Active density fluctuations in bacterial binary mixtures

Silvia Espada Burriel[1], Remy Colin[1,*]

[1] Max Planck Institute for Terrestrial Microbiology and Center for Synthetic Microbiology (SYNMIKRO), Karl-von-Frisch-strasse 10, 35043 Marburg, Germany
[*] remy.colin@synmikro.mpi-marburg.mpg.de

May 10, 2023

Abstract

In wild environments, intermixed motile and sessile microorganisms interact physically and chemically to give rise to complex community organization. Here, we investigate the role of physical interactions in these communities by studying the physical behavior of binary mixtures of motile and non-motile *Escherichia coli* bacteria. We systematically explored the phase diagram of the mixtures in microfluidics and microscopy experiments, tuning multiple parameters by biological and physical means. We report a novel phenomenon by which large density fluctuations of non-motile bacteria emerge when mixed with motile ones, in physiologically relevant conditions. Experimental results together with quantitative modeling and numerical simulations show that circular swimming of motile cells at surfaces generate recirculation flows that advect the non-motile cells through hydrodynamic interactions, and that sedimentation, by breaking the vertical symmetry of the system, is essential for local non-motile cell accumulation and the emergence of the large-scale density fluctuations. This behavior represents a new type of non-equilibrium self-organization in swimming bacterial populations, which is distinct from previously observed dry active matter structuration phenomena such as motility-induced phase separation, and appears crucial for complex microbial community organization.

Main

Many bacterial species, such as the model organism *Escherichia coli*, are motile. They rotate helical flagella to propel themselves forward in the medium at a few tens of microns per second. At the single cell level, the bacterium alternates seconds-long periods of straight runs (modulo rotational Brownian motion) with short reorientations, the tumbles, to describe a random walk that allows each cell to explore the environment [1]. Moreover, a sensory system, the chemotaxis pathway, modulates the tumbling frequency according to perceived changes in physicochemical environmental conditions (nutrients, pH, temperature, etc.) to bias the cell motion towards "preferred" conditions and thus navigate the environment [2–6].

The constant mechanical energy input due to swimming drives bacterial suspensions out of equilibrium. The bacteria can be thought of as force dipoles in a low Reynolds number fluid [7] which interact hydrodynamically with each other and with surfaces via the fluid they displace, on top of steric interactions upon direct collisions [8]. At single cell level, swimming bacteria tend to accumulate near surfaces due to both types of interactions [9–13]. The proximity to the surface also alters their swimming behavior, making them swim in circular trajectories [14–16]. At the population level, bacterial suspensions are active fluids in which a host of self-organization behaviors may take place [17]. The long-ranged hydrodynamic interactions between the rod-shaped bacteria, together with steric interactions [18], induce polar ordering which explain the emergence of swirling collective motion known as bacterial turbulence at high cell density (volume fraction $\phi > 1\,\%$) [19–23]. In contrast, if the cell speed decreases when density gets high, a motility induced phase separation (MIPS) can take place, where the active fluid phase-separate in slow/dense and fast/sparse regions [24–26]. This can arise because of steric interactions in absence of polar alignment in the case of active Brownian particles [25,27], but also in (polar) bacteria from density-induced biological downregulation of motility [28]. Moreover, motility combined with chemotaxis towards quorum-sensing chemicals released by other cells can produce density patterns [29], ranging in bacteria from autoaggregation on a 100 µm scale [30] to both static and dynamic density structures on a millimeter to centimeter scale [31–34].

These studies have focused on homogeneous systems, i.e. populations of phenotypically identical cells. However, natural communities are composed of many different species with diverse phenotypes, and they are observed to organize in a wide variety of complex spatial structures [35–37]. This complex organization is thought

to be a prominent determinant for the coexistence, ecology and evolution of multi-species communities[36–38]. Microbiologists have highlighted the importance of biochemical interactions between species, both cooperative (e.g. cross-feeding) and antagonistic (e.g. competition for resources or cross-killing), in the emergence of these spatial structures, especially when the colony forms a sessile (non-motile) surface attached biofilm[39–42].

However physical interactions between phenotypically different species might also matter strongly, particularly where active motility is involved. Indeed, motile cells are known to enhance the diffusion of passive particles[43–45], notably affecting their sedimentation[46]. Passive Brownian colloids may also experience effective attractive interactions when in suspension with motile cells[47–49]. Moreover, numerical simulations of mixtures of active and passive Brownian particles have shown the emergence of spatial co-organization patterns in two dimensions, both for disks[50–55] and rods[56]. The passive particles form dynamic clusters driven by the activity of the motile particles, which co-segregate, in a mechanism reminiscent of MIPS[57]. Active particles with different mobilities also seem to demix[58,59]. When passive particles are coupled to an active nematic, which commonly models dense active actin filament suspensions rather than bacteria[60], a fluctuation dominated phase ordering (FDPO) has been observed to emerge in simulations, where density fluctuations of the passive particles emerge as they surf on the two-dimensional active nematic[61–63]. Experimentally, few studies have investigated physical self-organization in bacterial mixtures, although it has been observed when motility is coupled to biochemical interactions, e.g. when two strains cross-regulate each other's motility[64] or compete for resources[65], when a non-motile strain recruit a motile one in its biofilm via chemotaxis[66] or when a non-motile strain hitchhike with a motile one[67].

In this paper, we systematically investigated the effect of motility on the spatial organization of binary mixtures of bacteria, where one component is motile and the other not. We observe that large fluctuations of the non-motile cell density emerge, under the action of the motile cells, while the motile cell distribution remained homogeneous. Systematically varying the density of both motile and non-motile cells, we found that the density fluctuations happened in a wide, biologically relevant range. We tuned the swimming pattern of the motile cells by genetic modifications and found that enhanced surface localization and highly spatiotemporally organized swimming patterns at the surface enhance the density fluctuations. We further found that non-motile cell sedimentation is an essential component of the emergence of the fluctuations. The density patterns are dynamic, forming and dissolving on the time scale of minutes. We characterized the collective dynamics of the non-motile cells, which highlighted the importance of vertical recirculation and of hydrodynamic interactions between cells in the emergence of the density patterns. Our observations can be recapitulated in a mechanism that is supported by numerical simulations: Surface-localized swimmers generate three-dimensional recirculation flows leading to convective fluxes of non-motile cells. The latter fluxes combine with sedimentation, which breaks the top-down symmetry of the system and thus allows for local cell accumulation, hence giving rise to the density fluctuations.

Emergence of large density fluctuations in bacterial binary mixtures

We study mixtures of motile and non-motile *E. coli*, which are fluorescently marked with respectively mCherry and mNeonGreen, in a microfabricated chamber of controlled height, $h = 50\,\mu$m (Methods), systematically varying the cell densities of both components of the binary mixture. All strains are stripped of their main surface adhesin Antigen 43 to prevent cell aggregation[30]. To keep biochemical interactions as trivial as possible, the strains are genetically identical, except for the non-motile strain being a knockout of the gene *fliC*, which encodes for the flagellar filament subunit. A passive system composed of two non-motile strains tagged with mCherry and mNeonGreen is used as a control.

In some binary mixtures, we observed large spatial fluctuations of the density of the non-motile cells in presence of motile cells (Fig. 1a), whereas the corresponding passive system is homogeneous (Fig. 1b). These large density fluctuations are characterized by the existence at steady state of dynamic patches of abnormally high and low densities of non-motile cells, which reorganize in a matter of minutes (Extended Movie 1). We quantified these density fluctuations using fluorescence intensity as a proxy for cell density (Methods). The spatial structure of the density of non-motile cells shows characteristics of giant fluctuations (Extended Figs. S1b and S1d). Correspondingly, the spatial autocorrelation of the non-motile fluorescence intensity shows a non-exponential decay at large length scales, $(10 - 50)\,\mu$m, well above the $\sim 2\,\mu$m cell length (Extended Fig. S1e). We used the normalized spatial variance of the cell mass on this typical length scale, called the heterogeneity level further on, as a quantification of the amplitude of the density fluctuations (Methods).

During the 30 minutes after a freshly stirred mixture is introduced in the observation chamber ($t = 0$), the heterogeneity level increases until it reaches a steady state, this dynamic being accelerated when more motile cells are present (Extended Movie 2 and Extended Fig. S2). We constructed a full phase diagram of the amplitude of the fluctuations at steady state (i.e. for $t \simeq 1$ hour) as a function of the volume fractions of motile (ϕ_M) and non-motile (ϕ_{NM}) cells (Fig. 1c). The density fluctuations are observed only when non-

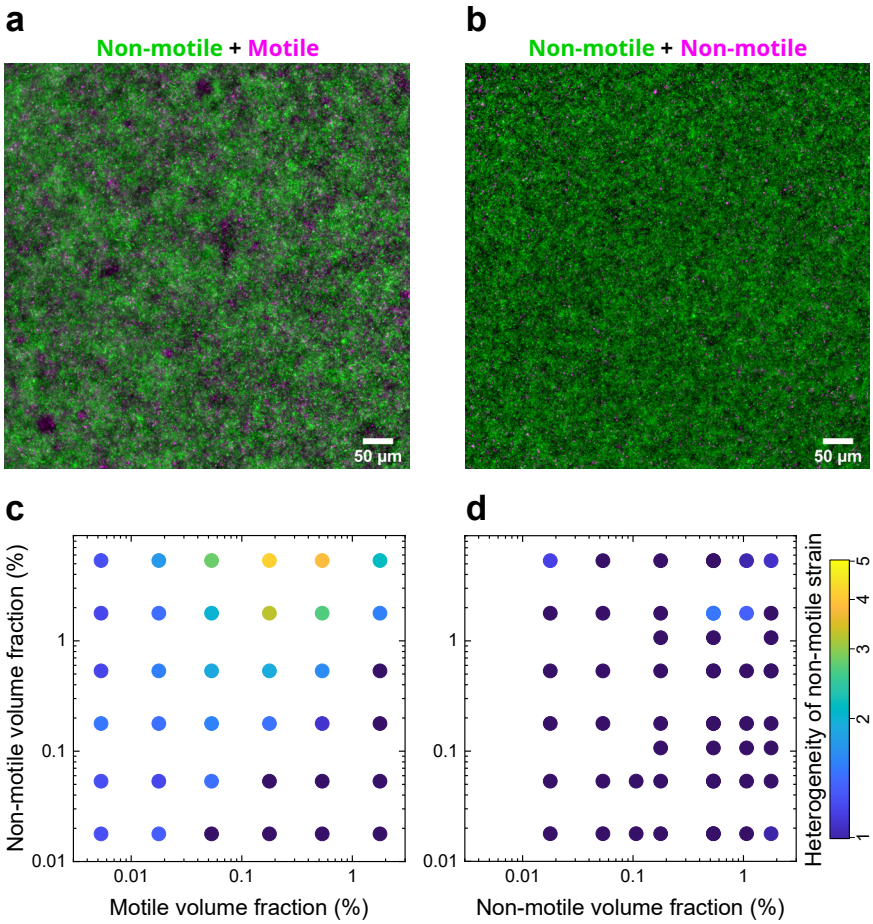

Figure 1: Density fluctuations in bacterial binary mixtures. a) Microscopy images of bacterial mixtures of motile strain (mCherry) and non-motile strain (mNeonGreen) of *E.coli* (motile cells volume fraction $\phi_M = 0.18\%$, non-motile $\phi_{NM} = 5.3\%$). The density of the non-motile strain exhibits strong spatial fluctuations in presence of the motile strain. b) The corresponding control where two non-motile strains of *E.coli* were mixed at the same volume fractions. c-d) Phase diagrams of the density fluctuations of the non-motile strain in binary mixtures of bacterial populations depending on the volume fraction of both strains. The density fluctuations are quantified as described in the Methods section by a heterogeneity level, which is represented by a color scale. Each point represents the geometric mean of at least three biological replicates. c) Mixture of a motile strain (abscissa) with a non-motile strain (ordinate). d) Mixture of two non-motile strains as a passive control system.

motile cells are in a majority ($\phi_{NM} > \phi_M$) and they reach a maximum for a volume fraction of the non-motile strain $\phi_{NM} \simeq 5\,\%$ and a volume fraction of the motile strain $\phi_M \simeq 0.2\,\%$ (Fig. 1c). The range of densities over which the fluctuations are observed corresponds to typical planktonic culture densities for the motile cells ($0.01 < \phi_M < 1\,\%$) and typical biofilm densities for the non-motile ones ($0.1 < \phi_{NM} < 10\,\%$), suggesting that this phenomenon is highly relevant to natural bacterial communities. Moreover, the heterogeneity level of the non-motile strain decreases at high volume fraction of motile cells ($\phi_M \geq 1\,\%$), where collective motion begins to occur[19,21,22]. Interestingly, the spatial distribution of the motile strain was always homogeneous (Extended Fig. S3), and the density fluctuations only appear in the non-motile strain when the motile strain was present (Fig. 1b, d).

Therefore, the swimming activity of the motile cells drives the emergence of these out-of-equilibrium density fluctuations for the non-motile cells, without them being affected in any obvious way. Moreover, no change in the swimming speed was noticed. This is contrary to the behavior in simulated MIPS-like phase separation of mixtures[24,25,52–57], suggesting that a different mechanism is at play. To gain more insight into this mechanism, we systematically explored possible control parameters of the behavior.

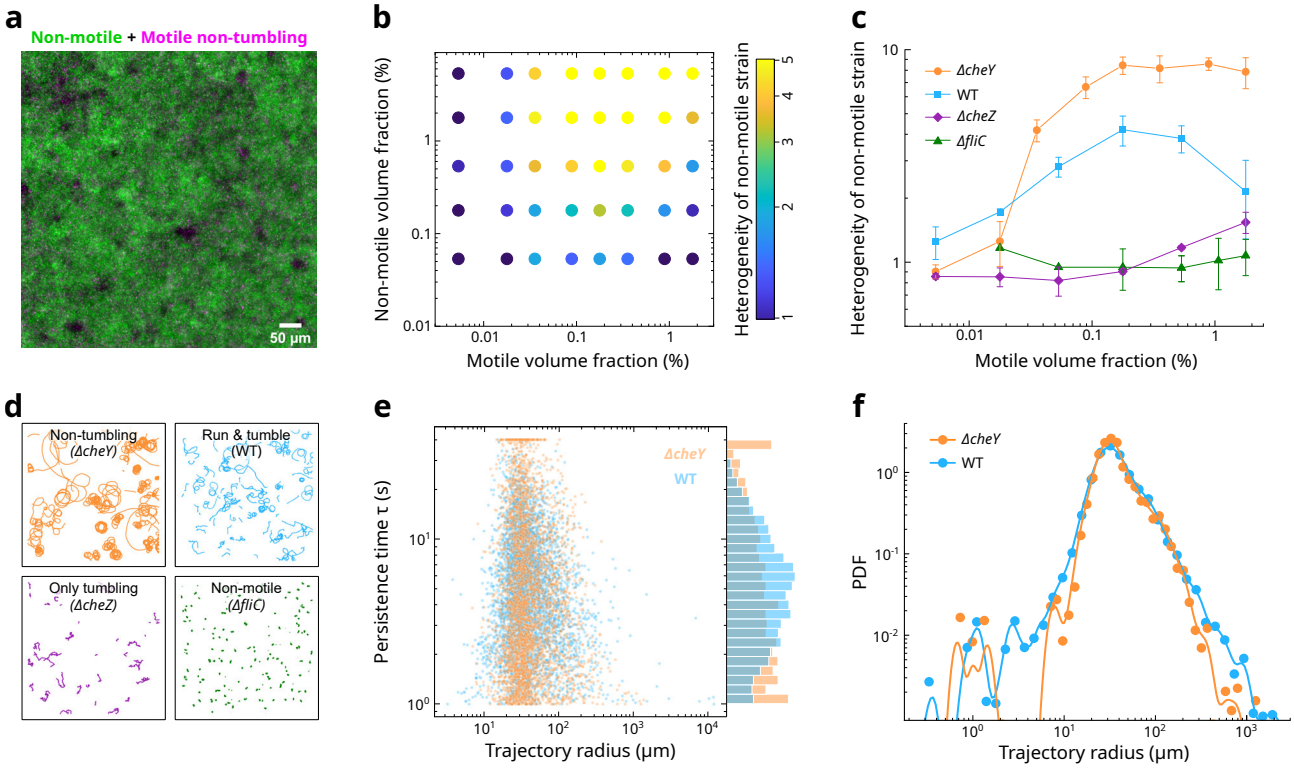

Figure 2: Tumbling affects the density fluctuations. a) Microscopy image of *E.coli* mixtures of a motile non-tumbling strain ($\Delta cheY$ mCherry) and a non-motile strain (mNeonGreen) where the non-motile strain presents large density fluctuations in the presence of the motile strain. b) Phase diagram of the density fluctuations in mixtures of a non-tumbling motile strain ($\Delta cheY$) (volume fraction in abscissa) with a non-motile strain (volume fraction in ordinate). Density fluctuations are quantified as described in Methods by the heterogeneity level, which is represented with the same color scale as Fig. 1c,d. c) Heterogeneity levels of the non-motile cells in a binary mixture with different motile strains as indicated, as a function of the volume fraction of the motile strain. The volume fraction of the non-motile strain is kept constant ($\phi_{NM} = 5.3\%$). (b,c) Each point represents the geometric mean of three replicates. (c) Error bars indicate the standard deviation. d) Swimming trajectories of the different strains used as the motile strain in the binary mixtures. e) Scatter plot of the persistence time τ of individual circular swimming trajectories of WT and $\Delta cheY$ motile cells, which were tracked at the bottom of an experimental chamber, as a function of the radius of curvature of the trajectory, and probability density distribution of τ (right side). f) Probability density distribution of the radii of curvature of the same trajectories as (e). Each trajectory is weighted by its length. Lines represent the kernel density estimate of the data. (e-f) The dataset comprises 4077 ($\Delta cheY$) and 6444 (WT) individual swimming trajectories in 8 biological replicates.

Tumbling, but not chemotaxis, affects the density fluctuations via the persistent circular swimming of surface localized motile cells

Since chemotaxis is known to be involved in the formation of some bacterial density patterns[30–34,66,67], we investigated whether this is the case in our system. We targeted the diffusible response regulator CheY, which plays the essential role of transducing the sensory signal from the chemoreceptors, which monitor environmental conditions, to the flagellar motor within the cytoplasm. There, CheY binds and triggers tumbles when phosphorylated, thus modulating swimming behavior[2–4]. We first used a CheY knock-out ($\Delta cheY$) motile strain, which is non-tumbling and non-chemotactic. To our surprise, the density fluctuations of the non-motile strain were found to be stronger and to appear on a wider range of cell densities in a mixture with the $\Delta cheY$ motile strain (Fig. 2a, b, c) than when the motile strain is wild-type (WT) for chemotaxis (Fig. 1c). Although the typical size of the density patterns is largely unaffected (Extended Figs. S1d,e), the local extrema of density are much more accentuated, with numerous depletion zones forming next to very dense patches (Fig. 2a, Extended Movie 3). However, a heterogeneity level similar to the WT strain case was found across the whole phase diagram for the motile $cheY^{**}$ mutant (Extended Fig. S4a), for which a double point mutation (D13K, Y106W) in the $cheY$ gene allows WT level of tumbling at proper expression level, but prevents chemotactic signaling[68]. This indicates that chemotaxis *per se* is not important for the emergence of the density fluctuations, but that the ability to tumble is a control parameter of the phenomenon.

We tested whether the tumbles affect the surface localization of motile cells[13] by measuring their vertical distribution via confocal microscopy. Motile cells localize preferentially near surfaces as expected[10,11], and the $\Delta cheY$ strain indeed shows an enhanced localization near the top and bottom surfaces compared to WT, which is however relatively moderate (Extended Fig. S4b). Close to the surfaces, the flagellated bacteria were observed to swim in clockwise circles when observed from above (Fig. 2d) as previously reported, because of the hydrodynamic torque induced by the no-slip wall[14]. In the $\Delta cheY$ strain, these circles are perturbed only by rotational diffusion (Fig. 2d). They thus tend to persist in place much longer than in the WT strain, where tumbles interrupt them. We indeed quantified the time needed for a cell to "break the circle" by measuring a circle persistence time defined as the typical time τ each swimming cell takes to move a linear distance $dr(\tau) > 3R$, with R the mean radius of curvature of the trajectory. The distribution of τ is strongly shifted to higher τ for the $\Delta cheY$ strain compared to WT, with a large fraction of $\Delta cheY$ cells staying in the $3R$ vicinity for the whole duration (40 s) of our cell tracking movies (Fig. 2e). This higher persistence corresponds to a higher heterogeneity level of the non-motile cells when mixed with $\Delta cheY$ (Fig. 2c). Conversely, we increased tumbling rate well above WT level by using a tumbling-only motile strain, $\Delta cheZ$. This strain's flagella rotate, but only clockwise, causing constant tumbling and Brownian-like random walk (Fig. 2d), while still pouring energy into the fluid. The heterogeneity level of the non-motile strain of the mixture drops to almost control levels for all tumbling-only volume fractions (Fig. 2c). This suggests that the ability to swim persistently in circles at the surface is key for the emergence of the density fluctuations.

