Methods

Experimental Protocols

Bacterial strains and fluorescent labelling

The clean deletion mutants and the corresponding WT strains[17,31] used in this study were labelled with either CFP or YFP using a Gm^r mini-Tn7 vector[32], using a three-strain mating protocol. Briefly, *P. aeruginosa* colonies were grown overnight at 42°C on an LB (Lennox, 20 g/l, Fisher Scientific) plate containing 1.5% (w/v) agar (Difco brand, BD). These were then mixed with both the mini-Tn7 donor *E. coli* strain and an SM10 λpir *E. coli* helper strain on the surface of a fresh LB agar plate. The resulting mixed three strain plate was incubated overnight at 30°C. Cells were then resuspended in liquid LB and transformants selected on LB agar plates containing both gentamicin (30 mg/l) and kanamycin (25 mg/l). The resulting CFP and YFP labelled strains were then directly competed with unlabelled strains to confirm that the impact of the labelling process on growth rate and motility was negligible (Extended Data Fig. 8a).

Bacterial cell culture

We streaked -80°C freezer stocks onto LB agar plates and incubated them overnight at 37°C. Single colonies were picked and grown overnight in liquid LB at 37°C under continuous shaking. The following day, overnight cultures were diluted 30-fold in fresh LB broth and returned to the 37°C shaking incubator for two hours to obtain cells in exponential phase. Immediately before being used in colony experiments, the optical density at 600 nm (OD_{600}) was adjusted to 0.05 (approximately 12,500 cells μl^{-1}) using fresh LB. For co-culture experiments, the optical densities of cultures of each individual strain were adjusted to OD_{600} = 0.05. These were then mixed in a single tube to ensure both strains were present in equal proportion. All colony-based assays were conducted at room temperature using LB agar composed of 5 g/l NaCl (Fisher Scientific), 5 g/l yeast extract (Bacto brand, BD) and 10 g/l tryptone (Bacto brand, BD), solidified with either 0.8% (w/v, subsurface assays unless otherwise specified) or 1.5% (w/v, surficial assays) agar.

Microscopic imaging

We used three microscopes in this study: two higher magnification inverted microscopes (denoted as "Zeiss inverted" and "Nikon inverted" within this text and our Supplementary Methods section) and a lower magnification stereo zoom microscope ("Zeiss zoom"). Specifications of these instruments are provided in Supplementary Table 1.

Surficial colony competition assays

Surficial colonies were inoculated with a liquid culture containing an equal fraction of two different strains, as described above. We spotted 10 µl of exponential phase culture onto a freshly poured 1.5% (w/v) LB agar plate and sealed the plate lids with Parafilm (Bemis) to prevent evaporation. The surficial colonies in which we evaluated the distribution and

orientation of cells using scanning laser confocal microscopy (Extended Data Fig. 2b, c) were incubated at room temperature for 16 h, ensuring that the colony was still sufficiently thin that its entire thickness could be imaged. In contrast, the surficial colonies used to quantify the relative fitness of WT/Δ*pilH* cells and their macroscopic distribution (Fig. 2a, b) were incubated at room temperature for 48 h.

We calculated the average number of cell divisions in each strain over the 48 hours of competition using the expression

$$\text{number of cell divisions} = \log_2\left(\frac{\text{number of cells of strain inoculated onto surface}}{\text{number of cells of strain in colony after 48 h}}\right). \quad (1)$$

We measured the number of cells of each strain used to inoculate the surface by diluting the mixed liquid cultures used for inoculation and then spreading them onto LB plates. The resulting colonies were counted manually after incubation overnight using the Zeiss zoom microscope, which allowed us to distinguish YFP and CFP expressing colony forming units (CFUs).

A similar technique was used to estimate the number of cells of each strain in each colony after 48 h of incubation. Whole surficial colonies were scraped, resuspended in fresh media, and vortexed. The resulting suspensions were then diluted, spread onto LB plates, and incubated overnight. We again used the Zeiss zoom microscope to manually enumerate the number of CFUs expressing either YFP or CFP.

We imaged the colonies after 48 h of competition to visualize the distribution of the two different strains. *P. aeruginosa* natively produces secretions called siderophores that have similar excitation and emission spectra to CFP[33,34]. While individual CFP and YFP labelled cells can easily be distinguished in the monolayer of subsurface colonies (Fig. 2g), surficial colonies are thicker, incubated for longer, and must be imaged with lower resolution objectives (owing to the lack of a coverslip). These three factors made it difficult to distinguish CFP-labelled cells from siderophore secretions in surficial colonies. To circumvent this problem, we imaged surficial colonies using a combination of brightfield and YFP fluorescence, such that regions with a larger fraction of CFP cells appeared darker grey in the merged brightfield/YFP images (Fig. 2a).

