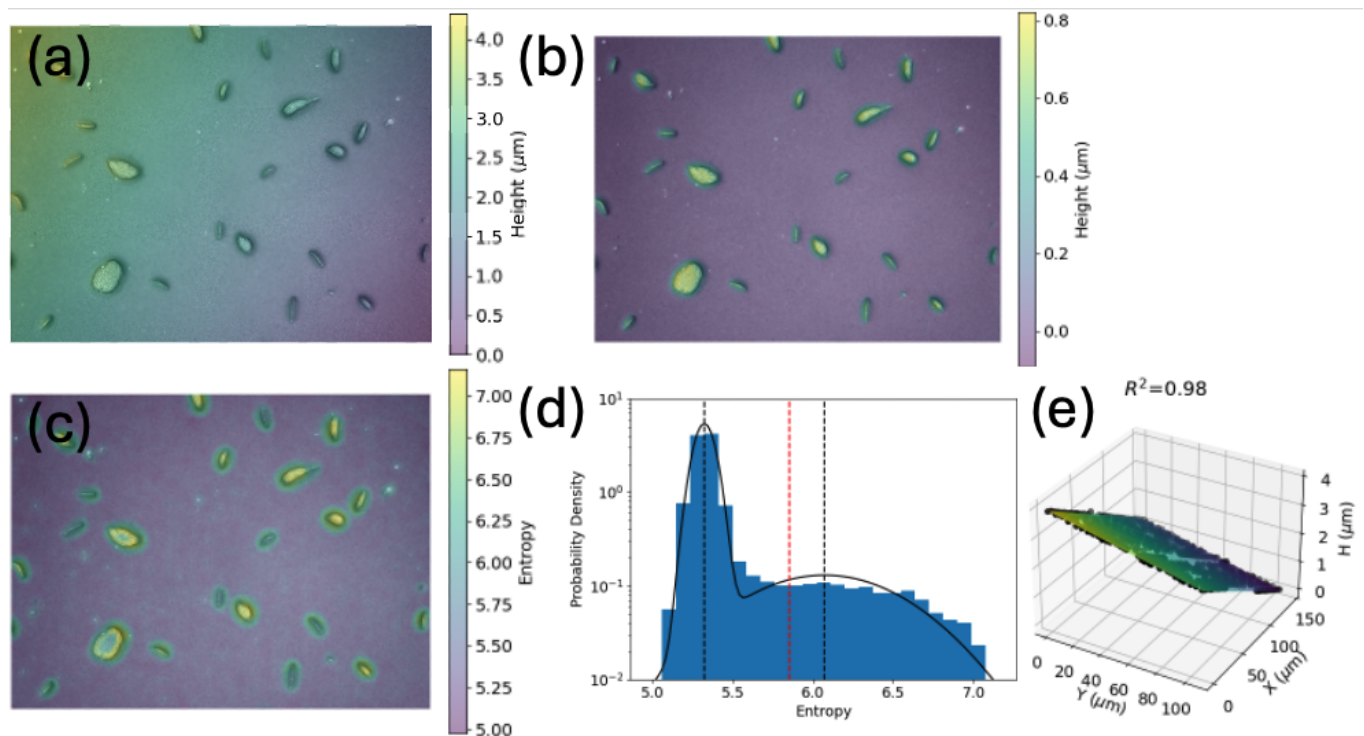
Supplementary Fig. 26. Procedure for correcting Keyence height images for sample tilt and surface unevenness. **(a)** Example uncorrected height height map overlaid on its corresponding reflectance image. Image is of small groups and single cells of *M. xanthus*. **(b)** (a) but with the corrected height map overlaid. **(c)** Output of entropy filtering the reflectance image of (a). **(d)** Histogram of entropy values from (c), with fit of the distribution of entropy values to a 2 component Gaussian mixture model (black, solid line). Black, dashed lines are the fitted means of the 2 components. Dashed red line is the threshold value, below which corresponding pixels are called as part of the "surface." **(e)** Fit (solid surface) of raw height values (black points, subsampled 1:1000 from original data).
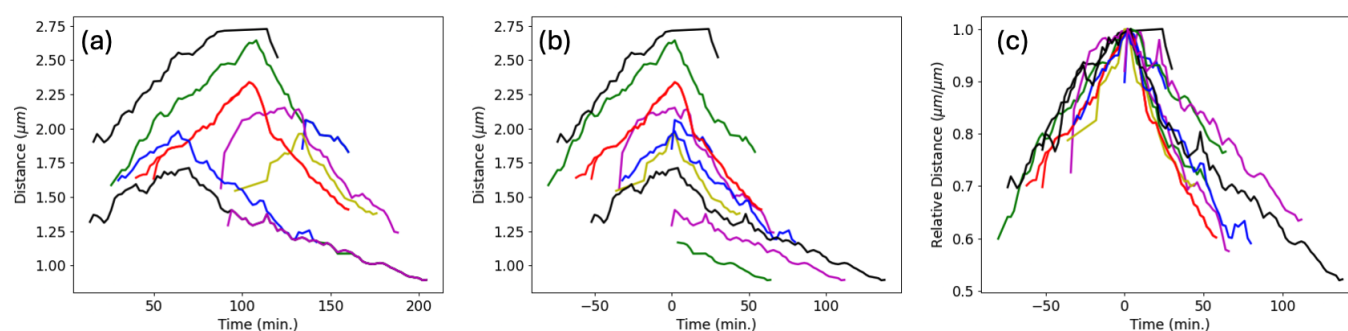
Supplementary Fig. 27. Calculation of relative centroid-to-centroid distance for non-motile cell pairs. **(a)** Smoothed traces of absolute centroid-centroid distance between pairs of cells. Each line represents a single pair of daughter cells. **(b)** Traces are aligned such that time $t = 0$ min. corresponds to the time at which the centroids are maximally-separated. **(c)** Relative centroid-centroid distance is calculated by dividing the distances in (b) by their value at $t = 0$ min.. The curve shown in Fig. 2i was computed by averaging, at each time, the data shown here.