The distribution of radius of curvature of the circular trajectories was very broad and similar for the WT and non-tumbling strains (Fig. 2f). This broadness is not unexpected since the radius of curvature of the trajectory depends strongly on the variable distance between the cell and the surface[14,15]. Since cell body length should also affect the radii of curvature[14], we altered their distribution by elongating both WT and $\Delta cheY$ motile cells via a short treatment with the antibiotic cephalexin[21]. Elongated cells trajectories show an average increase of the radii of curvature, but their distribution remains broad, presumably due to the effect of the distance to the surface (Extended Fig. S5a-d). In particular, the abundance of very curved trajectories (small radii) is similar in the elongated and untreated cells. When mixed with elongated motile cells, the heterogeneity level of the non-motile cells is unchanged for $\Delta cheY$ and slightly increased for WT motile cells (Extended Fig. S5e). Taken together, these results suggest that the persistent and strongly curved swimming trajectories contribute to the formation of the density fluctuations.

Sedimentation of non-motile cells is essential for the density fluctuations

We next considered the vertical distribution of non-motile cells, which is similar whether the motile strain of the mixture is WT or $\Delta cheY$ (Extended Fig. S4c). In both conditions, non-motile cells localize towards the bottom of the sample because of sedimentation, their volumetric mass density ($\rho_C = 1.1$ g.cm^{-3}) being higher than the one of the aqueous medium ($\rho_w = 1.0$ g.cm^{-3}). The (Boltzmann) steady state vertical distribution of sedimenting particles results from a balance between sedimentation and diffusion. The latter is enhanced as the volume fraction of motile cells increases (Extended Fig. S6a,b), as previously observed[43,44,46], and the vertical distribution of non-motile cells therefore becomes more homogeneous in presence of an increasing density of motile cells (Extended Fig. S6c,d).

Inspired by previous works on sedimentation of high-density colloidal suspensions[69,70], we decided to in-

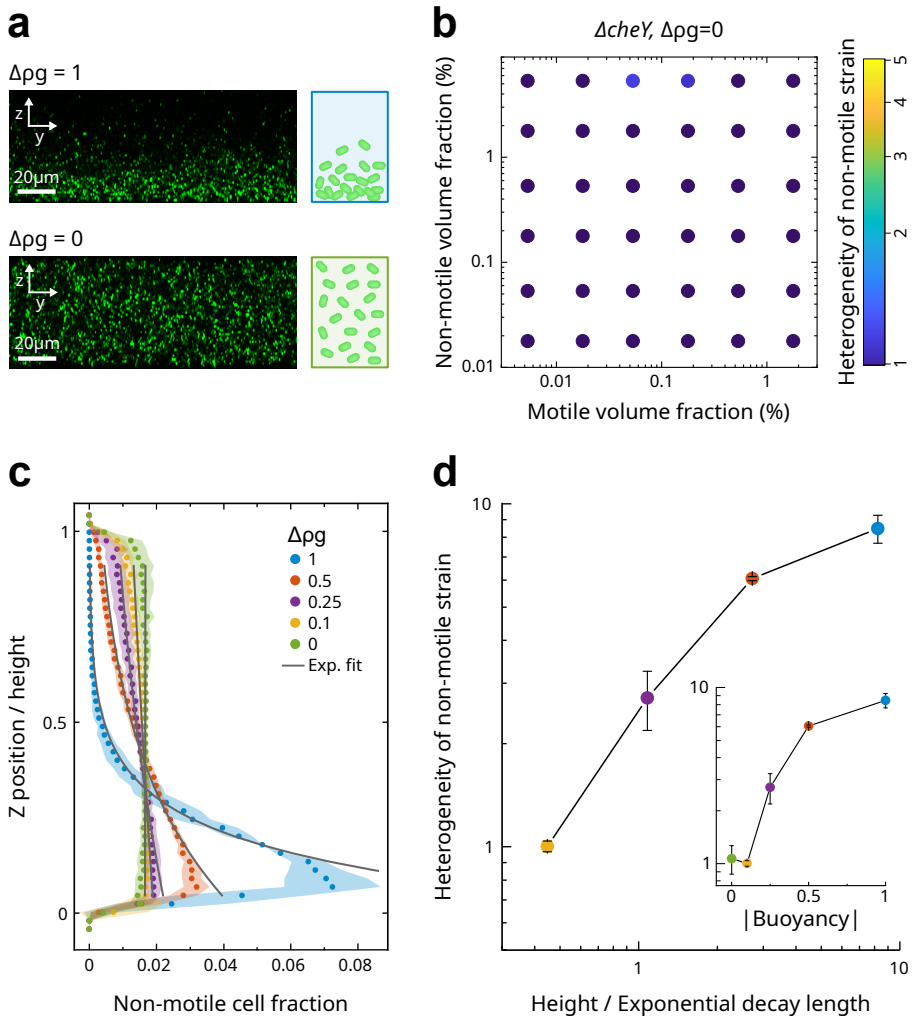

Figure 3: Sedimentation controls the density fluctuations. a) Confocal images of the vertical distribution of non-motile cells at the two extreme buoyancies $\Delta\rho g = 1$ and $\Delta\rho g = 0$. b) Phase diagram of the density fluctuations in a mixture of $\Delta cheY$ motile strain (volume fraction in abscissa) with a non-motile strain (volume fraction in ordinate) in density-matched suspending media ($\Delta\rho g = 0$). The density fluctuations are quantified by the heterogeneity level (Methods) which is represented by a color scale. c-d) Dose-dependence effect of buoyancy on the sedimentation (c) and the density fluctuations (d) of the most heterogeneous binary mixture ($\phi_M = 0.18\%$, $\phi_{NM} = 5.3\%$, motile strain $\Delta cheY$). c) Non-motile cell fraction (abscissa), measured by cell counting in confocal microscopy, as a function of the vertical position (z) in a channel of 50 μm height (ordinate), for suspensions at different buoyancies (the buoyancy level $\Delta\rho g$ is normalized for motility buffer ($\Delta\rho g = 1$)). The grey lines are exponential fits of the cell fraction as a function of z. d) Density fluctuations of the non-motile strain as a function (main figure) of the inverse exponential decay length of the sedimentation profile and (inset) of the buoyancy normalized for motility buffer ($\Delta\rho g$). a-d) Each point represents the geometric mean of three replicates. Error areas (c) and bars (d) indicate the standard deviation.

vestigate how sedimentation affects the density fluctuations in our system. We matched the volumetric mass density of the suspending medium to the one of the bacteria by adding iodixanol, a non-ionic, non-cytotoxic iodine compound that is commonly used as a density matching agent and density gradient medium. It also has little effect on the suspension's viscosity and is metabolically inert to *E. coli*. Coherently, adding iodixanol has little to no effect on the swimming speed of motile cells (Extended Fig. S7a,b) and on the diffusion of non-motile cells (Extended Fig. S7c). In the neutrally buoyant medium ($\Delta\rho g = 0$), the vertical distribution of non-motile cells becomes homogeneous, and thus top-down symmetric (Fig. 3a and Extended Fig. S7d), contrary to the asymmetric distribution in the negatively buoyant medium ($\Delta\rho g = 1$, in normalized units).

In absence of sedimentation, the density fluctuations of the non-motile cells are completely annihilated, whether the motile strain in the mixture is tumbling (WT, Extended Fig. S7e) or not ($\Delta cheY$, Fig. 3b). The level of heterogeneity is indeed equal to the one in the non-motile-only control (Fig. 1d). Focussing on one of the most heterogeneous mixtures when $\Delta\rho g = 1$ (motile $\Delta cheY$ with $\phi_M = 0.18\%$ and $\phi_{NM} = 5.3\%$), we

investigated the effect of varying the difference of volumetric mass density between the cells and the medium, i.e. the sedimentation speed, and therefore the slope of the vertical cell density profile (Fig. 3c). The amplitude of the density fluctuations increased as a function of this density mismatch (Fig. 3d), showing that the cell sedimentation speed is a control parameter of the system.

Furthermore, no matter how high we increased the density of non-motile cells, we never observed any density fluctuations in the $\Delta\rho g = 0$ case (Extended Fig. S7f). We also confined the mixtures in a $20\,\mu$m high channel, for which the half channel height matches the scale of the vertical gradient of non-motile cell density (Extended Fig. S8a). No density fluctuations were observed in this case (Extended Fig. S8b), further highlighting the importance of the three-dimensionality of the system. These results strongly indicate that the effect of sedimentation is not merely due to an increased density of cells near the bottom surface, but that the vertical gradient of cell density resulting from sinking under gravity is essential for the formation of the density fluctuations.

Non-motile cells experience space-correlated flows due to the motile cells

We then studied the dynamics within the density fluctuations at steady state (Extended Movie 1 & 3). The non-motile cells show collective movements with many converging and diverging convective flows, somewhat resembling boiling water. The movements were more pronounced in presence of non-tumbling motile cells, with cell depletion zones ("holes") intermittently opening and closing. The horizontal cell motion was quantified using particle image velocimetry, with the non-motile cells directly serving as tracers (Fig. 4a, b). Even at volume fractions of motile cells far below the emergence of bacterial turbulence, the velocity fields revealed large spatial correlations in the motion of non-motile cells (Fig. 4a). To quantify these correlations, the normalized spatial power spectral density of the velocity field, $E(q) = q|\tilde{v}(q)|^2/\langle v^2\rangle$, was computed as a function of the spatial wave number q (Methods)[20,21]. It represents the fraction of the kinetic energy ($\langle v^2\rangle$) in flow structures of typical size π/q. E(q) shows a pronounced peak, the hallmark of spatial correlation, at a wavelength $\pi/q = 50\,\mu$m, corresponding to the height of the channel (Fig. 4b). This value of the correlation length indicates that hydrodynamic interactions drive these collective movements of the non-motile cells, similarly to what was observed for swirling collective motion of homogenous populations of swimmers[21] and suggests that the motile cells generate structured flow fields that are able to stir the non-motile ones. The temporal correlations of the velocity field decayed rapidly in the tumbling motile case, whereas long term correlations are observed in the non-tumbling one (Extended Fig. S9a), coherently with the longer persistence of circular swimming in the latter case (Fig. 2e).

The level of heterogeneity is however not well predicted by the strength of the collective flows alone, as measured by the height of the peak of $E(q)$ (Extended Fig. S9b). Moreover, a peak of $E(q)$ was also observed in the density-matched suspending medium, which appears at the same length scale and with barely lower amplitude than the one in the corresponding negative buoyancy case (Fig. 4b), indicating quasi-identical spatial flow structures whether $\Delta\rho g = 0$ or $\Delta\rho g = 1$. However, the total amount of kinetic energy ($\langle v^2\rangle$) is somewhat lower in the neutral buoyancy case (Fig. S9c). Hence, regardless of the buoyancy, the motile cells are found to generate spatially correlated movements of the non-motile cells via hydrodynamic interactions, which however only give rise to strong heterogeneities of density in the negative buoyancy case, when the vertical symmetry of the non-motile cell density is broken.

We thus next considered the vertical motion of the non-motile cells. We tracked the three-dimensional motion of a few marked non-motile tracers with confocal microscopy in mixtures with unmarked non-motile and non-tumbling motile cells. We focused on areas where cell depletion zones ("holes") are visible with transmitted light, in the $\Delta\rho g = 1$ case (Fig. 4c), and quantified the velocity of the tracers as a function of the distance to the center of a hole (r) (Fig. 4d,e and Extended Fig. S9d,e). Close to the hole ($r < 10\,\mu$m), the mean vertical velocity of the tracers (V_z) is large and negative, i.e they move down fast on average, while they move on average rapidly away from the hole in the horizontal direction ($V_r > 0$). Further away from the hole ($r \sim 20\,\mu$m), the mean vertical velocity reaches a maximum, and the typical horizontal motion ($|V_{2D}|$) marks a lull. Far from the hole ($r > 30\,\mu$m), the radial velocity relative to the hole (V_r) is zero on average, while the mean vertical velocity is consistent with the sedimentation speed of the non-motile cells ($\simeq 0.05\,\mu$m/s). These findings suggest that the non-motile cells are exposed to a recirculatory flow, being dragged down and expelled away from the hole when close to it.

Recirculative flows generated by circular swimming generate density fluctuations in simulations

We turned to simulations to understand how the fluid flows generated by cells swimming in circle close to the surface allow for density fluctuations to emerge. We used a $N \log N$ algorithm[71,72] to compute fluid flows under

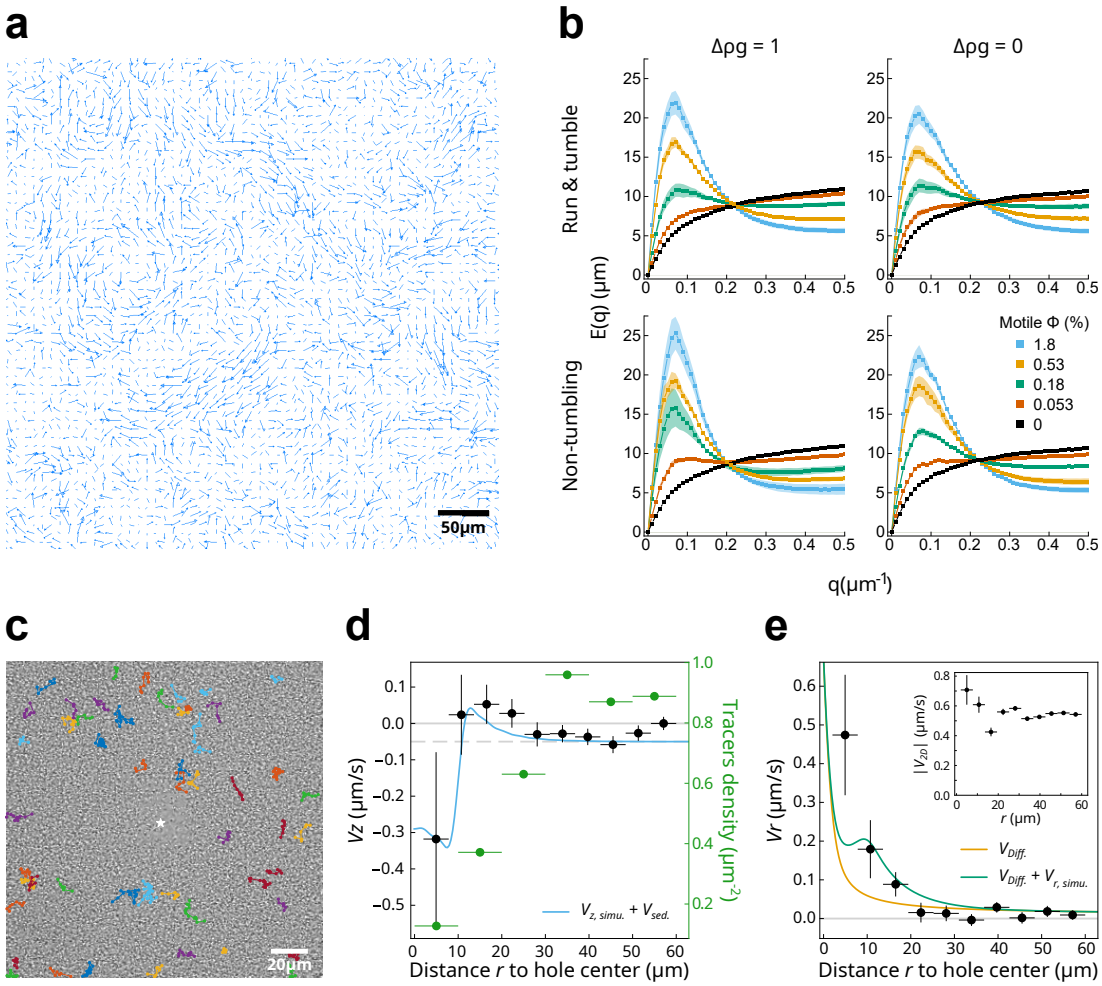

Figure 4: Dynamics of the non-motile cells in the density fluctuations. a) Velocity field of the non-motile strain at $\phi_{NM} = 5.3\%$ in mixture with the motile strain $\Delta cheY$ at $\phi_M = 0.53\%$ and at buoyancy $\Delta\rho g = 1$. b) Flow structure factor E(q) of the non-motile strain flow field for increasing cell densities and for mixtures with the indicated motile strains and buoyancies. c) Field example of the three-dimensional tracking of non-motile cells close to a hole performed in the mixture $\phi_{NM} = 5.3\%$, $\Delta cheY$ $\phi_M = 0.53\%$, at buoyancy $\Delta\rho g = 1$. The white star represents the mean position of the hole center over the whole recording time, and the colored points represent the XY projection of the 3D trajectories of the tracers (tagged non-motile cells). d) Tracer vertical velocity (V_z) (black circles) and their density (green circles) as a function of their distance r to a hole. e) Tracer radial velocity (V_r) (black circles) and horizontal speed ($|V_{2D}| = \sqrt{V_x^2 + V_y^2}$, inset) as function of the distance r to a hole. d-e) The velocities were quantified from more than 600 individual 3D trajectories. Each data point represents the binned mean and the vertical error bars represent the standard error. The continuous lines are simulated flows generated by circular swimmers ($R = 10\,\mu$m, $a^S v_0 = 70\,\mu\text{m}^2/\text{s}$). (d) The vertical velocity sums the flow of the circular trajectories ($V_{Z,simu.}$) and the sedimentation ($V_{sed} = 0.05\,\mu$m) (blue line), (e) and the horizontal velocity result from the diffusive flux ($V_{Diff.}$) alone (yellow line) and with the effect of the circular trajectories ($V_{r,simu.}$) (green line).

confinement by the two no-slip parallel walls at the top and bottom of the simulation domain, with periodic boundary conditions in the other directions (Methods and Supplementary Notes). The swimming cells are approximated as force dipoles, with a dipole length $l_p = 4\,\mu$m, and a reduced force $f_0/6\pi\eta = a^S v_0$, with a^S the swimmer hydrodynamic radius and v_0 the swimming speed. We first computed the normalized flow field generated by a single dipole that describes a perfectly circular trajectory, averaged over one circle revolution, $v_{1r}(r,z)/a^S v_0$ (Fig. 5a). The flow field is axisymmetric around the center of the circle, and features a vertical recirculation, where the fluid is sucked down in the middle of the circle to be expelled out and up on its outside (Fig. 5a and Extended Fig. S10a-f). The curvature of the dipole trajectory sets the strength of the recirculation, with the magnitude of the typical fluid velocity scaling as $\langle|v_{1r}|\rangle_{r,z} \propto 1/R^3$ with the recirculation radius R (Methods and Extended Fig. S10g-i), while the maximal velocity scaled as $\max|v_{1r}| \propto 1/R^2$ (Fig.