Subsurface assays

We prepared subsurface colonies using a protocol similar to one previously described[35]. Briefly, a pad of LB agar was cut from a freshly poured plate and transferred to a glass slide. We used an agar concentration of 0.8% (w/v), unless otherwise specified (Extended Data Fig. 10). The pad was spotted with a 1 μl drop of bacterial culture adjusted to an optical density of $OD_{600} = 0.05$, which was then allowed to dry until fully evaporated. The pad was then carefully inverted and placed into a coverslip-bottomed Petri dish (175 μm coverslip thickness, MatTek), sandwiching the cells between agar and the glass coverslip. By fully enclosing the agar pad, these dishes prevented evaporation and agar shrinkage over the course of the experiment. We found it was essential to use freshly poured agar to ensure consistency between experiments. We note that the bacterial culture was spotted onto the side of the agar that was facing up when

it was initially poured (i.e. the side that was exposed to air rather than the side against the plastic Petri dish).

Experiments with cells that lack pili (Δ*pilB*[17], Fig. 1f, g) and flagella (Δ*flgK*[36], Extended Data Fig. 1a, b) indicate that the cells in our assays move solely via pili-based motility. While cells can swim in agar at low concentrations[37], we found that the lowest concentration of agar used in our experiments (0.8%) supressed flagella-based motility. Δ*flgK* cells, which lack flagella, form subsurface colonies under 0.8% agar that actually expand at a faster rate than WT cells, verifying that the colony expansion observed in our assay is not driven by flagellar motility (Extended Data Fig. 1a, b). This finding is consistent with previous work that shows mutants lacking flagella exhibit more rapid twitching motility than WT cells[18], likely because flagella tend to adhere to surfaces.

A detailed description of each type of assay performed in the subsurface environment is provided in the Supplementary Methods.

Imaging rosette formation

Quantifying the movement of both defects and individual cells during the process of rosette formation (Fig. 4d-f, Extended Data Fig. 7, Supplementary Video 9) was exceptionally challenging, as it required imaging the monolayer at high spatial resolution (63X magnification, two frames per minute) at precisely the time and place that rosettes begin to form. It was difficult to estimate *a priori* where rosette formation would occur and thus where to place the Zeiss inverted microscope's relatively small field of view to capture these events.

To maximize our chances of success, we inoculated multiple Δ*pilH*-YFP/WT-CFP subsurface colonies with 10 µl of culture at a range of different optical densities (OD_{600}) in a 6-well coverslip-bottomed plate (175 µm coverslip thickness, MatTek). We then imaged the monolayer of each colony in turn, starting from the colony initiated with cells at the highest optical density. As rosettes form earlier in colonies inoculated at higher densities, this provided multiple opportunities to image the monolayer precisely at the moment that rosette formation begins.

Initially, we attempted to take time-lapse images of rosette formation using fluorescent confocal microscopy so we could use their YFP and CFP labels to continuously follow how the two strains were distributed. However, this bleached the cells and adversely affected their movement. Instead, we imaged the dynamics of rosette formation using brightfield microscopy for a period of one hour (Supplementary Video 9). After rosettes had formed, we then immediately switched over to fluorescent confocal imaging, which allowed us to quantify the distribution and orientation of the two different strains within the same rosette (Fig. 4f). Both brightfield and confocal microscopy were performed on the Zeiss inverted microscope.

Liquid culture competition assay

To compare the growth rate of mutants (Extended Data Fig. 8b), we grew the different strains in liquid culture and estimated their fitness relative to a WT reference strain by counting CFUs. We mixed a CFP labelled WT reference strain with YFP labelled Δ*pilH*, Δ*pilB*, and WT test strains in a 1:1 ratio. Liquid cultures were started at OD_{600} = 0.02 and placed in a shaking

incubator at 23°C, the same temperature used in the subsurface colony experiments. We counted the number of YFP and CFP expressing CFUs after t = 0, 210, and 420 mins of competition to calculate the relative fitness, w, of the YFP test strain compared to that of the CFP control:

$$w = \frac{\ln(C_Y(t)/C_Y(0))}{\ln(C_C(t)/C_C(0))}, \quad (2)$$

where $C_Y(0)$ and $C_C(0)$ are respectively the numbers of YFP and CFP cells measured at the beginning of the competition, and $C_Y(t)$ and $\mathrm{C_C(t)}$ are the numbers of YFP and CFP cells at time t.