SUPPLEMENTARY VIDEOS

Supplementary Video 1. Movie for the time series shown in Fig. 2d, showing the merger of non-motile *M. xanthus* cells due to capillary attraction.

Supplementary Video 2. Movie showing merging and splitting of motile *M. xanthus* groups. Colored regions are groups of cells marked as a single groups. Overlaid numbers are the detected number of cells in that group.

Supplementary Video 3. Movie showing an example of simulated cells in the "gas" phase. Simulation parameters are $\tilde{\ell} = 10^{3.5}$, $\tilde{F} = 10$, and $\tilde{\tau} = 10$. Color bar is the same as in main Fig. 3.

Supplementary Video 4. Movie showing an example of simulated cells in the "droplets" phase. Simulation parameters are $\tilde{\ell} = 10^2$, $\tilde{F} = 0.3$, and $\tilde{\tau} = 3$. Color bar is the same as in main Fig. 3.

Supplementary Video 5. Movie showing an example of simulated cells in the "polar clusters" phase. Simulation parameters are $\tilde{\ell} = 10^{2.5}$, $\tilde{F} = 1$, and $\tilde{\tau} = 10^2$. Color bar is the same as in main Fig. 3.

Supplementary Video 6. Movie showing an example of simulated cells in the "streams" phase. Simulation parameters are $\tilde{\ell} = 10^{2.5}$, $\tilde{F} = 1$, and $\tilde{\tau} = 1$. Color bar is the same as in main Fig. 3.

Supplementary Video 7. Movie showing an example of *M. xanthus* cells in the "gas" phase.

Supplementary Video 8. Movie showing an example of *M. xanthus* cells in the "droplets" phase.

Supplementary Video 9. Movie showing an example of *M. xanthus* cells in the "polar clusters" phase.

Supplementary Video 10. Movie showing an example of *M. xanthus* cells in the "streams" phase.

Supplementary Video 11. Movie for the time series in Fig. 3f, showing neighbor exchange in the "stream" phase.

Supplementary Video 12. Movie for the time series shown in Fig. 4a, showing the large-scale morphologies of *M. xanthus* cells upon changes in water availability.

Supplementary Video 13. Movie showing the large-scale morphologies of *F. johnsoniae* cells upon changes in water availability.

REFERENCES

[1] Flory, P. J. & Rehner Jr, J. Statistical mechanics of cross-linked polymer networks I. Rubberlike elasticity. *J. Chem. Phys.* **11**, 512–520 (1943).

[2] Doi, M. *Soft matter physics* (Oxford University Press, 2013).