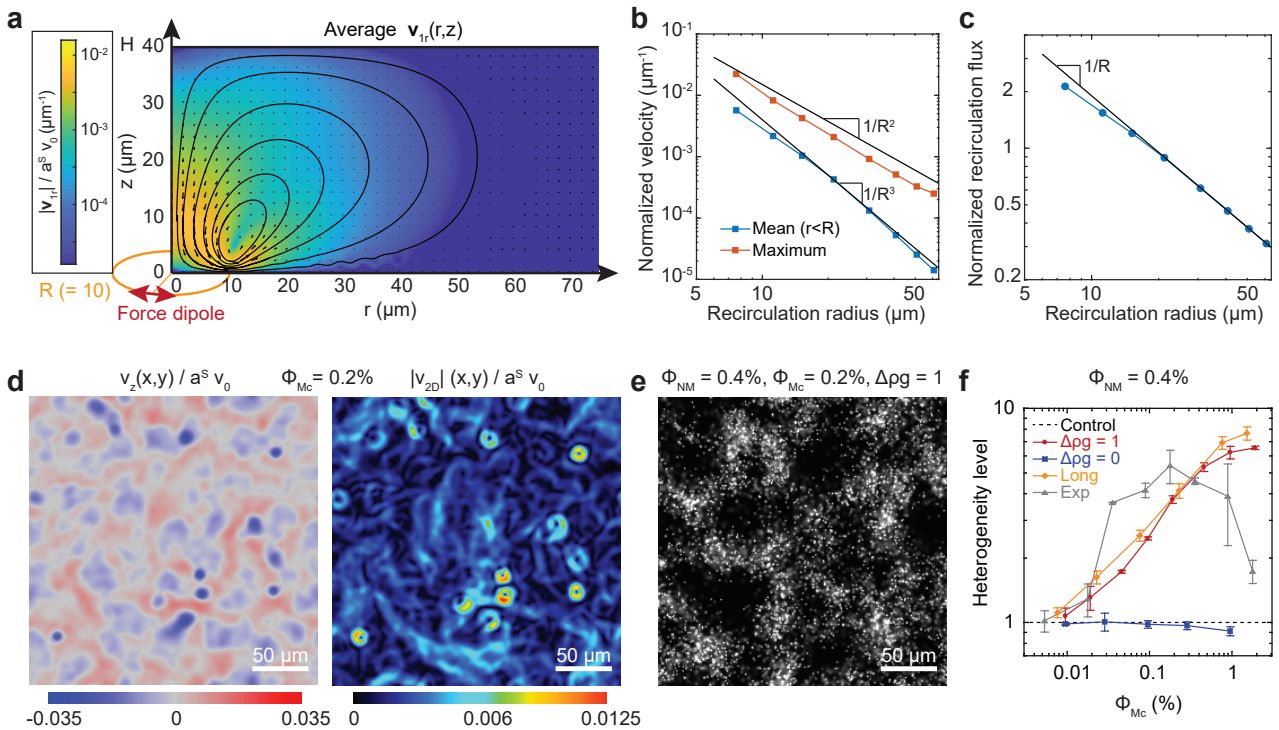

Figure 5: Recirculation flows lead to heterogeneities of non-motile cells in simulations of circular swimmers at surfaces. a, Map of the axisymmetric mean flow field generated by a circularly swimming force dipole. The flow field is computed between two no-slip walls spaced by $H = 40\,\mu$m, and the dipole circulates $z_M = 1\,\mu$m (about one cell body) above the bottom wall. Shown are the flow speed normalized to dipole force $a^S v_0$ (color), the velocity vectors (arrows) and streamlines in the (r,z) plane (solid lines), featuring a recirculation around a fixed point close to the circle radius. b, Mean and maximum of the flow speed inside the circle as a function of the position of the fixed point (recirculation radius). Solid lines indicate power law dependences. c, Normalized recirculation flux, $\max_{\{r\}} \Phi(r)/a^S v_0$ (equation 11, in μm) as a function of the recirculation radius. The solid line indicates $1/R$ dependence. d, Example of mean flow field generated by a population of circular surface swimming dipoles (density $\phi_{Mc} = 0.2\,\%$) for the distribution of circle radii of Fig. 2f. The vertical (v_z) and horizontal ($|v_{2D}| = \sqrt{v_x^2 + v_y^2}$) components are shown in the (x, y) plane, $z = 8\,\mu$m above the bottom wall. Color scales indicate $v/a^S v_0$ in μm^{-1}. e, Reconstructed microscopy image of negatively buoyant diffusive tracers ($v_{sed} = 0.05\,\mu$m/s, $D = 0.3\,\mu\text{m}^2$/s) that followed the mean circular swimmers flow field for 100 seconds, for the same conditions as (d). f, Heterogeneity level of the non-motile tracers, computed in simulations as a function of the volume fraction of motile circular swimmers ϕ_{Mc} for non-motile cells (NM) with volume fraction $\phi_{NM} = 0.4\%$. Control: Homogeneous distribution, $\Delta\rho g = 1$: sedimenting NM, $\Delta\rho g = 1$: non-sedimenting NM, Long: elongated motile cells. Exp: experimental data for non-tumbling motile cells, $\Delta\rho g = 1$ and $\phi_{NM} = 0.53\%$. Error bars indicate standard deviation over 3 replicates.

5b). Consistently, the recirculation flux (equation 11 and Methods), was found to be proportional to $1/R$ (Fig. 5c). The curvature of the dipole trajectory thus creates a net outward force perpendicular to the circle that drives the recirculation.

The shape of the simulated flow curves matched well with our experimental measurements close to depletion zones (Fig. 4d,e and Extended Fig. S10b,d,f). We vertically averaged $v_{1r}(r,z)/a^S v_0$ over the experimental distribution of tracer's z-positions in the 3D-tracking experiments, for a circle radius $R = 10\,\mu$m, which corresponds well to the apparent size of the typical experimentally observed depletion zone. This average simulated velocity field was in good agreement with the experimental velocity field around a depletion zone, when sedimentation and diffusion away from the hole center are also taken into account (Fig. 4d,e). The only fit parameter was the reduced dipole force, $a^S v_0 = 70\mu\text{m}^2$/s, which is 5 times larger than the expected value $a^S v_0 \sim 14\mu$m/s ($a^S = 0.7\,\mu$m and $v_0 = 20\,\mu$m/s as measured for our motile cells). This discrepancy might come from the steep dependence of v_{1r} on the radius R and/or from several circular swimmers contributing to various degrees to the hole formation.

The vertically recirculating flow field might thus perturb the bottom–heavy distribution of passive cells to form depletion zones surrounded by higher densities as observed experimentally. To understand how this

effect translates for thousands of interlaced circular swimmers, we simulated the mean flow field generated by a population of dipoles that swim in circle, with an experimentally-grounded distribution of circle radii (Extended Fig. S11a-b and Methods). The trajectory centers are evenly distributed at both surfaces and independent from each other. Even for large swimmer densities (e.g. $\phi_{Mc} = 0.2\%$, $n_{Mc} = 5\,10^3$ circles, for which the experimental heterogeneity level is maximum), the normalized total flow field $v_{tot}/(a^S v_0)$ shows a few prominent recirculation patterns over a more noisy background (Fig. 5d). Indeed, the typical contribution of a circle of radius R_i being of order $v_{i,1r} \propto 1/R_i^3$, even for very large numbers of swimmers, the relatively few smaller circles in the distribution contribute massively to the total mean flow field and make it spatially heterogeneous (Fig. 5b and Extended Fig. S5f). The flow speed, using $a^S = 0.7\,\mu$m and $v_0 = 20\,\mu$m/s, could be as high as $v_z \sim 0.4\,\mu$m/s in the fast areas, which is now compatible with our experimental measurements (Fig. 4d,e), and suggests that heterogeneities arise from the flow field.

We thus computed the trajectories of passive tracers, which experienced gravity ($v_{sed} = 0.05\,\mu$m/s) and Brownian motion ($D = 0.3\,\mu\text{m}^2$/s) at the level of non-motile cells in $\Delta\rho g = 1$ conditions, in the mean flow field generated by fixed configurations of circular swimmers with $a^S v_0 = 14\,\mu\text{m}^2$/s. For motile and non-motile densities in the experimental range, we observed the formation of strong patterns of density fluctuation that are similar to the experimental ones on reconstructed images of non-motile cells (Fig. 5e). They show giant number fluctuations (Extended Fig. S11c), which we used to define a level of heterogeneity (Methods), comparably to the experiments (Extended Fig. S1). The level of heterogeneity increases as a function of the time τ during which the tracers are exposed to a fixed configuration, corresponding in experiments to the time during which circles would persist (Extended Fig. S11d). For the highest fractions of circular swimmers, we also observed a partial vertical homogenization of the density distribution of non-motile cells (Extended Fig. S11e-f), similarly to experiments (Extended Fig. S6c,d). For a persistence time of $\tau = 100$ s, which is consistent with the experimental non-tumbling swimmer case (Fig. 2e and Extended Fig. S9a), the dependence of the simulated heterogeneity level on circular swimmer density is in fairly good quantitative agreement with experiments (Fig. 5f and Extended Fig. S11g-h), modulo a purely geometric scaling factor (Methods). If gravity is set to zero, no heterogeneities are observed in simulations, in agreement with experiments again, because non-motile cells are homogeneously distributed and the flows are incompressible (Fig. 5f). Normal and elongated circular swimmers also elicit equal levels of non-motile heterogeneities (Fig. 5f) in agreement with our experiments with the non-tumbling motile strain (Extended Fig. S5). Our simulations therefore recapitulate our experimental results and provide a mechanism for the emergence of the density fluctuations in binary mixtures: Circular swimming trajectories close to surfaces generate a vertical recirculation that disturbs the asymmetric distribution of sedimenting non-motile cells.

Discussion

In this paper, we report the emergence of density fluctuations in binary mixtures of motile and non-motile bacteria, over a broad range of physiologically relevant volume fractions of both species, and we identified the purely physical mechanism of this emergence, based on the combination of heterogeneous active flow fields with a symmetry breaking by gravity. Our experiments and modeling indeed demonstrated that persistent swimming in circles of the motile cells at the surface of the observation chamber, where they accumulate, creates recirculation fluid flows with regions of downwards suction surrounded by upward lift of the fluid, which collectively convect the non-motile species in three dimensions. At the population level, these flows are heterogeneous because the distribution of radii of curvature R of the swimming trajectories is broad and the typical fluid velocity produced by a circular swimmer scales as $1/R^3$, resulting in infrequent small R trajectories strongly contributing to the total flow. The second ingredient for the active density fluctuations to emergence is sedimentation under gravity, which breaks the vertical symmetry of the system, and allows cells to accumulate in the regions of upward lift, while they are depleted from the suction regions. In contrast, the fluid incompressibility prevents density heterogeneities to emerge in absence of gravity. Finally, this mechanism is strongly supported by numerical simulations of sedimenting non-motile particles in the flow generated by populations of surface circular swimmers, which could recapitulate all our experimental observations.

Our experiments clearly establish the sedimentation inverse length ($1/z_{sed} = v_{sed}/D_{NM}$) as one of the control parameters of the system (Fig. 3). The decrease of the heterogeneity level of the non-motile cells at high motile cell density can then be understood as a consequence of the co-occurrent increase in the effective diffusion coefficient of non-motile cells, which reduces $1/z_s$ and homogenizes them in the vertical direction (Extended Fig. S6). Our simulations and experiments agreed very well despite our rather strong approximations. Steric interactions become important at the non-motile densities we considered. Including them might thus improve this agreement further. Collective sedimentation effects, which can produce heterogeneities by themselves[69,70], are also likely to reinforce the heterogeneities produced by the motile cells and might explain the lower het-

erogeneity level in simulations compared to experiments (Fig. 5f). Our model also assumes that the swimmer force dipoles are parallel to the surface, although evidences suggest that the flagellum sticks up with an angle at the surface[9,12]. The dipole flow might then be slightly asymmetric, with a stronger contribution of the flagellum. The implied additional azimuthal component to the recirculation flow field was however not visible in our experiments, suggesting that such an effect is small at best. Our simulations, as well as experiments, indicate that the persistence time of the circular swimming controls the strength of the heterogeneities (Fig. 2e). This persistence time must decrease as a function of tumbling rate, tumbling angle and rotational diffusion[11,13,15,16]. The latter two are reduced in elongated cells[73], a probable explanation for the increased heterogeneity level upon elongation of WT cells (Extended Fig. S5). Interaction with the surface are furthermore known to affect tumbling[74] and possibly diffusion[16]. Therefore, the interaction with the surface might also reinforce the recirculative flows by allowing some long persistence of the circular motion in the WT case.

The emergent active density fluctuations we uncovered constitute a new type of collective structuration for bacteria. These density fluctuations indeed differ from MIPS-based phenomena[24–28,57,58], given that the spatial localization of the motile cells is not affected (Extended Fig. S3) and their influence on the non-motile cells can be explained only considering hydrodynamic interactions (Fig. 4). Our results thus highlight the importance of hydrodynamic interactions even below the onset of swirling collective motion known as bacterial turbulence[19,21–23]. In the most heterogeneous mixtures, the fluid flows are not strong enough to visibly affect other swimmers, but a flow structure has already emerged, which is revealed by the non-motile cells. This flow structure might have other signatures, which were up-to-now unrecognized, for example previously observed transient enhancement of chemotactic drifts[21]. Although our system requires three-dimensionality (Fig. S8), the behavior we observe shares similarities with FDPO, which is conceptualized on fluctuating surfaces[61–63], including giant number fluctuations of cell density and non exponentially decaying correlation functions (Extended Fig. S1). The temporally varying space-structured noise created by the swimmers, together with the gravity constraint, might then be interpretable in terms of a fluctuating surface[61]. Experiments with rapidly sedimenting or two-dimensionally constrained beads have shown that swimming bacteria can generate attractive interactions between them[47], which seem to lead to collective aggregation[48,49]. While these interactions were observed in two dimensions and mainly interpreted as active depletion interaction effects, the shape of the fluid flow generated by surface swimmers may also contribute to the observed behavior, in a low diffusion and infinite sedimentation limit of our experiments.

Active particles attached to surfaces can affect local dynamics, a phenomenon coined active carpet[75], by enhancing diffusion and creating fluid transport[76,77], mostly observed in eukaryotic cells like in the synchronously beating cilia[78,79] but also in such prokaryotic systems as engineered surfaces covered with attached flagellated bacteria[80]. Our observations fit in this general framework, further highlighting the importance of microorganism-surface interactions. However, in our case, the structure emerges from a population of freely swimming cells interacting with the surface.

The results presented in this study give insight into the potential impact of motility on microbial community dynamics in natural settings. Given the magnitude of the flow fields generated by the swimming *E. coli*, non-motile but also slowly moving species (at speeds $\lesssim 1\mu$m/s) should be affected by the density fluctuations. We expect the phenomenon we uncovered to play a role in the enhancement of self-aggregation[66,81], and also to affect the formation of biofilms in complex multispecies communities. Further investigation is required to determine whether the structure generated by motile cells provides any hindrance or benefits to non-motile cells. Additionally, while chemotaxis is not necessary for the emergence of this phenomenon, it is likely that interactions via chemical cues between the community members modulate the behavior in more natural settings[30–34,64–66]. Overall, these results shed light on the complex interplay between individual cells and their environment, and contribute to our understanding of how microorganisms interact and organize in nature.

Methods

Strains and plasmids

All strains used in this study are derived from *Escherichia coli* wild-type strain W3110 (RpoS+), for which the gene *flu*, which encodes for the major adhesin in our growth conditions, Antigen 43, was deleted[30]. Gene deletions are carried out using the plasmid pSIJ8, which carries the lambda Red recombineering genes inducible with arabinose and the flippase recombinase gene inducible with rhamnose[82]. The target gene is substituted with a PCR product that carries the Kanamycin (Kan) resistance cassette with lambda Red recombinase by adding 0.15% of arabinose. Recombinant cells are selected with the Kan resistance and then the marker is removed with the flippase recombinase by adding 50mM L-rhamnose to the culture. The deletion of the gene is checked by colony PCR and then the plasmid pSIJ8 is removed by growing the cells at 42°C. The non-motile strains lack the flagellin *fliC*. The chemotaxis response regulator *cheY* was deleted in the non-tumbling strain,

and the phosphatase *cheZ* was deleted in the tumbling-only strain.

All strains are fluorescently labeled with fluorescent proteins mCherry or mNeonGreen expressed from a plasmid with pTRc99a backbone with a strong ribosome binding site (RBS2)[83]. Motile strains were usually tagged with mCherry and non-motile strains with mNeonGreen. Reversing the tags gave identical results (Extended Fig. S1c). Growth medium is supplied with 100 μM Isopropyl β-d-1-thiogalactopyranoside (IPTG) for fluorescent protein expression induction and 0.1 mg/mL Ampicillin as antibiotic resistance marker. The non-chemotactic but tumbling strain expresses $cheY^{D13K,Y106W}$ ($cheY^{**}$) from a plasmid with pTRc99a backbone transformed in the $\Delta cheY$ strain. No IPTG was added, the expression of CheY** resulted only from the leaky expression of the plasmid. This strain was tagged with mCherry expressed from pOSK237, a plasmid with pBAD backbone, adding 0.004% L-arabinose for protein induction and 50 μg/mL Kanamycin as antibiotic resistance marker in the growth medium. All the strains and plasmids used in this study are listed in Table 1.

Microfabrication

Molds were fabricated using standard photolithography and microfabrication techniques. The SU8 photoresist (Microchem^{TM}) was spin-coated on a silicone wafer, prebaked, shaped by patterned UV exposure, postbaked and washed according to manufacturer instruction. The cast was then passivated with silane and checked using a profilometer. Poly-di-methylsiloxane (PDMS) was poured on the cast in a 1:10 crosslinker to base ratio, degassed, baked overnight at 65 °C, peeled off, cut to shape and covalently bound on isopropanol-rinsed microscopy high precision glass slides after oxygen plasma treatment. PDMS to glass covalent bonds were allowed to form for 20 minutes at room temperature and the devices were then filled with the samples.

Phase diagram assays

Cell growth and Sample preparation

Cells are grown overnight in liquid tryptone broth (TB) medium (1% Bacto tryptone + 0.5% NaCl) with appropriate antibiotics at 37 ºC. The overnight culture is diluted 1/100 in 10 mL fresh TB supplemented with antibiotics and inducers and grown at 34 ºC under shaking at 270 rpm for about 4h, until the optical density at 600 nm of the suspension reached $OD_{600} \simeq 0.5$.
For elongated cells, 100μg/mL of cephalexin, which blocks cell division, are added 60 minutes (i.e. about 1.5 cell cycles) before harvesting the cells, yielding cells about 2-4 times as long as without treatment.
The whole culture is then centrifuged at 3.5krpm for 5 minutes to prevent any damage to the flagella of motile cells. The pellet is resuspended in motility buffer (MB) (10mM KPO_4 + 0.1mM EDTA + 67mM NaCl + 0.01% Tween 80, pH=7.0) + 1% Glucose. This buffer allows cells to swim but not to grow. Tween 80 is a surfactant that prevents cell-surface adhesion[84,85]. Cells are brought to a high density (OD_{600}=60) by a subsequent centrifugation (3.5krpm, 5 minutes) and then 6 serial dilutions are done for each of the strains in MB. Each dilution of each strain is mixed with a dilution of the other strain in a proportion 1:1, to a total of 36 binary mixtures. Each mixture is then loaded in a chamber of a previously fabricated microfluidic device. The chamber consists of an inlet connected to an outlet by a straight channel of 50 μm height, 1 mm width and 1 cm length. The channel is then sealed with grease to prevent fluid flows. For steady state measurements, images are acquired after letting the mixtures settle in the microfluidic device for at least 40 minutes. For time dependent measurements, only two channels are filled at the same time, and brought immediately under the microscope. Each condition was probed in at least three independent biological replicates.

Physically modified media

When varying the buoyancy of the medium, cells are grown and processed as described above, except that the pellet of cells is resuspended in density matching motility buffer after the last centrifugation and the serial dilutions are made in this same density matching medium. The latter is prepared using iodixanol to increase the density of the MB up to the same density of a single *E.coli* cell (1.11g/ml). The density matching motility buffer contains 26.6% iodixanol, the concentration of the other components being unchanged (10mM KPO_4 + 0.1mM EDTA + 67mM NaCl + 0.01% Tween 80 + 26.6% iodixanol). For the experiments where different buoyancies were tested, the volume of iodixanol was increased to increase the density of the buffer: 0% ($\Delta\rho g = 1$), 13.2% ($\Delta\rho g = 0.5$), 19.9% ($\Delta\rho g = 0.25$), 23.9% ($\Delta\rho g = 0.1$), 26.6% ($\Delta\rho g = 0$).

Epifluorescence Microscopy

The density fluctuations are visualized using the Nikon Ti-E inverted fluorescence widefield microscope with a 20x (NA 0.75) objective and the mCherry (excitation filter 572/25, emission 645/90) and GFP (excitation

filter 470/40, emission 525/50) filters. Images are acquired in the focal plane 9 μm above the bottom of the microfluidic channel. Focal plane is maintained using the Nikon perfect focus system.

Density fluctuations quantification

A stack of images of 9 different positions spanning across the microfluidic channel is acquired 7 times for each fluorescence channel with the Andor Zyla sCMOS Camera and NIS software (1 px = 0.32 μm, field of view = 2048 x 2048 px^2, 50ms exposure). Further image processing is realized with Fiji (https://fiji.sc/). Each image in the stack is corrected for inhomogeneous illumination. For this, an illumination background image is computed at each cell density as the median fluorescence of a large set of images taken at this density. The intensity of this background image is then normalized by its mean. Each image of the stack is divided by the illumination background image corresponding to its cell density. To further correct for small remaining systematic biases in the intensity profile of each image, a quadratic polynomial filter is applied.

The spatial autocorrelation function of the corrected intensity is computed as:

$$C(\mathbf{r}) = \frac{1}{n(\mathbf{r})} \sum_{\mathbf{r_0}}^{n(\mathbf{r})} \left(\frac{I(\mathbf{r}+\mathbf{r_0})I(\mathbf{r_0})}{\langle I^2 \rangle} - 1 \right) \tag{1}$$

where $n(\mathbf{r})$ is the number of valid initial pixel positions $\mathbf{r_0}$ for a given spatial shift $\mathbf{r}$. In homogeneous control, this function decays over a length scale of about 1 μm corresponding to the bacterial size, whereas a second decay appears at several tens of microns when density fluctuations are present (Extended Fig. S1e).

The density fluctuations are quantified by calculating the heterogeneity level of the images. For this, each image is divided into non-overlapping squares of size N, allowed to vary from $N = (2^0)^2$ px^2 to $N = (2^9)^2$ px^2 (Extended Fig. S1a). The sum of the fluorescence intensity $I_{\mathbf{k},N} = \sum_{\mathbf{j}}^{N} I_{\mathbf{k}+\mathbf{j}}$ is computed within each square (indexed by $\mathbf{k}$). The spatial variance σ_N^2 of this total intensity is then computed across the squares of size N,

$$\sigma_N^2 = \mathrm{var}_{\mathbf{k}}(I_{\mathbf{k},N}). \tag{2}$$

The latter is then normalized by the pixel size of the slice (N) and by the variance over squares of size $N = 1$, σ_1^2 (Extended Fig. S1b), since one expects from the central limit theorem that $\sigma_N^2 = N\sigma_1^2$ for an image made of pixels that assume random non-correlated values. The ratio $\sigma_N^2/N\sigma_1^2$ increasing as a power of N is then a marker of large scale correlations in the spatial distribution of fluorescence intensity, often called giant number fluctuations[86].