Measurement of cell length in liquid cultures

To measure the lengths of cells in liquid culture (Extended Data Fig. 6a), we combined exponentially growing cultures of Δ*pilH*-YFP and WT-CFP at a 1:1 ratio. These were fixed with 3% paraformaldehyde and then diluted in phosphate buffered saline (PBS, Fisher Scientific) in 96-well plates with optical bottoms (Nunc brand, Thermo Scientific). We then centrifuged plates to ensure that cells were oriented flat against the optical bottoms of the wells and imaged them in brightfield, YFP and CFP channels at 63X magnification using the Zeiss inverted microscope. Cell lengths were then measured using our FAST software (see below).

Analysis of Experimental Data

Quantification of subsurface colonies

Our automated measurements of colony edge positions (Figs. 1f, g, 2d, Extended Data Figs. 1a, b, 9b), cell packing fractions (Fig. 2e, Extended Data Fig. 10a) and strain composition (Fig. 2f, Extended Data Figs. 9c, 10b) were performed using custom Matlab (Mathworks) scripts, which have been made publicly available as the colEDGE package (see Code Availability statement below). For a detailed description of the techniques used in this package, please see the Supplementary Methods and Supplementary Fig. 1.

Single-cell tracking

We attempted to use existing software packages to track movement of individual cells within the monolayer of our subsurface colony experiments. As a single monolayer image can contain more than 10,000 cells, we found existing software packages (most of which are designed for tracking sparse objects at low density) were prohibitively slow and/or were incapable of correctly segmenting individual cells when they are tightly packed together. Obtaining an accurate segmentation under these densely packed conditions is particularly challenging because the boundaries that separate neighbouring cells tend to be both thin and very faint.

To overcome these problems, we developed a new Matlab-based tracking platform named FAST[38] (Feature-Assisted Segmenter/Tracker). In brief, FAST uses a standard tracking by detection framework[39]. Firstly, individual cells are isolated from their neighbours using a sequence of segmentation routines. Next, we measure the "features" of each cell within each frame (including cell position, orientation, morphology, and fluorescence intensity). Finally, we use these features to follow individual cells between frames using an algorithm that automatically trains itself using unsupervised machine learning to optimize tracking based on the available feature information.

In the segmentation stage, we use brightfield images to identify individual cells in the monolayer. Our software uses a combination of automated ridge detection[40], topographical watershed[41] and intensity thresholding to generate black and white binary images of cells that are not connected to one another. Using this binary image as a mask, we then extract each cell's features from the original brightfield image.

Tracking is achieved via a two-stage algorithm. In the first stage, a low-fidelity nearest-neighbour tracking algorithm is used to generate a set of putative links between objects in consecutive frames. The subset of links with the smallest corresponding frame-frame object displacements is then classified as "correct", typically forming around half of the total putative links. This subset forms the training dataset. Statistical parameters are then extracted from this training dataset, allowing us to quantify how robust each feature is as a marker of object identity and thus quantify its utility in linking objects between sequential frames. In the second stage of the tracking algorithm, these measurements are used to dynamically adjust the weighting of each feature such that unreliable features have a reduced weighting compared to more reliable features. Tracking is then repeated, using these reweighted features as inputs. This approach allowed us to obtain extremely large tracking datasets, for example yielding a total of 161,769 cell trajectories for a single movie of the WT monolayer.

We note that, while FAST has primarily been developed for tracking of single cells in dense monolayers, its capabilities can also be leveraged to analyse other datasets. For example, FAST can also be used to track topological defects through time and space (see following section).

Detection and analysis of topological defects

Comets and trefoils occur at singularities in cell orientation. Our automated approach for locating these singularities is similar to that described in[42]: in the first stage, we used the OrientationJ plugin for Fiji[43] to measure the local orientation of cells using the tensor method[44]. Experimental images of the monolayer were loaded into Fiji and directly analysed using OrientationJ. For consistency, we also used the same defect analysis pipeline to analyse the output of our 2D SPR model: a timeseries of images was created by drawing rods as grey ellipses on a white background. These images were then processed using OrientationJ in the same way as for our experimental images. To facilitate direct comparison between our experiments and simulations, we set the size of the structure tensor window, which defines the spatial scale over which the orientation field is calculated, to a length equivalent to two cell/rod widths. This process yields the orientation of cells, $\theta = [-\pi/2, \pi/2)$ at each pixel in the input image (Fig. 3b).

The position, orientation and topological charge of defects was quantified using our custom package known as Defector (see Code Availability statement below and Supplementary Methods). We next used FAST to track the movement of defects using these three quantities as "features" (see above). To reduce noise, defects that were present for fewer than five timepoints were excluded from our analyses. This analysis yielded a total set of 1344 trefoil trajectories and 1382 comet trajectories.