[3] Mao, B., Divoux, T. & Snabre, P. Impact of saccharides on the drying kinetics of agarose gels measured by in-situ interferometry. *Sci. Rep.* **7**, 1–9 (2017).

[4] Beroz, F. *et al.* Verticalization of bacterial biofilms. *Nature physics* **14**, 954–960 (2018).

[5] Pécréaux, J., Döbereiner, H.-G., Prost, J., Joanny, J.-F. & Bassereau, P. Refined contour analysis of giant unilamellar vesicles. *Eur. Phys. J. E* **13**, 277–290 (2004).

[6] Harshey, R. M. & Matsuyama, T. Dimorphic transition in *Escherichia coli* and *Salmonella typhimurium*: surface-induced differentiation into hyperflagellate swarmer cells. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 8631–8635 (1994).

[7] Wilson, D. L., Baddeley, A. J. & Owens, R. A. A new metric for grey-scale image comparison. *Int. J. Comput. Vis.* **24**, 5–17 (1997).

[8] Natterer, F. *The mathematics of computerized tomography* (SIAM, 2001).

[9] Camley, B. A. *et al.* Polarity mechanisms such as contact inhibition of locomotion regulate persistent rotational motion of mammalian cells on micropatterns. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 14770–14775 (2014).

[10] Xiong, L. *et al.* Flower-like patterns in multi-species bacterial colonies. *eLife* **9**, e48885 (2020).

[11] Qin, B. *et al.* Cell position fates and collective fountain flow in bacterial biofilms revealed by light-sheet microscopy. *Science* **369**, 71–77 (2020).

[12] Huang, E. *et al.* OpenDrop: Open-source software for pendant drop tensiometry & contact angle measurements. *J. Open Source Softw.* **6**, 2604 (2021).

[13] Van Rossum, G. & Drake Jr, F. L. *Python tutorial* (Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands, 1995).

[14] Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020). URL `https://doi.org/10.1038/s41586-020-2649-2`.

[15] Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **17**, 261–272 (2020).

[16] van der Walt, S. *et al.* scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014). URL `https://doi.org/10.7717/peerj.453`.

[17] Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* **9**, 90–95 (2007).

[18] Kuglin, C. & Hines, D. The phase correlation image alignment method. *Proc. IEEE Int. Conf. Cybernet. Soc.* 163–165 (1975).

[19] Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, 234–241 (Springer, 2015).

[20] He, K., Zhang, X., Ren, S. & Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, 1026–1034 (2015).

[21] Paszke, A. *et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems 32*, 8024–8035 (Curran Associates, Inc., 2019). URL `http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf`.

[22] Guerrero-Pena, F. A. *et al.* Multiclass weighted loss for instance segmentation of cluttered cells. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2451–2455 (IEEE, 2018).

[23] Consul, P. C. & Jain, G. C. A generalization of the Poisson distribution. *Technometrics* **15**, 791–799 (1973).

[24] Consul, P. & Jain, G. On some interesting properties of the generalized Poisson distribution. *Biom. Z.* **15**, 495–500 (1973).

[25] Kaiser, D. Social gliding is correlated with the presence of pili in *Myxococcus xanthus*. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 5952–5956 (1979).

[26] Wu, S. S. & Kaiser, D. Markerless deletions of *pil* genes in *Myxococcus xanthus* generated by counterselection with the *Bacillus subtilis sacB* gene. *J. Bacteriol.* **178**, 5817–5821 (1996).

[27] Keilberg, D., Wuichet, K., Drescher, F. & Søgaard-Andersen, L. A response regulator interfaces between the Frz chemosensory system and the MglA/MglB GTPase/GAP module to regulate polarity in *Myxococcus xanthus*. *PLoS Genet.* **8** (2012).

[28] Islam, S. T. *et al.* Unmasking of the von Willebrand A-domain surface adhesin CglB at bacterial focal adhesions mediates myxobacterial gliding motility. *Sci. Adv.* **9**, eabq0619 (2023).

[29] Rhodes, R. G., Pucker, H. G. & McBride, M. J. Development and use of a gene deletion strategy for *Flavobacterium johnsoniae* to identify the redundant gliding motility genes remF, remG, remH, and remI. *J. Bacteriol.* **193**, 2418–2428 (2011).

[30] Keseler, I. M. *et al.* The EcoCyc database in 2021. *Front. Microbiol.* **12**, 711077 (2021).