Since our images have a texture that comes from the cells being several pixels large, even in homogeneous samples, $\sigma_N^2/N\sigma_1^2$ transiently increases with N until the latter reaches the area of a single cell ($N \sim 100$ px^2), above which $\sigma_N^2/N\sigma_1^2$ becomes constant as a function of N. In contrast, in samples with density fluctuations, $\sigma_N^2/N\sigma_1^2$ keeps increasing for $N > 100$ px^2, marking the large scale fluctuations of the cell density (Extended Fig. S1b). To reduce the corresponding heterogeneity level to a single number, we use $\sigma_N^2/N\sigma_1^2$ for a square of area $N = (2^7)^2$ px^2 = 16384 px^2. The width of this square is indeed about 42 μm, which corresponds to the large scale decay of the spatial autocorrelation function of the pixel intensity (Extended Fig. S1e). Since the value of $\sigma_N^2/N\sigma_1^2$ for $N = (2^7)^2$ in the control passive systems varies little across non-motile volume fractions and is close to 400 on average, we further defined the heterogeneity level as $(\sigma_N^2/N\sigma_1^2)/400$ for $N = (2^7)^2$, such that passive systems have a heterogeneity level of 1 on average, and heterogeneity is defined relative to this homogeneous case.

Image velocimetry analysis

A movie of 7500 binned frames is acquired at 5 frames per second with the Andor Zyla sCMOS Camera and NIS software (1 px = 1.3 μm, field of view = 512 x 512 px^2, 50ms exposure). The velocity field, the flow structure factor and the kinetic energy of the non-motile cells are measured using a local image velocimetry algorithm based on Fourier image analysis, following previously described protocol[21] (Github code: 10.5281/zenodo.3516258). In short, local velocities are computed on a rectangular grid with 4 px spacing, using Gaussian filtered (half width 3 px) squared subframes (32×32 px^2), with a 10 frames long sliding window. The spatial Fourier transform of the velocity field is computed as

$$\mathbf{v}(\mathbf{q}, t) = \sum_{\mathbf{x}} \mathbf{v}(\mathbf{x}, t) \exp\left(-i\mathbf{q}\cdot\mathbf{x}\right) \tag{3}$$

and the power spectrum as

$$E(q) = q \left\langle |\mathbf{v}(\mathbf{q}, t)|^2 \right\rangle_{|\mathbf{q}|=q,t} / \left\langle |\mathbf{v}(\mathbf{x}, t)|^2 \right\rangle_{|\mathbf{x}|,t} \tag{4}$$

Confocal microscopy

The binary mixtures are prepared following the above-mentioned Sample preparation and are visualized using a Zeiss LSM-880 confocal microscope equipped with a C-Apochromat 40x (NA 1.2) water immersion objective. Motile cells (tagged with mCherry) and non-motile cells (tagged with mNeonGreen) are visualized using the 594nm Helium-Neon laser and 488nm Argon laser, and detected with a photomultiplier at 490-535nm and 598-660nm, respectively.

Cell fraction vertical distribution

Z-stack images of the whole microfluidic channel are acquired using the ZEN Black software (Zeiss) (1px = $0.2\,\mu$m in X and Y, 1px = $1\,\mu$m in Z; field of view = 303.64 x 303.64 x 70 μm^3, $0.35\,\mu$s/px exposure). The number of cells in each Z plane is quantified by the connected components labeling algorithm for ImageJ [87]. Because the height and the tilt of the microfluidic channels slightly vary from sample to sample ($< 10\%$), the Z position is binned and the mean of the cell fraction over the bins is calculated.

3D tracking

To track the horizontal and vertical motion of the non-motile cells, the binary mixtures are prepared following the above-mentioned Sample preparation. To be able to track single non-motile cells, only the 0.03% of them are fluorescently labeled. Low density regions ("holes") of the mixtures are identified with transmitted light and Z-stacks are acquired at the highest possible speed around these regions using the ZEN Black software (Zeiss) (1px = $0.54\,\mu$m in X and Y, 1px = $2\,\mu$m in Z; field of view = 163.5 x 163.5 x 20 μm^3, $1.3\,\mu$s/px exposure, $\sim$0.3fps).

The three-dimensional trajectories are reconstructed using the Mosaic plugin for ImageJ [88]. From each trajectory is then extracted their relative position to the hole center (defined by transmitted light) and their velocities in the vertical and horizontal planes.

Motility assay and circular trajectories measurements

From an overnight culture, cells are grown in TB for 4h at 34ºC. The culture is centrifuged (3.5krpm for 5 minutes), and the pellet is resuspended in MB + 1% Glucose. Cells are centrifuged again (3.5krpm for 5 minutes) and resuspended in the adequate buffer (MB, MB + 26.6% iodixanol). The sample is then loaded in a previously fabricated microfluidic channel.

Cell motility is visualized using the Nikon TI with a 10x (NA 0.3) objective. A stack of 10000 frames at 100 frames per second is acquired under phase contrast illumination with an Eosens 4cxp CMOS camera and StreamPix Multicamera software. The swimming speed of cells, the fraction of motile cells and the diffusion coefficient of non-motile cells are measured using Differential Dynamic Microscopy [89], an image analysis method based on Fourier image analysis, following the protocol described in [21].

To reconstruct the horizontal trajectories of the cells close to the surface, a stack of 2000 frames at 50 frames per second is acquired. Single 2D trajectories are reconstructed using the plugin ParticleTracking for ImageJ [90] (https://github.com/croelmiyn/ParticleTracking). Cell trajectories are categorized into runs and tumbles as previously described [90]. To calculate the radius of curvature of the trajectories, only runs are considered. The parametric curvature $C(t)$ is calculated for frame t as

$$C(t) = \left| \frac{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}}{(\dot{x}\ddot{y} - \ddot{x}\dot{y})} \right| \tag{5}$$

where the first and second derivatives of $x(t)$ and $y(t)$ are extracted from the cell tracking data by Taylor expansion. For this, a 50 frame sliding window is considered around time t, over which $x(t')$ and $y(t')$ are least-squares fitted to a second-order polynomial to obtain their first and second derivatives, as given by Taylor's expansion:

$$f(t') = \sum_{n=0}^{2} f^{(n)}(t) \frac{(t'-t)^n}{n!} + O((t'-t)^3) \tag{6}$$

The mean curvature is computed for each run, and for each trajectory as the average of $C(t)$ over all times t for which the cell is categorized as running. The radius is then the inverse of the curvature, $R = 1/\langle C(t) \rangle$. When computing histograms, the mean radius R of a trajectory has a weight equal to the duration of this trajectory.

To estimate the circle persistence time of a trajectory of duration T, we compute the mean linear displacement

$$dr(t) = \left\langle \sqrt{\left(x(t+t_0) - x(t_0)\right)^2 + \left(y(t+t_0) - y(t_0)\right)^2} \right\rangle_{t_0 \in [1, T-t]} \tag{7}$$

and find the smallest lag time $t = \tau$ for which $dr(t)$ exceeds the distance $3R$ with R the above defined radius of curvature, ie 1.5 times the average circle diameter. If $dr(t)$ never exceeds $3R$, we assign $\tau = T$.

Numerical simulations

All simulations were performed in a three dimensional rectangular box of width and length W and of height H. Boundary conditions are periodic on the lateral edges of the simulation box, and reflexive no-slip at the top and bottom walls. Cell volume fractions are $\phi = nV/HW^2$, with n the number of cells and the cell volume V being twice higher for elongated cells. All parameter values are given in Table S1.

Mean flow field generated by a circular swimmer

The motile cell is modeled as a pusher dipole of point forces, representing the cell body and the flagellum, both of magnitude $f_0/6\pi\eta = a^S v_0$, with η the fluid viscosity, a^S the hydrodynamic radius of the cell body and v_0 the swimming speed[8]. The fluid flow $\mathbf{v}^{i\leftarrow j}$ in position $\mathbf{x}_i$ generated by a point force $\mathbf{f}_j$ positioned in $\mathbf{x}_j$ is computed as

$$\mathbf{v}^{i\leftarrow j} = \mathbf{\Omega}^{i\neq j} \cdot \frac{\mathbf{f}_j}{6\pi\eta} \tag{8}$$

where $\mathbf{\Omega}^{i\neq j}$ is the normalized hydrodynamic-interaction tensor, i.e. the normalized Green function, in our geometry, of the Stokes equation:

$$-\nabla p + \eta \Delta \mathbf{v} + \mathbf{f} = 0\,. \tag{9}$$

The tensor $\mathbf{\Omega}^{i\neq j}$ and therefore the hydrodynamic flows $\mathbf{v}^{i\leftarrow j}$ are computed using the two-dimensional Fourier-space expansion of the hydrodynamic flow introduced in Mucha *et al.*[71], using the N log(N) algorithm of Hernandez-Ortiz *et al.*[72], which allows an efficient calculation of the fluid flow generated by multiple point forces in our geometry via the additivity of the solutions to Stokes equation, and which we implemented in Java (Supplementary notes). The dipole flow $v^D(\mathbf{x}; \mathbf{x}_0, \mathbf{u})$ is the sum of the flow generated by the two point forces, $+\mathbf{u}f_0$ located at $\mathbf{x}_0 + l_d\mathbf{u}$ (cell body), and $-f_0\mathbf{u}$ located at $\mathbf{x}_0 - l_d\mathbf{u}$ (flagellum), with $\mathbf{x}_0$ the position of the center of the cell, $\mathbf{u}$ the unit vector pointing in the swimming direction and l_d the half dipole length. To describe circular surface swimmers, the dipoles are horizontal and located at a height $z_M = 1\,\mu$m from the surface over which they circulate. The mean flow field generated over one revolution is computed as the average of the dipole flow:

$$v_{1r}(\mathbf{x}; R) = \frac{1}{n\theta} \sum_{k=0}^{n\theta-1} v^D(\mathbf{x}; \mathbf{x}_k, \mathbf{u}_k) \tag{10}$$

over $n\theta = 100$ equally spaced positions of the dipole on a circle of radius R, $\mathbf{x}_k = (R\cos\theta_k, R\sin\theta_k, z_M)$, with $\theta_k = 2\pi k/n\theta$, the dipole being tangential to the circle, $\mathbf{u}_k = (-\sin\theta_k, \cos\theta_k, 0)$. Because this configuration of forces is rotationally symmetric, the mean flow field only depends on the longitudinal (z) and radial (r) coordinates in the cylindrical coordinate system relative to the center of the circle, and has no azimuthal component.

The flow field was analyzed using MATLAB (2020a). We define the inward flux generated by the circular swimmer at a distance r by:

$$\Phi(r) = 2\pi r \int_0^H \left|\mathbf{v}_{1r}(r, z \,|\, \mathbf{v}_{1r} \cdot \mathbf{u}_r < 0) \cdot \mathbf{u}_r\right| \, dz \tag{11}$$

with $\mathbf{u}_r$ the radial unit vector. The recirculation radius is defined as the distance r which maximizes Φ, corresponding to the position of the center of the recirculation whirl. The recirculation flux is defined as the maximal flux $\max_r \Phi(r)$.

The vertically averaged $v_{1r}(\mathbf{x}; R)$ of Fig. 4d,e is computed as $\langle v \rangle(r) = \int_0^{Hc} dz\, v_{1r}((r,z); R) p(z)$, with the fitted tracer distribution $p(z) = \exp(-z/\lambda)/\int_0^{Hc} dz \exp(-z/\lambda)$, $\lambda = 10\,\mu$m the experimental decay length and $Hc = 20\,\mu$m the height of the confocal volume.

Distribution of tracers in the mean flow field of a population of circular swimmers

The mean flow field created by a population of n_{Mc} (250 to $5\,10^4$) circular swimmers sums the mean flows v_{1r} of the individual circular swimmers. Swimmers are assigned with equal probability to the top ($z = H - z_M$) or bottom ($z = z_M$) wall, mimicking experiments (Extended Fig. S7b). The centers of the circles are distributed uniformly in the horizontal (x, y) directions. The circle radius is drawn from a long-tailed distribution $p(R)$, which was parametrized to fit the experimental distribution of circle radii for normal-sized and elongated cells (Extended Fig. S11a-b). We perform a simplified Brownian dynamics simulation of the motion of n_{NM} ($5\,10^3$ to $2\,10^4$) non-motile cells that are subjected to the mean flow field of the circles, gravity and Brownian motion. The evolution of the position of tracer i obeys

$$d\mathbf{r}_i = \left[\mathbf{v}_{sed} + \sum_{j}^{\text{dipoles}} v_{1r}(\mathbf{r}_i - \mathbf{x}_j; R_j) \right] dt + \sqrt{2D_{NM}}\, d\mathbf{w}_i \tag{12}$$

with $\mathbf{v}_{sed}$ the downward-oriented sedimentation speed, $\mathbf{x}_j$ and R_j the center and radius of circle j, which are kept fixed throughout the simulation, dt the simulation time step, D_{NM} the diffusion coefficient of the non-motile cells, and $d\mathbf{w}_i = \sqrt{dt}\sigma_i$ the normalized Brownian noise. The Gaussian-distributed three-dimensional random variable σ_i has unit variance and no cross-correlation with $\sigma_{j\neq i}$. The initial distribution of non-motile cells is uniform horizontally and obeys the Boltzmann distribution in confined geometry along the vertical axis

$$p(z_i, t = 0) = \frac{v_{sed}}{D_{NM}} \frac{\exp\left(-z_i\, v_{sed}/D_{NM}\right)}{1 - \exp(-H v_{sed}/D_{NM})}\,. \tag{13}$$

Steric interactions, direct hydrodynamic interactions between non-swimmers and the effect of confinement on self-diffusion are neglected (Supplementary notes). The equation of motion is time integrated using an Euler scheme, with $dt = 0.1$ s in equivalent real time. The configuration of the system is saved every step, for a total duration of the simulation of 1000 steps. Because the circle configuration and thus the velocity field are fixed, the time elapsed since the beginning of the simulation can be interpreted as a persistence time of the circles, and the configuration after this time as a realization of the structure of the population of non-motile cells in the mixture for said persistence time.

A pseudo image of the suspension is produced for each time step of the simulation from the positions of the non-motile tracers. It renders the microscopy image that is expected under our experimental conditions for the given non-motile cells configuration, assuming a Gaussian point spread function for our objective. Each particle, positioned in $(\mathbf{r_i} = (x_i, y_i), z_i)$ contribute to the intensity in position $\mathbf{r} = (x, y)$ in the rendered image at $h = 8\,\mu$m above the bottom of the channel as:

$$I_i = \frac{1}{1 + 3s^2} \exp\left(\frac{2\left(\mathbf{r} - \mathbf{r_i}\right)^2}{a_d^2(1 + 3s^2)} \right)\,, \tag{14}$$

with $s = (z_i - h)/dh$ the distance to the focal plane, normalized to the depth of field dh. The images are printed with a pixel size of $\text{px}_{sim} = 1\,\mu$m, which is larger than in our experimental setup, to save computer memory. They are then analyzed with the same algorithms as the experimental data. We used a box size of $32 \times 32\,\mu\text{m}^2$ to compute the heterogeneity level as $\sigma_N^2/N\sigma_1^2$ normalized to its value in the purely passive homogeneous control, similarly to experiments. The value of $\sigma_N^2/N\sigma_1^2$ in the simulated passive system is 8.9, smaller than in experiments owing to the smaller pixel size there ($1\,\text{px}_{exp} = 0.325\,\mu$m).

The simulations for each parameter set are repeated at least three times and quantifications are averaged over these repeats. The fully parallelized java code for running the simulations is available in Github (https://github.com/croelmiyn/SimulationMixtures, DOI:10.5281/zenodo.7919119).

Acknowledgements

The authors disclose support from the Deutsche Forschungsgemeinschaft (Grant No. CO 1813/2-1) and the Max Planck Gesellschaft. The authors thank L. Laganenka, O. Schauer and G. Scarinci for providing plasmids, G. Malengo for technical microscopy support, Chia-Fu Chou for helping with wafer fabrication, and V. Sourjik, S. Murray and H. Niederholtmeyer for helpful discussions.

References

[1] Berg, H. C. & Anderson, R. A. Bacteria swim by rotating their flagellar filaments. *Nature* 380–382 (1973).

[2] Tu, Y. Quantitative modeling of bacterial chemotaxis: signal amplification and accurate adaptation. *Annu Rev Biophys* **42**, 337–59 (2013).

[3] Parkinson, J. S., Hazelbauer, G. L. & Falke, J. J. Signaling and sensory adaptation in escherichia coli chemoreceptors: 2015 update. *Trends in Microbiology* **23**, 257–266 (2015).

[4] Colin, R. & Sourjik, V. Emergent properties of bacterial chemotaxis pathway. *Current Opinion in Microbiology* **39**, 24–33 (2017).

[5] Waite, A. J., Frankel, N. W. & Emonet, T. Behavioral variability and phenotypic diversity in bacterial chemotaxis. *Annu Rev Biophys* **47**, 595–616 (2018).

[6] Colin, R., Ni, B., Laganenka, L. & Sourjik, V. Multiple functions of flagellar motility and chemotaxis in bacterial physiology. *FEMS Microbiol Rev* fuab038 (2021).

[7] Drescher, K., Dunkel, J., Cisneros, L. H., Ganguly, S. & Goldstein, R. E. Fluid dynamics and noise in bacterial cell-cell and cell-surface scattering. *Proceedings of the National Academy of Sciences* **108**, 10940–10945 (2011).

[8] Elgeti, J., Winkler, R. G. & Gompper, G. Physics of microswimmers—single particle motion and collective behavior: a review. *Reports on Progress in Physics* **78**, 056601 (2015).

[9] Vigeant, M. A.-S., Ford, R. M., Wagner, M. & Tamm, L. K. Reversible and irreversible adhesion of motile Escherichia coli cells analyzed by total internal reflection aqueous fluorescence microscopy. *Applied and Environmental Microbiology* (2002). Publisher: American Society for Microbiology.

[10] Berke, A. P., Turner, L., Berg, H. C. & Lauga, E. Hydrodynamic attraction of swimming microorganisms by surfaces. *Physical Review Letters* **101**, 038102 (2008).

[11] Li, G. & Tang, J. X. Accumulation of microswimmers near a surface mediated by collision and rotational Brownian motion. *Physical Review Letters* **103**, 078101 (2009). Publisher: American Physical Society.

[12] Bianchi, S., Saglimbeni, F. & Di Leonardo, R. Holographic imaging reveals the mechanism of wall entrapment in swimming bacteria. *Physical Review X* **7**, 011010 (2017).

[13] Junot, G. *et al.* Run-to-tumble variability controls the surface residence times of *E. coli* bacteria. *Physical Review Letters* **128**, 248101 (2022).

[14] Lauga, E., DiLuzio, W. R., Whitesides, G. M. & Stone, H. A. Swimming in Circles: Motion of bacteria near solid boundaries. *Biophysical Journal* **90**, 400–412 (2006).

[15] Li, G. L., Tam, L. K. & Tang, J. X. Amplified effect of brownian motion in bacterial near-surface swimming. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 18355–18359 (2008).

[16] Perez Ipiña, E., Otte, S., Pontier-Bres, R., Czerucka, D. & Peruani, F. Bacteria display optimal transport near surfaces. *Nature Physics* **15**, 610–615 (2019).

[17] Cates, M. E. Diffusive transport without detailed balance in motile bacteria: does microbiology need statistical physics? *Reports on Progress in Physics* **75**, 042601 (2012).

[18] Ginelli, F., Peruani, F., Bär, M. & Chaté, H. Large-scale collective properties of self-propelled rods. *Physical Review Letters* **104**, 184502 (2010).

[19] Sokolov, A., Aranson, I. S., Kessler, J. O. & Goldstein, R. E. Concentration dependence of the collective dynamics of swimming bacteria. *Physical Review Letters* **98**, 158102 (2007).

[20] Wensink, H. H. *et al.* Meso-scale turbulence in living fluids. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 14308–14313 (2012).

[21] Colin, R., Drescher, K. & Sourjik, V. Chemotactic behaviour of *Escherichia coli* at high cell density. *Nature Communications* **10**, 5329 (2019).

[22] Martinez, V. A. *et al.* A combined rheometry and imaging study of viscosity reduction in bacterial suspensions. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 2326–2331 (2020).

[23] Peng, Y., Liu, Z. & Cheng, X. Imaging the emergence of bacterial turbulence: Phase diagram and transition kinetics. *Science Advances* **7**, eabd1240 (2021).

[24] Cates, M. E., Marenduzzo, D., Pagonabarraga, I. & Tailleur, J. Arrested phase separation in reproducing bacteria creates a generic route to pattern formation. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 11715–11720 (2010).