The root mean square displacement (RMSD) of tracked trefoils and comets was calculated using:

$$R(\tau) = \sqrt{\langle \big(x(t+\tau) - x(t)\big)^2 + \big(y(t+\tau) - y(t)\big)^2 \rangle}. \quad (5)$$

where τ is the lag time, (x, y) is the position of the defect, and $\langle \cdot \rangle$ denotes an ensemble average across all times t and across all defect trajectories.

True to its name, pili-based "twitching" motility is jerky and highly unsteady, owing to the stochastic retraction and detachment of individual pili[6,45]. Obtaining a reliable measure of cell movement around defects thus required averaging of data across a large number of defects and cell trajectories so that the stochastic component of each cell's movement was averaged out.

To accomplish this, we first tracked the movement of defects and cells independently from one another. Next, we transformed the coordinate system of each cell trajectory so that its origin and orientation were measured relative to the centre and orientation of a nearby defect. This allowed us to combine cell trajectories collected from around a large number of different comet and trefoil defects. After the cell trajectories were aligned with one another in the same reference frame, we averaged the cell velocities in a two-dimensional array of bins. The size

of each bin was 3.2 μm × 3.2 μm. All velocity measurements were made with respect to the laboratory reference frame, not the reference frame of the defect.

We calculated the flowfield around defects in the SPR model in the same way as for the experiments, though rod trajectories were obtained directly from the model output, rather than from our FAST tracking software. To facilitate direct comparison between the SPR model and experiments, we normalized the flowfield velocities around defects by dividing them by the average speed of all cells within the simulation or field of view, respectively.

Quantification of collective cell motility during rosette development

While our FAST software can track horizontally oriented cells in the monolayer of colonies, once cells had reoriented perpendicular to the surface it was exceedingly difficult to distinguish individual cells. Instead of single-cell tracking, we therefore characterized the collective movement of cells during rosette formation using a technique known as particle image velocimetry (PIV), which operates on a coarse-grained level and does not require the segmentation of individual cells.

PIV analysis was conducted on a one-hour long timeseries of 63X magnification images recorded during the process of rosette formation. Images were pre-processed using contrast normalization and manually stabilized to remove thermally-induced drift in the *xy*-plane. The resulting images were analysed using PIVlab, an open-source Matlab-based software[46]. We filtered our results using PIVlab's built-in tools to remove spurious data points. Specifically, velocity vectors that exceeded 0.6 μm min^{-1} were removed and replaced with velocities interpolated from neighbouring points, which helped to reduce noise. The resulting measurements of instantaneous velocity were then averaged over the entire one-hour period (Fig. 4e, Extended Data Fig. 7a).

Self-propelled rod (SPR) simulations

Our individual-based SPR simulations are based on a previously described model[15,47]. In brief, cells are modelled as rigid rods, which repel each other through an interaction potential, U. Additionally, rods are self-propelled by a force, F, acting along their axis. In the 2D SPR model used to generate Fig. 3d, the equations of motion that determine how the position $\boldsymbol{r}_\alpha$ and orientation θ_α of rod α changes over time are given by:

$$\boldsymbol{f}_T \cdot \frac{\partial \boldsymbol{r}_\alpha}{\partial t} = -\frac{\partial U_\alpha}{\partial \boldsymbol{r}_\alpha} + F\hat{\boldsymbol{u}}_\alpha, \tag{6a}$$

$$f_\theta \frac{\partial \theta_\alpha}{\partial t} = -\frac{\partial U_\alpha}{\partial \theta_\alpha}, \tag{6b}$$

where $\boldsymbol{f}_T$ is the translational friction tensor, f_θ is the rotational friction constant and $\hat{\boldsymbol{u}}_\alpha$ is the orientational unit vector ($\hat{\boldsymbol{u}}_\alpha = (\cos(\theta_\alpha), \sin(\theta_\alpha))$.

During rosette formation, cells can reorient out of the plane. To model this process, we give the rods in our simulations an additional degree of freedom, by allowing them to change their polar angle ϕ with respect to the 2D plane. The equation of motion for ϕ is similar to that for the azimuthal angle θ, but must additionally take into account the elastic restoring force imposed by the overlaying agar and extracellular polymeric secretions (EPS) that acts to keep cells oriented flat against the surface[48,49]. Taking these into account, we arrive at the third equation of motion:

$$f_\phi \frac{\partial \phi_\alpha}{\partial t} = -\frac{\partial U_\alpha}{\partial \phi_\alpha} + \frac{k l_\alpha^2}{2} \cos\phi_\alpha \sin\phi_\alpha\,, \tag{7}$$

where k is the stiffness of the overlying substrate and l_α is the length of the rod. This additional governing equation is used in the simulations presented in Fig. 4a-c and Extended Data Figs. 5, 6b-e. Further details about these simulations and a more detailed derivation of Eq. (7) can both be found in the Supplementary Methods.