[25] Fily, Y. & Marchetti, M. C. Athermal phase separation of self-propelled particles with no alignment. *Physical Review Letters* **108**, 235702 (2012).

[26] Cates, M. E. & Tailleur, J. Motility-induced phase separation. *Annual Review of Condensed Matter Physics* **6**, 219–244 (2015).

[27] Buttinoni, I. *et al.* Dynamical clustering and phase separation in suspensions of self-propelled colloidal particles. *Physical Review Letters* **110**, 238301 (2013).

[28] Liu, C. *et al.* Sequential establishment of stripe patterns in an expanding cell population. *Science* **334**, 238–41 (2011).

[29] Agudo-Canalejo, J. & Golestanian, R. Active phase separation in mixtures of chemically interacting particles. *Physical Review Letters* **123**, 018101 (2019).

[30] Laganenka, L., Colin, R. & Sourjik, V. Chemotaxis towards autoinducer 2 mediates autoaggregation in *Escherichia coli*. *Nature Communications* **7**, 12984 (2016).

[31] Budrene, E. O. & Berg, H. C. Dynamics of formation of symmetrical patterns by chemotactic bacteria. *Nature* **376**, 49–53 (1995). Number: 6535 Publisher: Nature Publishing Group.

[32] Park, S. *et al.* Influence of topology on bacterial social interaction. *Proceedings of the National Academy of Sciences* **100**, 13910–13915 (2003). Publisher: National Academy of Sciences Section: Biological Sciences.

[33] Saragosti, J. *et al.* Directional persistence of chemotactic bacteria in a traveling concentration wave. *Proceedings of the National Academy of Sciences* **108**, 16235–16240 (2011). Publisher: National Academy of Sciences Section: Biological Sciences.

[34] Fu, X. *et al.* Spatial self-organization resolves conflicts between individuality and collective migration. *Nature Communications* **9**, 2177 (2018). Number: 1 Publisher: Nature Publishing Group.

[35] Cordero, O. X. & Datta, M. S. Microbial interactions and community assembly at microscales. *Current Opinion in Microbiology* **31**, 227–234 (2016).

[36] Nadell, C. D., Drescher, K. & Foster, K. R. Spatial structure, cooperation and competition in biofilms. *Nature Reviews Microbiology* **14**, 589–600 (2016).

[37] Azimi, S., Lewin, G. R. & Whiteley, M. The biogeography of infection revisited. *Nat Rev Microbiol* **20**, 579–592 (2022).

[38] Lion, S. & Baalen, M. Self-structuring in spatial evolutionary ecology. *Ecol Lett* **11**, 277–95 (2008).

[39] Liu, W. *et al.* Interspecific bacterial interactions are reflected in multispecies biofilm spatial organization. *Frontiers in Microbiology* **7** (2016).

[40] Stubbendieck, R. M., Vargas-Bautista, C. & Straight, P. D. Bacterial communities: Interactions to scale. *Front Microbiol* **7**, 1234 (2016).

[41] Yanni, D., Marquez-Zacarias, P., Yunker, P. J. & Ratcliff, W. C. Drivers of spatial structure in social microbial communities. *Current Biology* **29**, R545–R550 (2019).

[42] Eigentler, L., Davidson, F. A. & Stanley-Wall, N. R. Mechanisms driving spatial distribution of residents in colony biofilms: an interdisciplinary perspective. *Open Biol* **12**, 220194 (2022).

[43] Wu, X.-L. & Libchaber, A. Particle diffusion in a quasi-two-dimensional bacterial bath. *Physical Review Letters* **84**, 3017–3020 (2000).

[44] Kim, M. J. & Breuer, K. S. Enhanced diffusion due to motile bacteria. *Phys. Fluids* **16**, 5 (2004).

[45] Peng, Y. *et al.* Diffusion of ellipsoids in bacterial suspensions. *Physical Review Letters* **116**, 068303 (2016). Publisher: American Physical Society.

[46] Singh, J., Patteson, A. E., Maldonado, B. O. T., Purohit, P. K. & Arratia, P. E. Bacterial activity hinders particle sedimentation. *Soft Matter* **17**, 4151–4160 (2021). Publisher: The Royal Society of Chemistry.

[47] Angelani, L., Maggi, C., Bernardini, M. L., Rizzo, A. & Di Leonardo, R. Effective interactions between colloidal particles suspended in a bath of swimming cells. *Physical Review Letters* **107**, 138302 (2011).

[48] Gokhale, S., Li, J., Solon, A., Gore, J. & Fakhri, N. Dynamic clustering of passive colloids in dense suspensions of motile bacteria. *Phys Rev E* **105**, 054605 (2022).

[49] Bouvard, J., Moisy, F. & Auradou, H. Ostwald-like ripening in the two-dimensional clustering of passive particles induced by swimming bacteria. *Phys. Rev. E* **107**, 044607 (2023).

[50] Takatori, S. C. & Brady, J. F. A theory for the phase behavior of mixtures of active particles. *Soft Matter* **11**, 7920–7931 (2015).

[51] Kümmel, F., Shabestari, P., Lozano, C., Volpe, G. & Bechinger, C. Formation, compression and surface melting of colloidal clusters by active particles. *Soft Matter* **11**, 6187–6191 (2015).

[52] Stenhammar, J., Wittkowski, R., Marenduzzo, D. & Cates, M. E. Activity-induced phase separation and self-assembly in mixtures of active and passive particles. *Physical Review Letters* **114**, 018301 (2015).

[53] Wysocki, A., Winkler, R. G. & Gompper, G. Propagating interfaces in mixtures of active and passive brownian particles. *New Journal of Physics* **18**, 123030 (2016).

[54] Dolai, P., Simha, A. & Mishra, S. Phase separation in binary mixtures of active and passive particles. *Soft Matter* **14**, 6137–6145 (2018). Publisher: Royal Society of Chemistry.

[55] Rodriguez, D. R., Alarcon, F., Martinez, R., Ramírez, J. & Valeriani, C. Phase behaviour and dynamical features of a two-dimensional binary mixture of active/passive spherical particles. *Soft Matter* **16**, 1162–1169 (2020). Publisher: Royal Society of Chemistry.

[56] McCandlish, R. S., Baskaran, A. & F. Hagan, M. Spontaneous segregation of self-propelled particles with different motilities. *Soft Matter* **8**, 2527–2534 (2012).

[57] Wittkowski, R., Stenhammar, J. & Cates, M. E. Nonequilibrium dynamics of mixtures of active and passive colloidal particles. *New Journal of Physics* **19**, 105003 (2017).

[58] Weber, S. N., Weber, C. A. & Frey, E. Binary mixtures of particles with different diffusivities demix. *Physical Review Letters* **116**, 058301 (2016). Publisher: American Physical Society.

[59] Williams, S., Jeanneret, R., Tuval, I. & Polin, M. Confinement-induced accumulation and de-mixing of microscopic active-passive mixtures. *Nat Commun* **13**, 4776 (2022).

[60] Marchetti, M. C. *et al.* Hydrodynamics of soft active matter. *Reviews of Modern Physics* **85**, 1143 (2013).

[61] Das, D. & Barma, M. Particles sliding on a fluctuating surface: Phase separation and power laws. *Physical Review Letters* **85**, 1602–1605 (2000). Publisher: American Physical Society.

[62] Mishra, S. & Ramaswamy, S. Active nematics are intrinsically phase separated. *Physical Review Letters* **97**, 090602 (2006). Publisher: American Physical Society.

[63] Das, A., Polley, A. & Rao, M. Phase segregation of passive advective particles in an active medium. *Physical Review Letters* **116**, 068306 (2016). Publisher: American Physical Society.

[64] Curatolo, A. I. *et al.* Cooperative pattern formation in multi-component bacterial systems through reciprocal motility regulation. *Nature Physics* **16**, 1152 (2020).

[65] Gude, S. *et al.* Bacterial coexistence driven by motility and spatial competition. *Nature* **578**, 588–592 (2020).

[66] Laganenka, L. & Sourjik, V. Autoinducer 2-Dependent Escherichia coli biofilm formation is enhanced in a dual-species coculture. *Applied and Environmental Microbiology* **84** (2018).

[67] Muok, A. R. & Briegel, A. Intermicrobial hitchhiking: How nonmotile microbes leverage communal motility. *Trends in Microbiology* **29**, 542–550 (2020).

[68] Scharf, B. E., Fahrner, K. A., Turner, L. & Berg, H. C. Control of direction of flagellar rotation in bacterial chemotaxis. *Proc Natl Acad Sci U S A* **95**, 201–206 (1998).

[69] Royall, C. P., Dzubiella, J., Schmidt, M. & van Blaaderen, A. Nonequilibrium sedimentation of colloids on the particle scale. *Physical Review Letters* **98**, 188304 (2007). Publisher: American Physical Society.

[70] Wysocki, A. *et al.* Direct observation of hydrodynamic instabilities in a driven non-uniform colloidal dispersion. *Soft Matter* **5**, 1340–1344 (2009). Publisher: The Royal Society of Chemistry.

[71] Mucha, P. J., Tee, S. Y., Weitz, D. A., Shraiman, B. I. & Brenner, M. P. A model for velocity fluctuations in sedimentation. *Journal of Fluid Mechanics* **501**, 71–104 (2004).

[72] Hernandez-Ortiz, J. P., de Pablo, J. J. & Graham, M. D. N log n method for hydrodynamic interactions of confined polymer systems: Brownian dynamics. *Journal of Chemical Physics* **125**, 164906 (2006).

[73] Schauer, O. *et al.* Motility and chemotaxis of bacteria-driven microswimmers fabricated using antigen 43-mediated biotin display. *Scientific Reports* **8**, 9801 (2018).

[74] Molaei, M., Barry, M., Stocker, R. & Sheng, J. Failed escape: Solid surfaces prevent tumbling of escherichia coli. *Physical Review Letters* **113**, 068103 (2014).

[75] Guzmán-Lastra, F., Löwen, H. & Mathijssen, A. J. T. M. Active carpets drive non-equilibrium diffusion and enhanced molecular fluxes. *Nature Communications* **12**, 1906 (2021).

[76] Mathijssen, A. J. T. M., Guzman-Lastra, F., Kaiser, A. & Lowen, H. Nutrient transport driven by microbial active carpets. *Physical Review Letters* **121**, 248101 (2018).

[77] Xu, H. R., Dauparas, J., Das, D., Lauga, E. & Wu, Y. L. Self-organization of swimmers drives long-range fluid transport in bacterial colonies. *Nature Communications* **10**, 1792 (2019).

[78] Sleigh, M. A., Blake, J. R. & Liron, N. The propulsion of mucus by cilia. *American Review of Respiratory Disease* **137**, 726–741 (1988).

[79] Gilpin, W., Bull, M. S. & Prakash, M. The multiscale physics of cilia and flagella. *Nature Reviews Physics* **2**, 74–88 (2020).

[80] Darnton, N., Turner, L., Breuer, K. & Berg, H. C. Moving fluid with bacterial carpets. *Biophysical Journal* **86**, 1863–1870 (2004).

[81] Lushi, E., Goldstein, R. E. & Shelley, M. J. Collective chemotactic dynamics in the presence of self-generated fluid flows. *Physical Review E* **86**, 040902 (2012). Publisher: American Physical Society.

[82] Jensen, S. I., Lennen, R. M., Herrgård, M. J. & Nielsen, A. T. Seven gene deletions in seven days: Fast generation of *Escherichia coli* strains tolerant to acetate and osmotic stress. *Scientific Reports* **5**, 17874 (2015).

[83] Besharova, O., Suchanek, V. M., Hartmann, R., Drescher, K. & Sourjik, V. Diversification of gene expression during formation of static submerged biofilms by Escherichia coli. *Frontiers in Microbiology* **7**, 1568 (2016).

[84] Nielsen, C. K., Kjems, J., Mygind, T., Snabe, T. & Meyer, R. L. Effects of Tween 80 on growth and biofilm formation in laboratory media. *Frontiers in Microbiology* **7** (2016).

[85] Schwarz-Linek, J. *et al.* Escherichia coli as a model active colloid: A practical introduction. *Colloids and Surfaces B: Biointerfaces* **137**, 2–16 (2016).

[86] Dey, S., Das, D. & Rajesh, R. Spatial structures and giant number fluctuations in models of Active Matter. *Physical Review Letters* **108**, 238001 (2012). Publisher: American Physical Society.

[87] Legland, D., Arganda-Carreras, I. & Andrey, P. MorphoLibJ: integrated library and plugins for mathematical morphology with ImageJ. *Bioinformatics* **32**, 3532–3534 (2016).

[88] Gong, Y. & Sbalzarini., I. F. Image enhancement by gradient distribution specification. *In Proc. ACCV, 12th Asian Conference on Computer Vision, Workshop on Emerging Topics in Image Enhancement and Restoration* w7–p3 (2014).

[89] Wilson, L. G. *et al.* Differential dynamic microscopy of bacterial motility. *Physical Review Letters* **106**, 018101 (2011).

[90] Suchanek, V. M. *et al.* Chemotaxis and cyclic-di-GMP signalling control surface attachment of Escherichia coli. *Molecular Microbiology* **113**, 728–739 (2020). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mmi.14438.

[91] Amann, E., Ochs, B. & Abel, K.-J. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli. *Gene* **69**, 301–315 (1988).

[92] Serra, D. O., Richter, A. M., Klauck, G., Mika, F. & Hengge, R. Microanatomy at cellular resolution and spatial order of physiological differentiation in a bacterial biofilm. *mBio* **4**, e00103–13 (2013). Publisher: American Society for Microbiology.

Name	Relevant genotype and phenotype[1]	Source
Plasmid		
pTrc99a	Expression vector; pBR ori, pTrc promoter, IPTG inducible; AmpR	[91]
pOB2	*mCherry* in pTrc99A, IPTG inducible; AmpR	[83]
pVM42 (pOB35)	*mNeonGreen* in pTrc99A, IPTG inducible; AmpR	[83]
pOSK237	*mCherry* in pBAD, Arabinose inducible; KanR	[73]
pLeoL6	$cheY^{D13K,Y106W}$ in pTrc99a, IPTG inducible; AmpR	[68]
pSIJ8	pKD46, rhaRS-prha-FLP, AmpR	[82]
Strain		
W3110RH	W3110 derivative with functional RpoS, *rpoS396(Am)*	[92]
RC101 (WT)	W3110 Δflu pOB2; AmpR	This study
RC125 ($\Delta cheY$)	W3110 $\Delta cheY$ Δflu pOB2; AmpR	This study
VS1211 ($\Delta cheZ$)	W3110 $\Delta cheZ$	[83]
RC113 ($\Delta fliC$)	W3110 $\Delta fliC$ Δflu pOB35; AmpR	This study

1.AmpR: ampicillin-resistant, KanR: kanamycin resistant.

Table 1: Plasmids and strains used in this study.

Extended Data

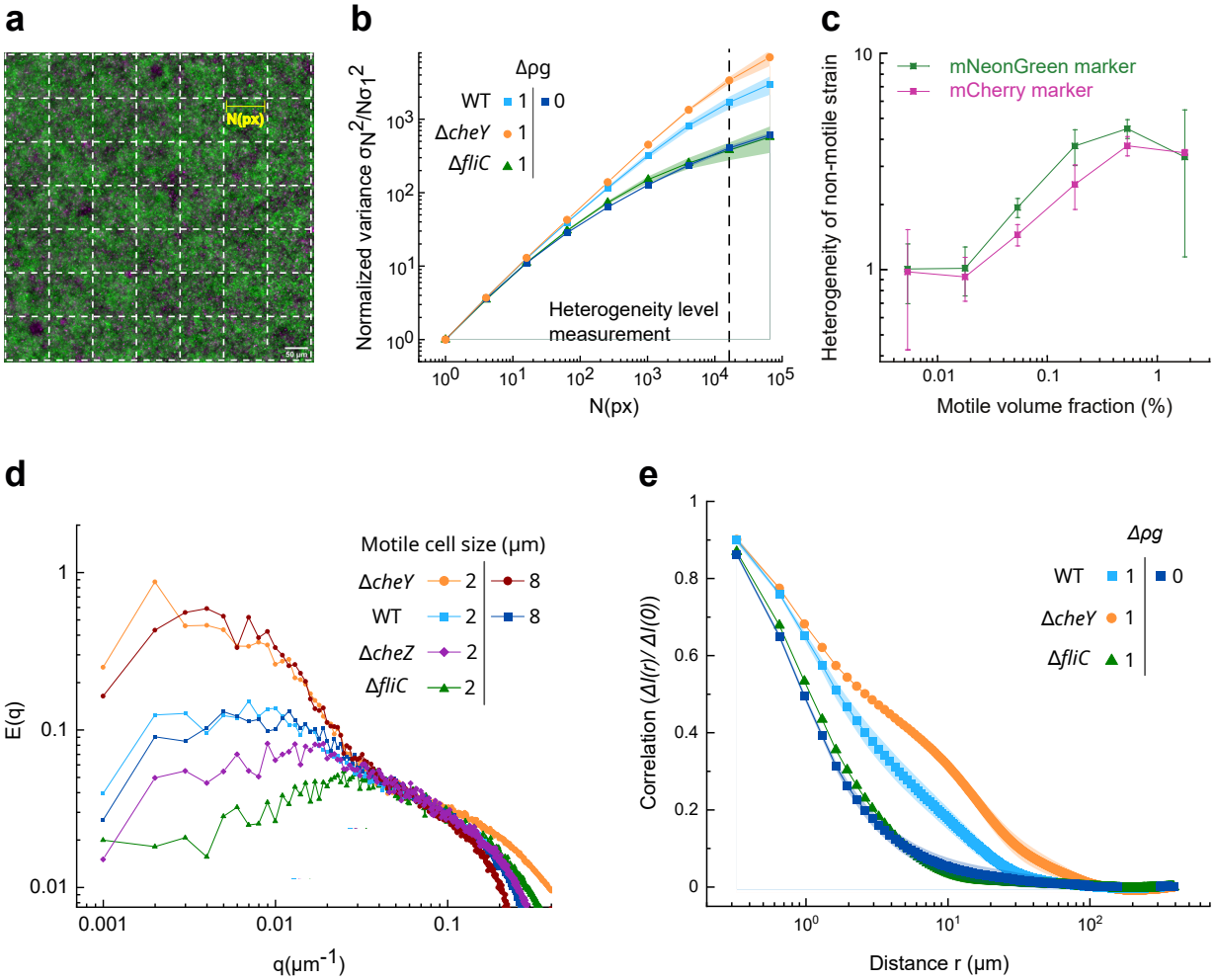

Figure S1: Characterization of the density fluctuations a) Fluorescent image divided into non-overlapping squares of area N px^2. b) Normalized variance of the fluorescence intensity of the non-motile strain $\sigma_N^2/N\sigma_1^2$ as a function of the number N of pixels of the division square, in the binary mixture $\phi_{NM} = 5.3\%$, $\phi_M = 0.18\%$. The dotted line represents the value of N for which heterogeneity level is calculated. c) Comparison of the heterogeneity level as a function of the density of motile cells, for a fixed non-motile cell density $\phi_{NM} = 5.34\%$, obtained when the non-motile cells are marked with either the mNeonGreen or the mCherry marker as indicated, the motile cells being marked with the other one. d) Structure factor E(q) of the non-motile strain in the binary mixture $\phi_{NM} = 5.3\%$, $\phi_M = 0.18\%$ with the different strains used in this study. e) Spatial autocorrelation function of the fluorescence intensity for the non-motile cells in the binary mixture $\phi_{NM} = 5.3\%$, $\phi_M = 0.18\%$, featuring a short scale decay around 1 μm, corresponding to the cell size, and a second non-exponential decay at $\sim$ 10-50 μm length scales in heterogeneous samples. b-e) The data points consist of the geometric mean of three biological replicates and error areas or bars represent the standard deviation.

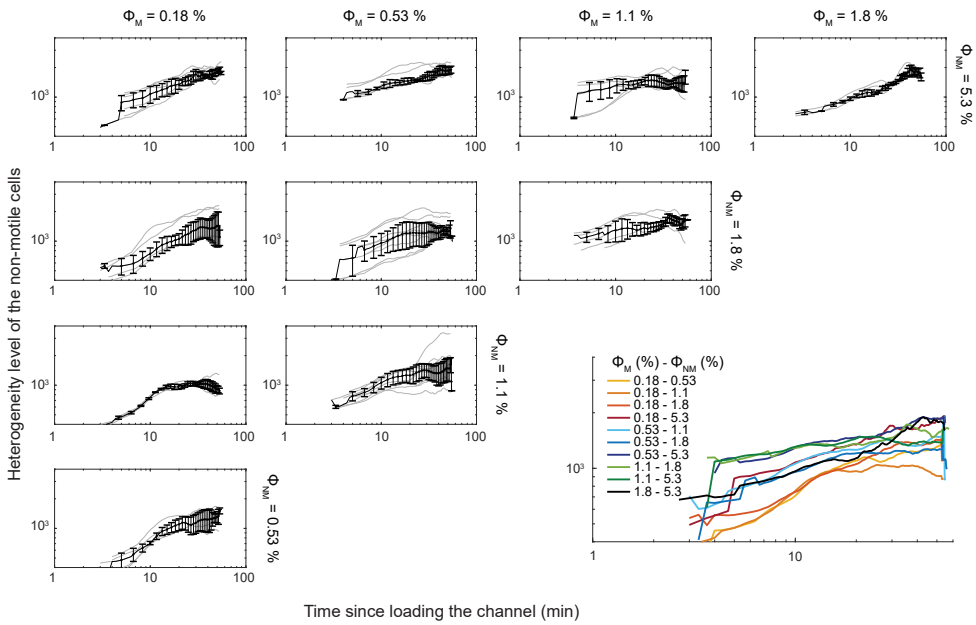

Figure S2: Dynamics of emergence of the density fluctuations. The heterogeneity level of the non-motile strain of the mixture is shown as a function of the time since the mixture was loaded in the channel of the microfluidic chip, for different combinations of volume fractions of motile and non-motile cells. Displayed is the geometric mean of three independent replicates and error bars represent the standard deviation.

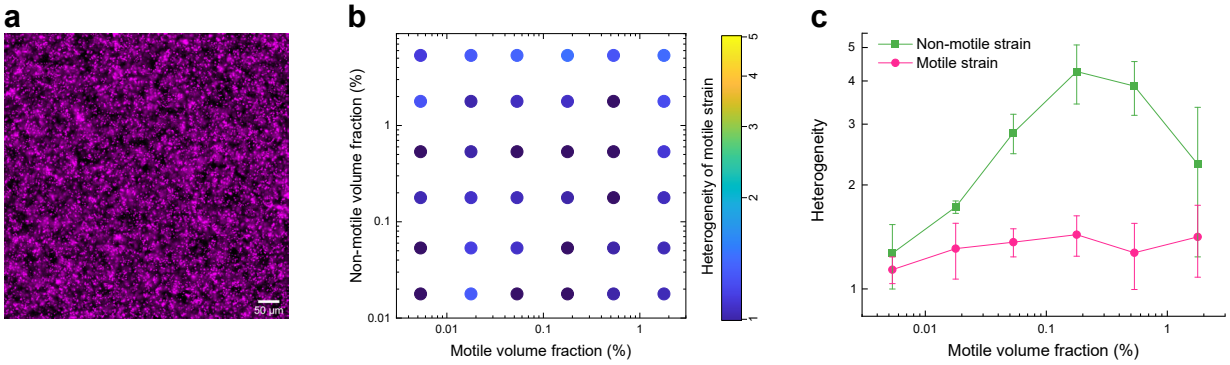

Figure S3: Behavior of the motile strain in the binary mixture. a) Microscopy image of the motile strain (WT *E.coli* mCherry) the binary mixture $\phi_{NM} = 5.3\%$, $\phi_M = 0.18\%$. b) Phase diagram of the density fluctuations of the motile strain in binary mixtures of bacterial populations as a function of the volume fraction (in %) of each strain. The heterogeneity level, which quantifies density fluctuations, is represented by a color scale. Each point represents the geometric mean of three independent replicates. c) Comparison between the heterogeneity level of the motile and the non-motile strains as a function of the volume fraction of the motile strain. The volume fraction of the non-motile strain is kept constant at 5.34%. Each point represents the geometric mean of three independent replicates and error bars indicate the standard deviation.

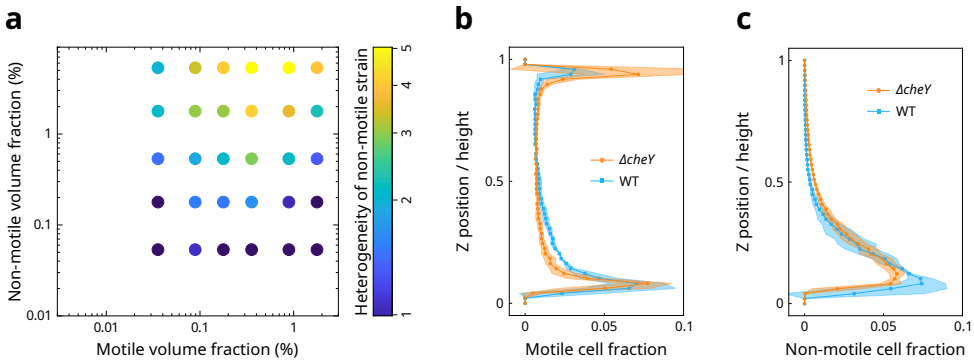

Figure S4: Effect of chemotaxis and tumbling. a) Phase diagram of the density fluctuations of the non-motile strain in binary mixtures with *cheY*** as the motile strain. The quantification of the density fluctuations as heterogeneity level is represented by a color scale. Each point represents the mean of three replicates. b-c) Motile (b) and Non-motile (c) cell fraction (abscissa), measured by cell counting in confocal microscopy, as a function of the vertical position in a channel of 50 μm of height (ordinate), in the binary mixtures with ϕ_{NM} = 5.3% and ϕ_M = 0.18%, with the Δ*cheY* and WT motile strains. Each point represents the mean and error area the standard deviation over three replicates.

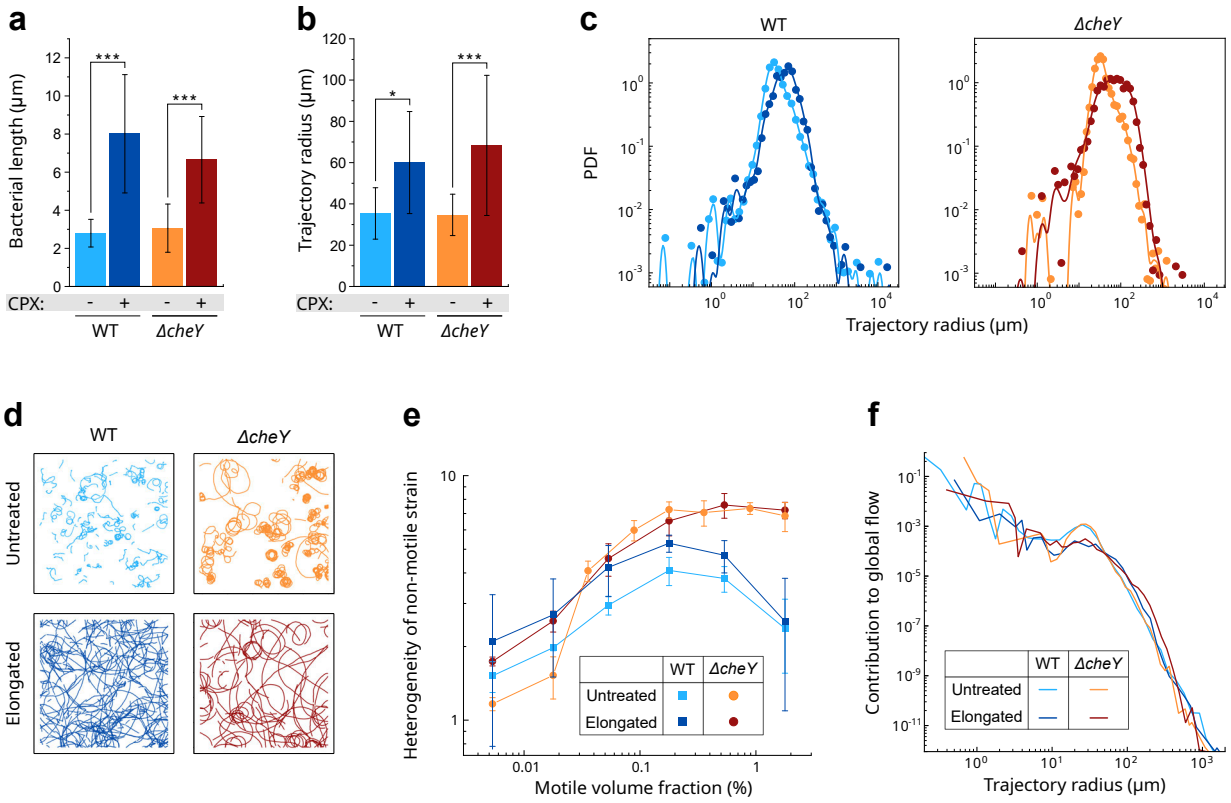

Figure S5: Effect of cell elongation. a-b) Median (a) bacterial cell length and (b) trajectory radius described by the motile cells (WT: run and tumble, $\Delta cheY$: non-tumbling), untreated and elongated, swimming close to the surface. The error bars represent the median absolute deviation. Welch's t-test: $*p < 0.05$, $***p < 0.0001$. c) Probability distribution of the radii from the circular trajectories. The radii are weighted by the trajectory length. a-c) The cell length and the trajectory radius were measured as described in the Methods section. Displayed are more than of 14000 individual swimming trajectories. d) Swimming trajectories of untreated and elongated cells, both the run and tumble strain (WT) and the non-tumbling strain ($\Delta cheY$). e) Comparison between the heterogeneity level of the elongated and untreated cells in both strains (WT and $\Delta cheY$). The volume fraction of the non-motile strain is kept fixed to 5.34% and the volume fraction of the motile strain is systematically varied. Each point represents the geometric mean of three independent replicates and error bars indicate the standard deviation. f) Estimation of the contributions to global flows of a cell swimming with a given radius of curvature. The estimation is computed as $p(R)/R^3$, i.e. the product of the probability $p(R)$ of an observing radius R by the typical magnitude of the fluid velocity produced by the circular swimmer for this radius ($\langle |v_{1r}| \rangle \propto 1/R^3$).

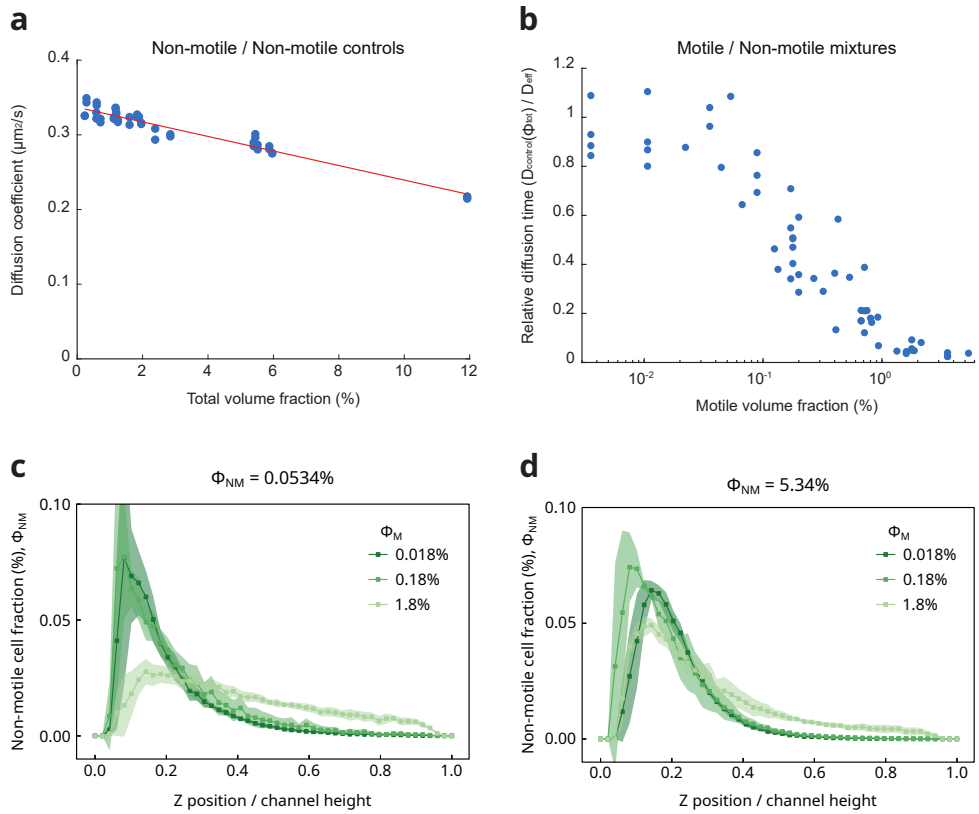

Figure S6: Effect of the motile cells on the diffusion of the non-motile cells. a-b) Diffusion of the non-motile cells measured using Differential Dynamic Microscopy. a) Diffusion coefficient of the non-motile cells in absence of motile cells (control) as a function of the total volume fraction of cells in the population. The straight line is a linear fit. b) Inverse of the diffusion coefficient of the non-motile cells (normalized by its value in the non-motile only case at the same density) as a function of the volume fraction of run and tumble (WT) motile cells in the mixtures. The increase in the diffusion of the non-motile cells becomes noticeable for volume fractions of motile cells above 0.03%. c-d) Non-motile cell fraction (abscissa), measured by cell counting in confocal microscopy, as a function of the vertical position in a channel of 50 μm of height (ordinate), for varying volume fractions of the motile strain (WT) in the mixture, in the cases of a low (c) or a high (b) total volume fraction of non-motile cells.

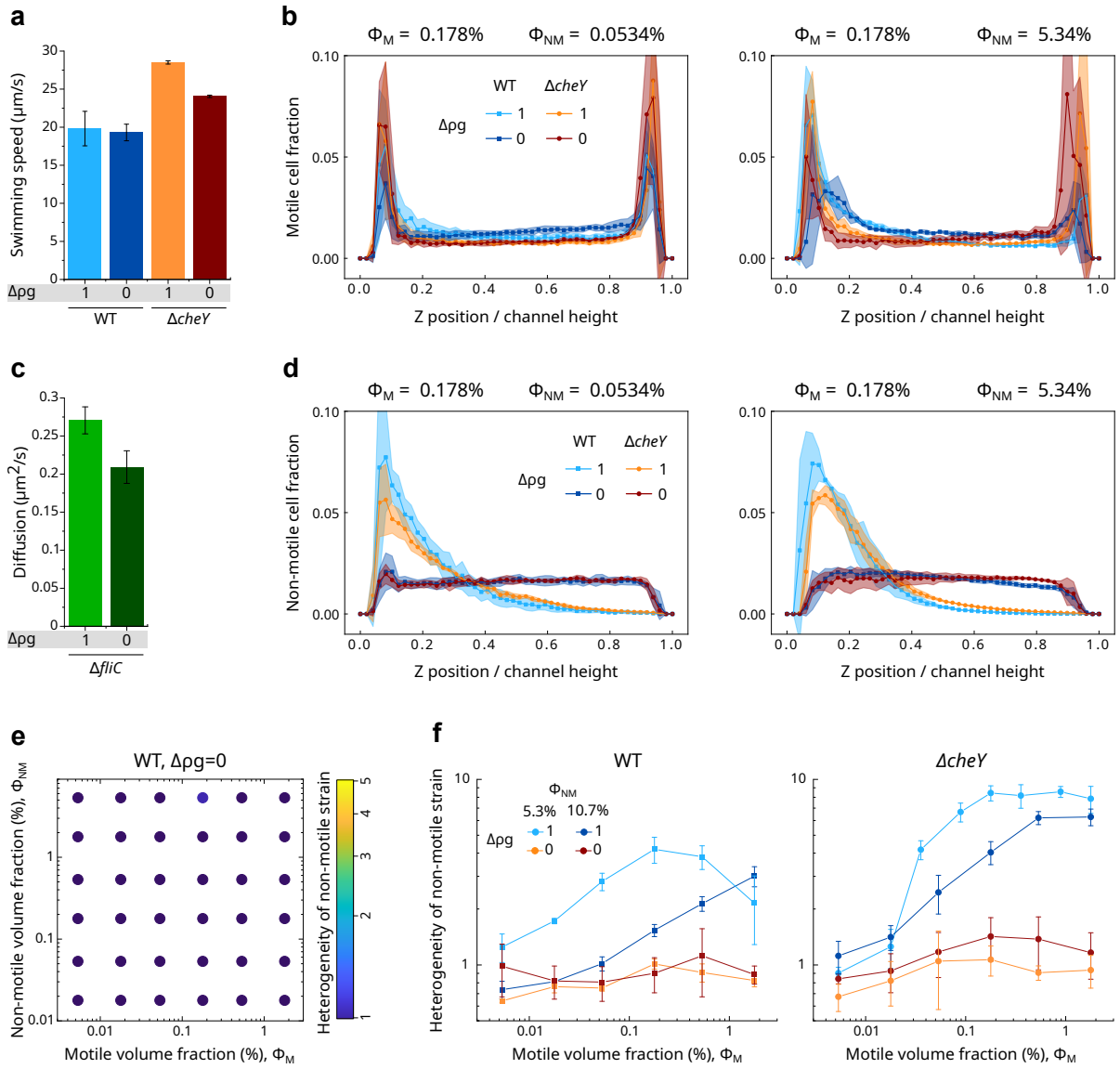

Figure S7: Effect of Iodixanol and neutral buoyancy. (a-d) Effect of Iodixanol on cell motion and position. (a) Swimming speed of motile cells (WT and $\Delta cheY$) and (c) diffusion coefficient of non-motile cells alone ($\Delta fliC$) were measured using Differential Dynamic Microscopy in regular motility buffer ($\Delta\rho g = 1$) and in neutrally buoyant medium, i.e. motility buffer with 26.6% of Iodixanol ($\Delta\rho g = 0$). (b, d) Vertical localization of the (b) motile and (d) non-motile cells in the binary mixture $\phi_{NM} = 5.3\%$, $\phi_M = 0.18\%$ and in a channel of 50 μm of height. The type of motile cells (WT or $\Delta cheY$) and the suspending medium conditions ($\Delta\rho g = 1$ or 0) are as indicated. The cell count, measured from confocal microscopy Z-stacks, was normalized to total cell number to obtain the cell fraction, shown as a function of the Z position. e) Phase diagram of the density fluctuations of the non-motile strain (volume fraction in ordinate) in mixture with the WT motile strain (volume fraction in abscissa) in density-matched suspending media ($\Delta\rho g = 0$). The heterogeneity level, which quantifies density fluctuations, is represented by a color scale. f) Effect of increased non-motile volume fraction (ϕ_{NM}) on the density fluctuations. The heterogeneity of the non-motile strain is plotted against the volume fraction of the motile strain for two ϕ_{NM}; 5.3 % and 10.7 %, for two sedimentation regimes; $\Delta\rho g = 1$ and $\Delta\rho g = 0$, and for two motile strains; WT and $\Delta cheY$. a-f) Each point represents the mean (a,c), binned mean (b,d) or geometric mean (e-f) of three independent replicates and error bars or areas indicate the standard deviation.

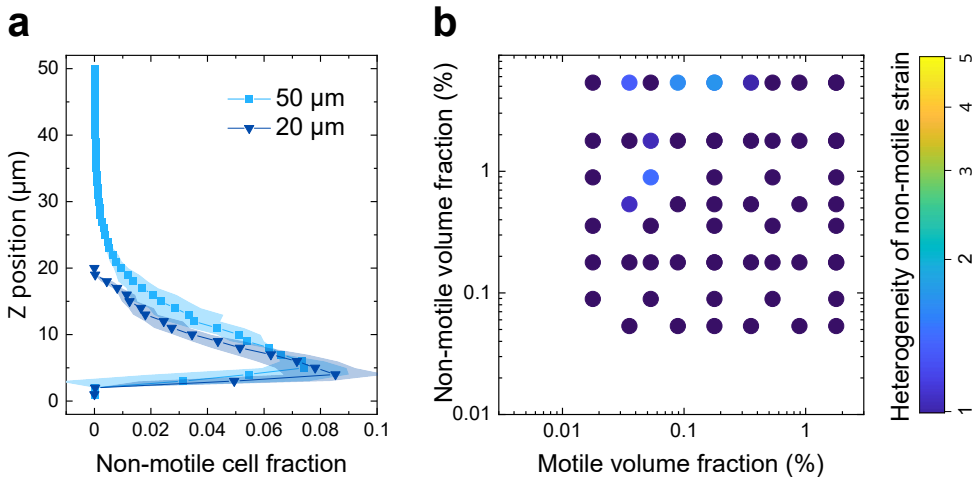

Figure S8: Effect of vertical confinement in a $20\,\mu$m high channel a) Vertical distribution of non-motile cells in the binary mixture $\phi_{NM} = 5.3\,\%$, $\phi_M = 0.18\,\%$, with a WT motile strain, for two channel heights ($h = 50\,\mu m$ and $h = 20\,\mu m$). b) Phase diagram of the density fluctuations of the non-motile strain in binary mixtures with WT as the motile strain in channels of $20\,\mu m$ in height. The quantification of the density fluctuations as heterogeneity level is represented by a color scale. Each point represents the mean of three replicates.