Data availability

Source data for figures 1-4 and extended data figures 1, 3-6, 8-10 are provided with the paper. Additionally, data that support the findings of this study can be accessed at https://doi.org/10.15131/shef.data.12735251.v1.

Code availability

The FAST cell tracking package can be accessed at https://doi.org/10.5281/zenodo.3630641, with extensive documentation on its use and functionality available at https://mackdurham.group.shef.ac.uk/FAST_DokuWiki/dokuwiki. The Defector defect detection package is available at https://doi.org/10.5281/zenodo.3974873, while the colEDGE colony composition package can be accessed at https://doi.org/10.5281/zenodo.3974875.

31. O'Toole, G. A. & Kolter, R. Flagellar and twitching motility are necessary for Pseudomonas aeruginosa biofilm development. *Mol. Microbiol.* **30**, 295–304 (1998).

32. Choi, K.-H. & Schweizer, H. P. mini-Tn7 insertion in bacteria with single attTn7 sites: example Pseudomonas aeruginosa. *Nat. Protoc.* **1**, 153–161 (2006).

33. Cox, C. D. & Graham, R. Isolation of an iron-binding compound from Pseudomonas aeruginosa. *J. Bacteriol.* **137**, 357–64 (1979).

34. Elliott, R. P. Some properties of pyoverdine, the water-soluble fluorescent pigment of the pseudomonads. *Appl. Microbiol.* **6**, 241–6 (1958).

35. Darzins, A. The pilG gene product, required for Pseudomonas aeruginosa pilus production and twitching motility, is homologous to the enteric, single-domain response regulator CheY. *J. Bacteriol.* **175**, 5934–44 (1993).

36. O'Toole, G. A. & Kolter, R. Initiation of biofilm formation in Pseudomonas fluorescens WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. *Mol. Microbiol.* **28**, 449–461 (1998).

37. Wolfe, A. J. & Berg, H. C. Migration of bacteria in semisolid agar. *PNAS* **86**, 6973 (1989).

38. Meacock, O. J. & Durham, W. M. FAST v0.9.1. (2020). doi:10.5281/ZENODO.3630642

39. Li, K. *et al.* Cell population tracking and lineage construction with spatiotemporal context. *Med. Image Anal.* **12**, 546–566 (2008).

40. Lindeberg, T. Edge Detection and Ridge Detection with Automatic Scale Selection. *Int. J. Comput. Vis.* **30**, 117–156 (1998).

41. Meyer, F. Topographic distance and watershed lines. *Signal Processing* **38**, 113–125 (1994).

42. Kawaguchi, K., Kageyama, R. & Sano, M. Topological defects control collective dynamics in neural progenitor cell cultures. *Nature* **545**, 327–331 (2017).

43. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

44. Püspöki, Z., Storath, M., Sage, D. & Unser, M. Transforms and Operators for Directional Bioimage Analysis: A Survey. In *Focus on Bio-Image Informatics* (eds. Vos, W. H. De, Munck, S. & Timmermans, J.-P.) 69–93 (Springer, Cham, 2016).

45. Jin, F., Conrad, J. C., Gibiansky, M. L. & Wong, G. C. L. Bacteria use type-IV pili to slingshot on surfaces. *PNAS* **108**, 12617–22 (2011).

46. Thielicke, W. & Stamhuis, E. J. PIVlab – Towards User-friendly, Affordable and Accurate Digital Particle Image Velocimetry in MATLAB. *J. Open Res. Softw.* **2**, e30 (2014).

47. Wensink, H. H. & Löwen, H. Emergent states in dense systems of active rods: from swarming to turbulence. *J. Phys. Condens. Matter* **24**, 464130 (2012).

48. Zhao, K. *et al.* Psl trails guide exploration and microcolony formation in Pseudomonas aeruginosa biofilms. *Nature* **497**, 388–91 (2013).

49. Nayar, V. T., Weiland, J. D., Nelson, C. S. & Hodge, A. M. Elastic and viscoelastic characterization of agar. *J. Mech. Behav. Biomed. Mater.* **7**, 60–68 (2012).

50. Shi, X. & Ma, Y. Topological structure dynamics revealing collective evolution in active nematics. *Nat. Commun.* **4**, 3013 (2013).