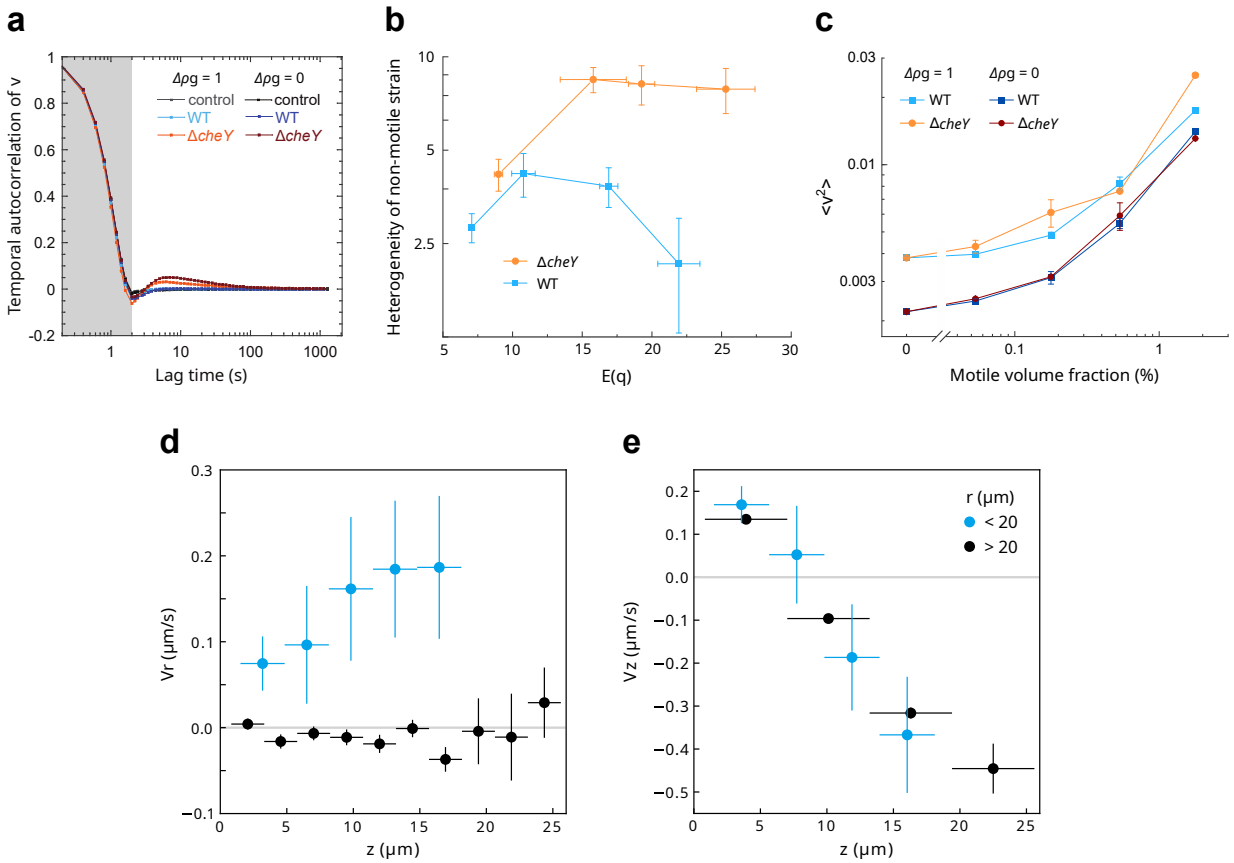

Figure S9: Dynamics of the non-motile cells in the density fluctuations. a) Temporal autocorrelation of the non-motile cell velocity in the binary mixture $\phi_{NM} = 5.3\%$, $\phi_M = 0.53\%$ with the indicated motile cells (WT, $\Delta cheY$ and control ($\Delta fliC$)) and buoyancy ($\Delta\rho g = 1$ and $\Delta\rho g = 0$). The grey area indicates the width of the sliding window that was used to estimate the velocity. The non-motile cells' velocity shows long-time ($t > 5$ s) positive correlations only in presence of $\Delta cheY$ motile cells. b) The heterogeneity level of the non-motile strain in mixture with two different motile strains, WT and $\Delta cheY$, as a function of the peak of the flow structure factor E(q) (Fig. 4b) that quantifies the amplitude of the flow patterns. The volume fraction of the non-motile strain is kept to $\phi_{NM} = 5.3\%$ and the one of the motile strain increases with E(q) (Fig. 4b). Each point represents the mean over a triplicate of the peak of E(q) with the standard deviation as the error bars, and the geometric mean of the heterogeneity level with the standard deviation as the error bars. c) Kinetic energy $\langle v^2 \rangle$ of the non-motile strain as a function of the motile volume fraction ϕ_M for mixtures with different motile strains (WT and $\Delta cheY$) and buoyancies ($\Delta\rho g = 1$ and $\Delta\rho g = 0$). d) Vertical (V_z) and e) radial (V_r) velocity relative to the center of the depletion zones (hole) of the non-motile tracer cells as a function of their vertical position (z), for cells that are less (blue), or more (black), than 20 μm away from the hole center (r). Each data point represents the binned mean, the vertical error bars represent the standard error, and the horizontal error bars represent the bins over which the mean was calculated. The data comprises more than 1500 individual 3D trajectories.

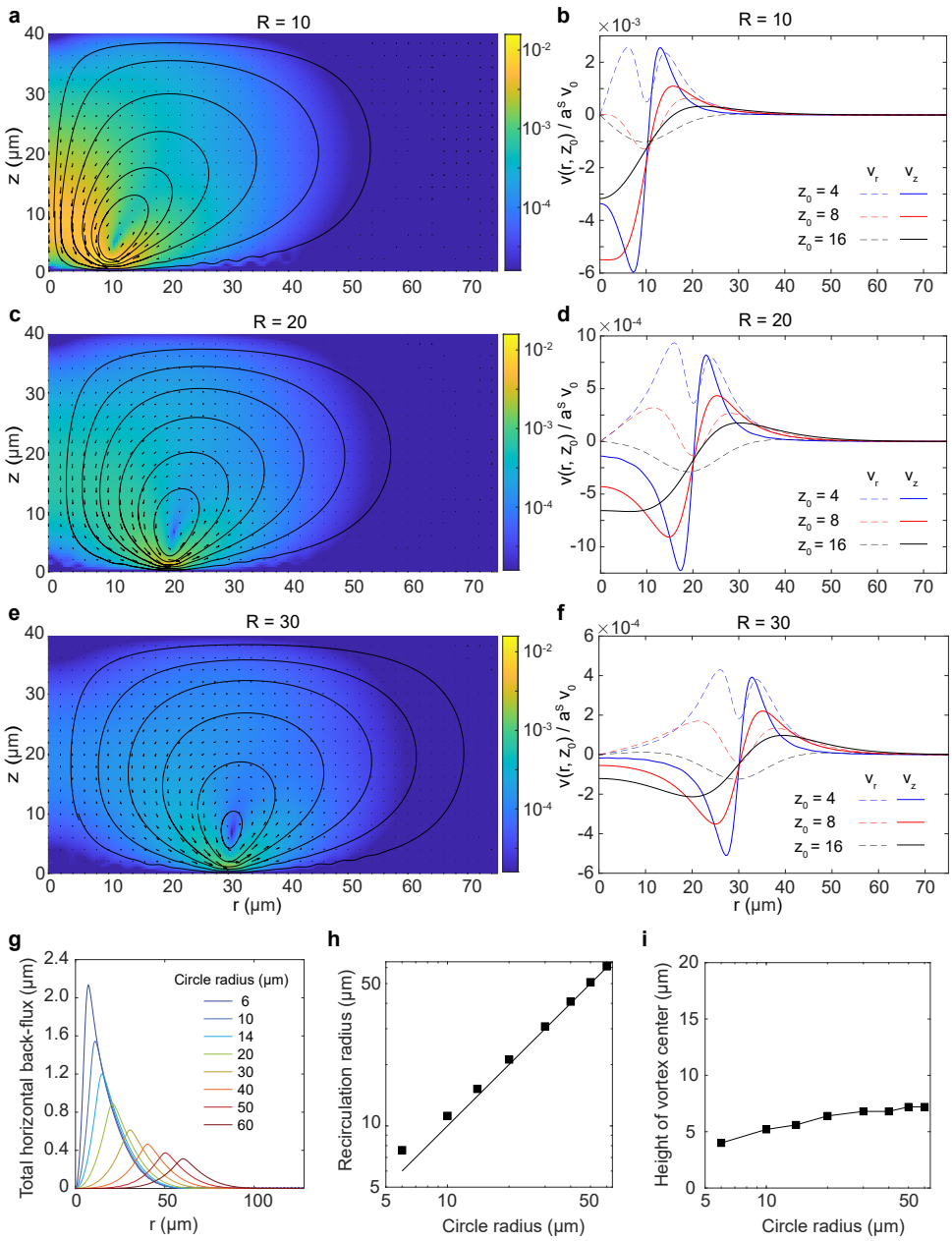

Figure S10: Computed flow field created by a circular swimming dipole for several circle radii R, situated $z_M = 1\,\mu$m above the bottom in a channel of height $H = 40\,\mu$m. a-f, Shown are the flow fields and some streamlines as a function of the horizontal distance to the center of the circle r and the height within the channel z (a,c,e), and the normalized vertical(v_z) and radial (v_r) velocities at the indicated heights z_0 (b,d,f), for the indicated radii. The normalized velocity $v(r,z)/a^S v_0$ is in μm^{-1}. g, Total back-flux $\Phi(r)/a^S v_0$ (equation 11) through a cylinder of radius r centered on the circle of indicated radius (R). The back-flux is maximum for r corresponding to the position of the eye of the recirculation vortex, called the recirculation radius. The maximum is the recirculation flux plotted in Fig. 5c. h, Recirculation radius and i, height of the eye of the recirculation vortex, as a function of the radius of the circle.

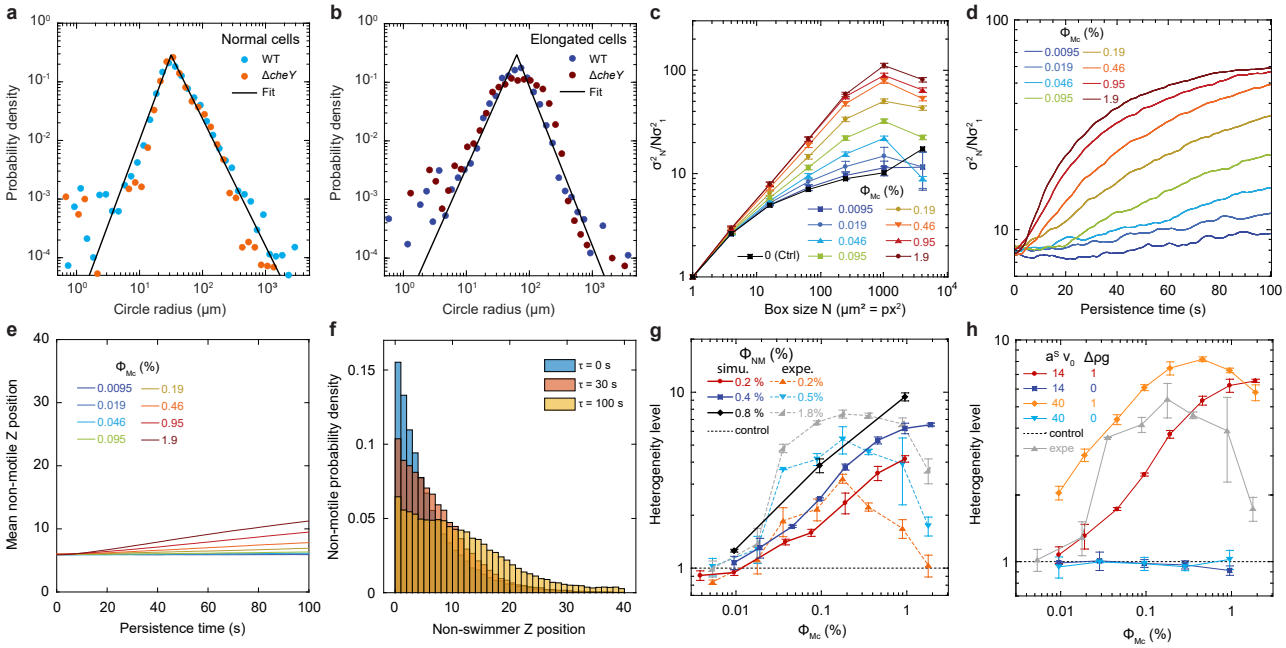

Figure S11: Extended data to the simulations of mixtures of motile circular swimmers (Mc) and non-motile tracer cells (NM). a-b, Comparison between the ansatz probability density of a circle having a given radius (Fit) and the experimental distributions for normal (a) and elongated cells (b). c, Heterogeneity measure $\sigma_N^2/N\sigma_1^2$ (equation 2) of the NM cells (volume fraction $\phi_{NM} = 0.4\,\%$) as a function of the box size N for the indicated volume fractions of motile circular swimmers of normal size (ϕ_{Mc}), averaged over the persistence time range $\tau = [90, 100]$ s. d, Heterogeneity measure $\sigma_N^2/N\sigma_1^2$ as a function of the time since the start, from a homogeneous state, of the simulation (persistence time) for indicated ϕ_{Mc} for $N = 256$. The heterogeneity level is the average of $\sigma_N^2/N\sigma_1^2$ over $\tau = [90, 100]$ s for $N = 256$. e, Average vertical position of the NM cells ($\phi_{NM} = 0.4\,\%$) throughout the simulation for indicated ϕ_{Mc}. The increase for the largest ϕ_{Mc} reflect the increased effective diffusion. f, Vertical distribution of non-motile cells for the indicated persistence times for the $\phi_{Mc} = 1.9\,\%$ simulation. **g**, Heterogeneity level as a function of the volume fraction of motile circular swimmers ϕ_{Mc}, for different values of the volume fraction of non-motile cells (ϕ_{NM}), compared to rescaled values (Method) for experiments with the non-tumbling mutant (expe.). Control is a homogeneous distribution of cells. h, Effect of increasing the dipolar force $f_0/6\pi\eta = a^S v_0$, all things being otherwise equal. Simulation in presence ($\Delta\rho g = 1$) and absence ($\Delta\rho g = 0$) of gravity are shown. Expe. corresponds to experiments with the non-tumbling motile cells and $\phi_{NM} = 0.5\,\%$.

Extended Movie 1: Dynamics of the non-motile strain (mNeonGreen, volume fraction $\phi_{NM} = 5\,\%$) in mixture with a motile WT strain (volume fraction $\phi_M = 0.18\,\%$). Scale bar: $50\,\mu$m. The movie is 60 times accelerated.

Extended Movie 2: Time evolution of the formation of the density fluctuation in a binary mixture of a motile WT strain (mCherry, volume fraction $\phi_M = 0.5\,\%$) and a non-motile strain (mNeonGreen, volume fraction $\phi_{NM} = 1.8\,\%$). Scale bar: $50\,\mu$m. The movie is 200 times accelerated.

Extended Movie 3: Dynamics of the non-motile strain (mNeonGreen, volume fraction $\phi_{NM} = 5\,\%$) in mixture with a motile $\Delta cheY$ strain (volume fraction $\phi_M = 0.18\,\%$). Scale bar: $50\,\mu$m. The movie is 60 times accelerated.

Supplementary notes: Extended simulation methods

Brownian dynamics simulations

Our simulations use a truncated and adapted version of the equation used for Brownian dynamics simulations to compute the time evolution of the positions of the mixture of motile and motile cells. The full Brownian dynamics of a set of N particles in a fluid otherwise at rest is given by :

$$d\mathbf{r}_i = \left[\frac{\mathbf{f}_i}{\zeta_i} + \sum_j \mathbf{\Omega}^{i,j} \cdot \frac{\mathbf{f}_j}{6\pi\eta} + \frac{k_B T}{6\pi\eta} \sum_j \frac{\partial}{\partial \mathbf{r}_j} \mathbf{\Omega}^{i,j} \right] dt + \sqrt{2}\mathbf{B}^{i,j} \cdot d\mathbf{w}_j \tag{15}$$

with dt the time step, $\mathbf{r}_i$ the positions of the particles indexed by i, $\mathbf{f}_i$ the forces on particle i, that it will also exert on the fluid, ζ_i the Stokes drag coefficient, η the fluid viscosity, $\mathbf{\Omega}^{i,j}$ the hydrodynamic-interaction tensor, and $\sqrt{2}\mathbf{B}^{i,j} \cdot d\mathbf{w}_j$ the Brownian term including random local perturbations $d\mathbf{w}_j = \sigma_j \sqrt{dt}$, with σ_i a Gaussian random variable with zero mean and unit variance. The tensor $\mathbf{B}^{i,j}$ is coupled to the hydrodynamic interactions via the fluctuation dissipation theorem $k_B T \left(\mathbf{I}/\zeta_i + \mathbf{\Omega}^{i,j}/6\pi\eta \right) = \mathbf{B}^{i,j} \cdot \left(\mathbf{B}^{i,j} \right)^T$.

Approximations

The full Brownian dynamics simulations of the binary mixtures would imply solving this set of equations for the motile and non-motile particles together. In our simple scheme, we only solve for the positions of the non-motile particles, the positions of the motile ones being an ansatz as explained in the Methods section. Furthermore, we simplify the treatment of hydrodynamic interactions.

The hydrodynamic-interaction tensor $\mathbf{\Omega}^{i,j}$ accounts for hydrodynamic interactions between particles, via its non-diagonal terms $\mathbf{\Omega}^{i,j}$ with i, j two particle indices, as well as for direct interactions with the boundaries of the liquid domain (in principle relevant in our case) through its diagonal terms $\mathbf{\Omega}^{i,i}$. We neglect the latter diagonal terms, $\mathbf{\Omega}^{i,i} = 0$, considering that the values of the diffusion coefficient of non-motile cells and the swimming speeds of the motile ones were measured within the $40 - 50\mu$m high channel, and thus already account for the mean increase in friction due to confinement. We further make a two-body approximation for the cross-particle interaction, i.e. $\mathbf{\Omega}^{i,j \neq i}$ depends only on the positions of i and j. In this case, the divergence of the interaction tensor vanishes, $\frac{\partial}{\partial \mathbf{r}} \mathbf{D} = 0$. We also neglect the hydrodynamic interactions between non-motile particles, considering only the terms $\mathbf{\Omega}^{i,j} \cdot \mathbf{f}_j / 6\pi\eta$ for which i is non-motile and j is the flagellum or cell body of a motile cell, modeled as a force dipole. As explained in Methods, we also consider only the average flow created by the dipoles when they perform one circular revolution. We further neglect the effects of hydrodynamic interactions in the Brownian term, which thus reduces to $\mathbf{B}^{i,j} = \mathbf{I}\sqrt{D_{NM}}$, with $D_{NM} = k_B T/\zeta$ the diffusion coefficient of the non-motile cells and $\mathbf{I}$ the identity matrix. Finally, the only bulk force on the non-motile cells is buoyancy-corrected gravity, since we neglect steric interactions between non-motile cells.

The simplified equation of motion for the non-motile cell i therefore reduces down to:

$$d\mathbf{r}_i = \left[\mathbf{v}_{sed} + \sum_{dipoles} \left\langle \mathbf{\Omega}^{i,j} \cdot \frac{\mathbf{f}_j}{6\pi\eta} \right\rangle_{1\mathrm{r}} \right] dt + \sqrt{2D_{NM}} d\mathbf{w}_i \tag{16}$$

with $\mathbf{v}_{sed} = \Delta m \mathbf{g}/\zeta$ the sedimentation speed, and the indices j run on the flagella and cell bodies of motile cells.

Hydrodynamic interactions

We focus in the following on the interparticle hydrodynamic interaction term of the equation of motion (16):

$$d\mathbf{r}_i^{\mathrm{Hydro}} = \left[\mathbf{\Omega}^{i,j} \cdot \frac{\mathbf{f}_j}{6\pi\eta} \right] dt \tag{17}$$

Although the interparticle hydrodynamic interaction tensor $\mathbf{\Omega}^{i,j}$ is simply given by the Oseen tensor in unbound space, $\mathbf{\Omega}^{i,j} = 3/4 \left(\mathbf{I} + \mathbf{r}_{ij}\mathbf{r}_{ij}/r_{ij}^2 \right)/r_{ij}$, its computation is significantly more difficult and expensive in the case of confined geometries. For this, we follow an algorithm first introduced by Mucha *et al.*[71] and further developed by Hernandez-Ortiz *et al.*[72] for our geometry with periodic boundary conditions in x and y and no-slip conditions on the top and bottom walls, which allows computation in $N \log N$ time scaling. We briefly remind here the principle of the algorithm and explain the details of our implementation.

In the following, we write $\mathbf{f}_j/6\pi\eta$ as "$\mathbf{f}_j$" to simplify the expressions. By definition, the flow created on particle i situated in $\mathbf{x}^i$ by the point force $\mathbf{f}^j$ in $\mathbf{x}^j$ is $\mathbf{u}^{i\leftarrow j} = \mathbf{\Omega}(\mathbf{x}^i, \mathbf{x}^j)\mathbf{f}^j$. Expliciting the vector and tensor components, with $\alpha, \beta = 1, 2, 3$ the indices for the dimensions of space, we have:

$$u_\alpha^{i\leftarrow j} = \Omega_{\alpha,\beta}(\mathbf{x}^i, \mathbf{x}^j) f_\beta^j \tag{18}$$

with an implicit sum over spatial indices. In our geometry, the interaction tensor can be written as a Fourier series:

$$\Omega_{\alpha,\beta}(\mathbf{x}^i, \mathbf{x}^j) = \sum_{\mathbf{k}} \hat{\omega}_{\alpha,\beta}(\mathbf{k}, x_3^i, x_3^j) e^{i\mathbf{k}(\mathbf{x}_{2D}^i - \mathbf{x}_{2D}^j)} \tag{19}$$

with $\mathbf{k} = (l, m)$ the wave vector and $\mathbf{x}_{2D} = (x_1, x_2)$ the position in the horizontal plane. The vertical position x_3 is confined between the two walls in $x_3^{\text{walls}} = \pm h$, so that the channel height is $H = 2h$. The Fourier-transformed interaction tensor $\hat{\omega}_{\alpha,\beta}(\mathbf{k}, x_3^i, x_3^j)$ is the Green function that is solution of the partially Fourier-transformed Stokes equation for a point source:

$$6\pi i\, l\, \hat{p}_\beta + \mathbf{k}^2\, \hat{\omega}_{1,\beta} - \frac{\mathrm{d}^2}{\mathrm{d}x_3^2}\hat{\omega}_{1,\beta} = \frac{6\pi}{W^2}\delta(x_3^i - x_3^j)\delta_{1,\beta} \tag{20a}$$

$$6\pi i\, m\, \hat{p}_\beta + \mathbf{k}^2\, \hat{\omega}_{2,\beta} - \frac{\mathrm{d}^2}{\mathrm{d}x_3^2}\hat{\omega}_{2,\beta} = \frac{6\pi}{W^2}\delta(x_3^i - x_3^j)\delta_{2,\beta} \tag{20b}$$

$$\frac{\mathrm{d}}{\mathrm{d}x_3}\hat{p}_\beta + \mathbf{k}^2\, \hat{\omega}_{3,\beta} - \frac{\mathrm{d}^2}{\mathrm{d}x_3^2}\hat{\omega}_{3,\beta} = \frac{6\pi}{W^2}\delta(x_3^i - x_3^j)\delta_{3,\beta} \tag{20c}$$

$$il\, \hat{\omega}_{1,\beta} + im\, \hat{\omega}_{2,\beta} + \frac{\mathrm{d}}{\mathrm{d}x_3}\hat{\omega}_{3,\beta} = 0 \tag{20d}$$

Here, $\hat{p}_\beta$ is the normalized contribution to the pressure field of the force component in direction β, W^2 is the horizontal area of the simulation box, the derivatives are taken relative to x_3^i, and Eq. (20d) is the incompressible flow constraint.

The gist of the contribution of Mucha *et al.*[71] is to realize that there is an analytical solution to this equation, which can be factorized in the following form:

$$\hat{\omega}_{\alpha,\beta}(\mathbf{k}, x_3^i, x_3^j) = \hat{c}_{\alpha,\gamma}^b(\mathbf{k}, x_3^i)\hat{a}_{\gamma,\beta}^b(\mathbf{k}, x_3^j) \tag{21}$$

where the Boolean index b indicates whether the point where the flow is calculated lies below the source of flow, i.e. $(x_3^i < x_3^j) \equiv (b = 1)$. The tensors $\hat{a}$ and $\hat{c}$ indeed have different analytical expressions in each case, which converge to the same well-defined value when $x_3^i \to x_3^j$. The analytical expressions of the tensors are given in a separate paragraph below. The contribution of the point force f_β^j in j to the flow in i can thus be written as:

$$u_\alpha^{i\leftarrow j} = \sum_{\mathbf{k}} \left[\hat{c}_{\alpha,\gamma}^b(\mathbf{k}, x_3^i) e^{i\mathbf{k}\mathbf{x}_{2D}^i}\right] \left[\hat{a}_{\gamma,\beta}^b(\mathbf{k}, x_3^j) e^{-i\mathbf{k}\mathbf{x}_{2D}^j} f_\beta^j\right] \tag{22}$$

and the total flow produced by the other particles on i can be written as a product of contributions that respectively depend only on the source and the target:

$$u_\alpha^i = \sum_{\mathbf{k}} \left[\hat{c}_{\alpha,\gamma}^1(\mathbf{k}, x_3^i) e^{i\mathbf{k}\mathbf{x}_{2D}^i}\right] \sum_{x_3^j > x_3^i} \left[\hat{a}_{\gamma,\beta}^1(\mathbf{k}, x_3^j) e^{-i\mathbf{k}\mathbf{x}_{2D}^j} f_\beta^j\right] + \sum_{\mathbf{k}} \left[\hat{c}_{\alpha,\gamma}^0(\mathbf{k}, x_3^i) e^{i\mathbf{k}\mathbf{x}_{2D}^i}\right] \sum_{x_3^j < x_3^i} \left[\hat{a}_{\gamma,\beta}^0(\mathbf{k}, x_3^j) e^{-i\mathbf{k}\mathbf{x}_{2D}^j} f_\beta^j\right] \tag{23}$$

We now describe the algorithm allowing the fast calculation of these sums for all particles[72]. At each time step, all the necessary coefficients of $\hat{a}$ and $\hat{c}$ are computed. The particles are sorted according to their vertical position. The algorithm initializes a vector $\Lambda_\gamma^1(\mathbf{k})$, which will contain all the contribution of the cells above the current one,

$$\Lambda_\gamma^1(\mathbf{k}) = \sum_{x_3^j > x_3^i} \left[\hat{a}_{\gamma,\beta}^1(\mathbf{k}, x_3^j) e^{-i\mathbf{k}\mathbf{x}_{2D}^j} f_\beta^j\right] . \tag{24}$$

The particles are scanned from top to bottom. For each particle i, the contribution to the flow in $\mathbf{x}^i$ from the particles above it is computed as:

$$u_\alpha^{i,1} = \sum_{\mathbf{k}} \left[\hat{c}_{\alpha,\gamma}^1(\mathbf{k}, x_3^i) e^{i\mathbf{k}\mathbf{x}_{2D}^i}\right] \Lambda_\gamma^1(\mathbf{k}) , \tag{25}$$

the $\Lambda^1_\gamma(\mathbf{k})$ is then updated with the contribution of particle i:

$$\Lambda^1_\gamma(\mathbf{k}) \leftarrow \Lambda^1_\gamma(\mathbf{k}) + \left[\hat{a}^1_{\gamma,\beta}(\mathbf{k}, x^i_3) e^{-i\mathbf{k}\mathbf{x}^i_{2D}} f^i_\beta\right] \tag{26}$$

and the next particle is considered. A first top-to-bottom swipe through the particles thus allows to compute the $b = 1$ contribution to the flow for each particle. A bottom-to-top swipe similarly calculates the $b = 0$ contribution. This allows to compute the hydrodynamic interactions in $n_k N$ time, with n_k the number of wave vectors contributing to the sum over $\mathbf{k}$. Importantly, this sum can be truncated because the large $|\mathbf{k}|$ have vanishing contributions to the sum. The sorting of the cells then imposes a global scaling of the computation time as $N(n_k + \log N)$, which is much faster than the typical $O(N^2)$ scaling for brute force computing hydrodynamic interactions.

In our implementation of the algorithm, the two following modifications to the scheme are made. First, the force f^i_β is replaced by a Gaussian-smoothed force $f^i_\beta e^{-k^2 a_0^2}$, with a_0 a cutt-off length. This is equivalent to replacing the point force by a Gaussian distribution of width a_0 in the horizontal plane. It accelerates the convergence of the sum of the Fourier components by regularizing the slowly converging contributions of some of the terms for which $l = 0$ or $m = 0$, with $\mathbf{k} = (l, m)$, and regularizes the divergences at $\mathbf{x}_i = \mathbf{x}_j$ without affecting the far field solution. This also allows us to consider a further reduced number of components n_k. Second, since the motile cells are located at positions $x^j_3 = \pm(h - z_M)$, we first compute $\left(\Lambda^1_\gamma, \Lambda^0_\gamma\right)$ for the three cases $x^i_3 > (h - z_M)$, $-(h - z_M) < x^i_3 < (h - z_M)$ and $x^i_3 < -(h - z_M)$ before scanning once through the non-motile particles, without the now useless sorting, to compute the total flow field they experience. Finally, dimensional analysis tells us that all distances can be expressed in units of W and that the hydrodynamic interaction tensor $\mathbf{\Omega}(\mathbf{x}^i, \mathbf{x}^j)$ has units of an inverse distance. To avoid numerical errors due to exponential factors in the expressions of $\hat{a}$ coefficients, we define a hydrodynamic pixel size px_H, such that $W = 128\,\mathrm{px}_H$, and express all distances in this unit for the purpose of calculating hydrodynamic interactions, $\mathbf{\Omega}$ being converted back in natural units for solving Eq. 16.

Interaction-tensor coefficients

We remind here the mathematical expressions for the tensors $\hat{a}^b_{\gamma,\beta}(\mathbf{k}, x_3)$ and $\hat{c}^b_{\alpha,\gamma}(\mathbf{k}, x_3)$ that allows computing the hydrodynamic interactions. The analytic expressions of all the coefficients were originally given in [72]. We start with the point force contribution coefficient $\hat{a}^b_{\gamma,\beta}$. Taking into account the symmetries of the problem, for the non-zero wave vectors $\mathbf{k} = (l, m) \neq \mathbf{0}$, it reduces to:

$$\hat{a}^b_{\gamma,\beta}(\mathbf{k}, x_3) = \begin{pmatrix} a^b_{1,1}(l,m) & a^b_{1,2}(m,l) & i\,a^b_{3,1}(l,m) \\ a^b_{1,2}(l,m) & a^b_{1,1}(m,l) & i\,a^b_{3,2}(l,m) \\ i\,a^b_{1,3}(l,m) & i\,a^b_{1,3}(m,l) & a^b_{3,3}(l,m) \end{pmatrix} \tag{27}$$

The six above-defined $a^b_{i,j}$ depend on $b = (0, 1)$ as well as on the position x_3. For wave vectors $\mathbf{k} \neq \mathbf{0}$, the later dependence takes the functional form:

$$a^b_{i,j}(l, m, x_3) = ca^{b,0}_{i,j} \left[\left(ca^{b,1}_{i,j}\, x_3 + ca^{b,2}_{i,j}\right) e^{kx_3} + \left(ca^{b,3}_{i,j}\, x_3 + ca^{b,4}_{i,j}\right) e^{-kx_3}\right] \tag{28}$$

The coefficients $ca^{b,n}_{i,j}$ only depend on $\mathbf{k} = (l, m)$ and b. They can be calculated once at the beginning of the simulation for speed. Because several typing errors were introduced in [72], their correct expressions are given in Table S2.

The tensor of coefficients accounting for the position where the flow is calculated, x^p_3, are given in the case of non-zero wave vectors $\mathbf{k} = (l, m) \neq \mathbf{0}$ by:

$$\hat{c}^b_{\alpha,\gamma}(\mathbf{k}, x^p_3) = \begin{pmatrix} c_{l2} & c_{lm} & -i\,c^b_l \\ c_{lm} & c_{m2} & -i\,c^b_m \\ -i\,c^b_l & -i\,c^b_m & c_{k2} \end{pmatrix} \tag{29}$$

with the six distinct coefficients assuming the following values:

$$c_{l2} = \sinh(kz) + \frac{l^2 z}{k}\cosh(kz) \tag{30a}$$

$$c_{m2} = \sinh(kz) + \frac{m^2 z}{k}\cosh(kz) \tag{30b}$$

$$c_{k2} = \sinh(kz) - \frac{k^2 z}{k}\cosh(kz) \tag{30c}$$

$$c_{lm} = \frac{lm\, z}{k}\cosh(kz) \tag{30d}$$

$$c_l^b = \epsilon^b\, l\, z \sinh(kz) \tag{30e}$$

$$c_m^b = \epsilon^b\, m\, z \sinh(kz) \tag{30f}$$

with $\epsilon^b = 1$ if $b = 1$ and $\epsilon^b = -1$ if $b = 0$, and the variable $z = h + \epsilon^b x_3^p$.

In the case of the contribution of the $\mathbf{k} = \mathbf{0}$ terms to the flow, which express the no-net-global-flow condition, the two tensors are given by:

$$\hat{a}^b_{\gamma,\beta}(\mathbf{k} = \mathbf{0}, x_3) = \delta_{\gamma,\beta}(1 - \delta_{\beta,3}) \begin{pmatrix} a_{0,1} \\ a^b_{0,2} \end{pmatrix} \tag{31}$$

$$\hat{c}^b_{\alpha,\gamma}(\mathbf{k} = \mathbf{0}, x_3^p) = \delta_{\alpha,\gamma} \begin{pmatrix} c_{0,1} & c^b_{0,2} \end{pmatrix} \tag{32}$$

with $\delta_{\alpha,\beta}$ being the Kronecker delta. The two vectors form a scalar product when $\Lambda = \sum_{\text{part}} \hat{a} f$ and $\hat{c}$ are multiplied to compute the flow in Equation (25), such that the $\mathbf{k} = \mathbf{0}$ contribution to the flow for a pair-wise interaction is

$$u_\alpha(\mathbf{k} = \mathbf{0}) = \omega^b_{\alpha,\beta}(\mathbf{k} = \mathbf{0}) f_\beta = (c_{0,1} a_{0,1} + c^b_{0,2} a^b_{0,2})(1 - \delta_{\alpha,3}) f_\alpha\,. \tag{33}$$

Note that the contribution of the vertical component f_3 is zero for $\mathbf{k} = \mathbf{0}$, reflecting the no-flow boundary condition at the walls. Finally, the coefficients $a_{0,n}$ and $c_{0,n}$ are given by:

$$a_{0,1} = -9\pi \frac{h^2 - x_3^2}{4W^2 h^3} \tag{34a}$$

$$a^b_{0,2} = -6\pi \frac{x_3 - \epsilon^b h}{2W^2 h} \tag{34b}$$

$$c_{0,1} = h^2 - (x_3^p)^2 \tag{35a}$$

$$c^b_{0,2} = x_3^p + \epsilon^b h \tag{35b}$$

with $\epsilon^b = 1$ if $b = 1$, $\epsilon^b = -1$ if $b = 0$, h the half-height and W^2 the horizontal area of the simulation box, as defined previously.

Parameter	Value		Unit
Global			
W	256		μm
H	40		μm
dt	0.1		s
Total duration	1000		steps
Motile cells	Wild–type	Elongated	
v_0	20		μm/s
a^S	0.7	1.1	μm
l_d	2	4	μm
z_M	1		μm
V_{Mc}	1	2	μm^3
Radii distribution function	$p(\log R) = \begin{cases} \exp(-\lvert \log R - \mu \rvert / l_1)/(l_1 + l_2) & \text{if } \log R > \mu \\ \exp(-\lvert \log R - \mu \rvert / l_2)/(l_1 + l_2) & \text{if } \log R \leq \mu \end{cases}$		
$\log R$	Basis 10 logarithm of the radius		
μ	1.5	1.8	
l_1	0.20	0.16	
l_2	0.15	0.18	
Non-motile cells			
v_{sed}	0.05 $(g = 1)$, $5\,10^{-5}$ (g=0)		μm/s
D_{NM}	0.30		μm^2/s
V_{NM}	1		μm^3
Hydrodynamics			
Hydrodynamic pixel (px)	2		μm
$n\theta$	100		
n_k	127		
a_0	0.3		px
Pseudo Image			
h	8		μm
dh	4		μm
a_{d}	1.0		μm
Image pixel size	1.0		μm

Table S1: Table of the coefficient values used in the numerical simulation

i,j/n	0	1	2	3	4
			b=1		
1,1	$C_h/(k^2(1-e^{4hk}))$	$kl^2(1+e^{8hk}(1-4kh)+e^{4hk}(-2+4kh))$	$(hkl^2+k^2+m^2)+e^{4hk}(2kl^2h-4h^2k^2(l^2+8m^2)-2(k^2+m^2))+e^{8hk}(-3l^2kh+4h^2k^2l^2+k^2+m^2)$	$kl^2(e^{10hk}+e^{6hk}(-2-4kh)+e^{2hk}(1+4kh))$	$e^{10hk}(hkl^2-k^2-m^2)+e^{6hk}(2kl^2h+4h^2k^2(l^2+8m^2)+2(k^2+m^2))+e^{2hk}(-3hkl^2-4h^2k^2l^2-k^2-m^2)$
1,2	$C_hml/(k^2(e^{4hk}-1))$	$-k+e^{8hk}(-k+4hk^2)+e^{4hk}(2k-4hk^2)$	$(-kh+1)+e^{8hk}(3hk-4h^2k^2+1)+e^{4hk}(-2kh-28h^2k^2-2)$	$e^{10hk}(-k)+e^{2hk}(-k-4hk^2)+e^{6hk}(2k+4hk^2)$	$e^{10hk}(-1-hk)+e^{2hk}(3hk+4h^2k^2-1)+e^{6hk}(-2hk+28h^2k^2+2)$
1,3	C_hl	$1+e^{4hk}(-1-4hk)$	$h+e^{4hk}(-h+4h^2k)$	$e^{6hk}(1)+e^{2hk}(-1+4kh)$	$e^{6hk}(h)+e^{2hk}(-h-4h^2k)$
3,1	C_hl	$1+e^{4hk}(-1+4hk)$	$h+e^{4hk}(-h-4h^2k)$	$e^{6hk}(1)+e^{2hk}(-1-4kh)$	$e^{6hk}(h)+e^{2hk}(-h+4h^2k)$
3,2	C_hm	$1+e^{4hk}(-1+4hk)$	$h+e^{4hk}(-h-4h^2k)$	$e^{6hk}(1)+e^{2hk}(-1-4kh)$	$e^{6hk}(h)+e^{2hk}(-h+4h^2k)$
3,3	C_h	$(-k)+e^{4hk}(k+4hk^2)$	$(1-kh)+e^{4hk}(-3hk-4h^2k^2-1)$	$e^{6hk}(k)+e^{2hk}(-k+4k^2h)$	$e^{6hk}(hk+1)+e^{2hk}(3hk-4h^2k^2-1)$
			b=0		
1,1	$C_h/(k^2(1-e^{4hk}))$	$-kl^2e^{10hk}+e^{6hk}(2kl^2+4k^2l^2h)+e^{2hk}(-l^2k-4l^2k^2h)$	$e^{10hk}(hkl^2-k^2-m^2)+e^{6hk}(2kl^2h+4h^2k^2(l^2+8m^2)+2(k^2+m^2))+e^{2hk}(-3l^2kh-4h^2k^2l^2-k^2-m^2)$	$(-kl^2)+e^{4hk}(2kl^2-4k^2l^2h)+e^{8hk}(-kl^2+4l^2k^2h)$	$(hkl^2+k^2+m^2)+e^{4hk}(2kl^2h-4h^2k^2(l^2+8m^2)-2(k^2+m^2))+e^{8hk}(-3hkl^2+4h^2k^2l^2+k^2+m^2)$
1,2	$C_hml/(k^2(e^{4hk}-1))$	$e^{10hk}k+e^{2hk}(k+4hk^2)+e^{6hk}(-2k-4hk^2)$	$e^{10hk}(-kh-1)+e^{2hk}(3hk+4h^2k^2-1)+e^{6hk}(-2kh+28h^2k^2+2)$	$k+e^{8hk}(k-4hk^2)+e^{4hk}(-2k+4hk^2)$	$(1-hk)+e^{8hk}(3hk-4h^2k^2+1)+e^{4hk}(-2hk-28h^2k^2-2)$
1,3	$-C_hl$	$e^{6hk}(-1)+e^{2hk}(1-4hk)$	$e^{6hk}h+e^{2hk}(-h-4h^2k)$	$(-1)+e^{4hk}(1+4kh)$	$(h)+e^{4hk}(-h+4h^2k)$
3,1	$-C_hl$	$e^{6hk}(-1)+e^{2hk}(1+4hk)$	$e^{6hk}h+e^{2hk}(-h+4h^2k)$	$(-1)+e^{4hk}(1-4kh)$	$(h)+e^{4hk}(-h-4h^2k)$
3,2	$-C_hm$	$e^{6hk}(-1)+e^{2hk}(1+4hk)$	$e^{6hk}h+e^{2hk}(-h+4h^2k)$	$(-1)+e^{4hk}(1-4kh)$	$(h)+e^{4hk}(-h-4h^2k)$
3,3	$-C_h$	$e^{6hk}(k)+e^{2hk}(-k+4hk^2)$	$e^{6hk}(-1-kh)+e^{2hk}(-3hk+4h^2k^2+1)$	$(-k)+e^{4hk}(k+4k^2h)$	$(hk-1)+e^{4hk}(3hk+4h^2k^2+1)$

Table S2: Expressions for the $ca_{i,j}^{b,n}(l,m)$ coefficients, with the constant $C_h = 3\pi\exp(kh)/(kW^2(2e^{4hk}(1+8h^2k^2)-e^{8hk}-1))$ and $\mathbf{k}=(l,m)$, $k=|\mathbf{k}|$ being non-zero.