
Supplementary information

Clonal-aggregative multicellularity tuned by salinity in a choanoflagellate

In the format provided by the
authors and unedited

Supplementary Information for

Clonal-aggregative multicellularity tuned by salinity in a choanoflagellate

Núria Ros-Rocher^{1*}, Josean Reyes-Rivera^{2,3*}, Uzuki Horo¹, Chantal Combredet¹,
Yeganeh Foroughijabbari¹, Ben T. Larson^{2,4}, Maxwell C. Coyle^{2,5}, Erik A.T. Houtepen⁶,
Mark J. A. Vermeij^{6,7}, Jacob L. Steenwyk², Thibaut Brunet¹

¹Institut Pasteur, Université Paris-Cité, CNRS UMR3691, Evolutionary Cell Biology and Evolution of Morphogenesis Unit, 25-28 Rue du Dr. Roux, 75015 Paris, France

²Howard Hughes Medical Institute and the Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA, USA

³Present address: Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA, USA

⁴Present address: Department of Biological Sciences and Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, 110 8th St, Troy, NY, 12180, USA

⁵Present address: Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA

⁶Caribbean Research and Management of Biodiversity foundation (CARMABI), Piscaderabaai z/n Willemstad, Curaçao

⁷Department of Freshwater and Marine Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, 1012 WP, The Netherlands

*Contributed equally

Correspondence: N. R.-R.: nuria.ros@pasteur.fr J. R.-R.: josean.reyes@berkeley.edu T.B.: thibaut.brunet@pasteur.fr

This PDF file includes (in order of appearance):

Supplementary Methods

Supplementary Figures S1 to S11

Supplementary Table S2

SI References

Supplementary Methods

All dilution percentages are (v/v) unless specified otherwise.

Cell strains and growth conditions

Choanoeca flexa clonal monoxenic cultures (strain 'ChoPs7', hereafter referred to as 'Strain 1'), established as described in¹ from single sheet cultures isolated near Boka Wandomi in Shete Boka National Park (see **Figure 6a** for location reference), were grown in 25 cm² tissue-culture treated flasks (#130189, ThermoFisher Scientific) in either in 1% Seawater Complete (SWC) or 5% Cereal Grass Medium 3 (CGM3), following the protocols below:

Culture in 1% SWC: ChoPs7 cells were cultured in 10 mL of filter-sterilised 1X Artificial Seawater (see recipe below) supplemented with 1% SWC (see recipe below), hereafter referred to as *SWC medium*, unless otherwise specified. Cultures were further supplemented with 1% *Halopseudomonas oceani* (formerly *Pseudomonas oceani*) resuspended food pellet (see below) and maintained at 25°C or 30°C (see specific experiments below) in an incubator with a 12-hour light-dark cycle (Memmert IPP410ecoplus). Incubator humidity was maintained with an open box of distilled water placed on the bottom shelf of the incubator. For most experiments, ChoPs7 cultures were scaled up in 75 cm² tissue-culture treated flasks (#130190, ThermoFisher Scientific) containing 20 mL of filter-sterilised 1X Artificial Seawater, 1% SWC and 1% *H. oceani* resuspended food pellet. For light-to-dark inversion experiments, ChoPs7 cultures grown in SWC medium were additionally supplemented with 0.1% *Algoriphagus machipongonensis* resuspended food pellet (see below), added as a source of retinal, which is required for light responsiveness in *C. flexa*¹.

Culture in 5% CGM3: ChoPs7 cells were cultured in 25 cm² tissue-culture treated flasks containing 10 mL of filter-sterilised 1X Artificial Seawater supplemented with 5% CGM3 (see below), hereafter referred to as *CGM3 medium*, unless otherwise specified. Cultures were further supplemented with 1% *H. oceani* resuspended food pellet (see below) and maintained at 25°C under a 12-hour light-dark cycle in a Caron low temperature incubator equipped with a Venoya Full Spectrum 150W Plant Growth LED lamp controlled by a programmable timer (Leviton VPT24-1PZ Vizia). Incubator humidity was maintained at approximately 60% as above.

Salpingoeca rosetta monoxenic clonal cultures (strain 'SrEpac'; American Type Culture Collection ATCC PRA-390), established as described in (Levin and King, 2013), were grown in 25 cm² tissue-culture treated flasks (#130189, ThermoFisher Scientific) containing 10 mL of filter-sterilised SWC medium. SrEpac cultures were supplemented with 1% *Echinicola pacifica* resuspended food pellet (see below) and maintained at 22°C in an incubator (HPP IPP^{Plus}, Memmert). Incubator humidity was maintained using an open box of distilled water placed on the bottom shelf of the incubator. For kin recognition experiments, SrEpac cultures were scaled up in 75 cm² tissue-culture treated flasks (#130190, ThermoFisher Scientific) with 20 mL of filter-sterilised SWC medium supplemented with 1% *E. pacifica* resuspended food pellet and grown for 2 days to obtain a mixed culture of slow and fast swimmers.

Culture medium preparation

1X Artificial Seawater (hereafter ASW) was prepared by diluting 32.9 g of Instant Ocean (#218035, Aquarium Systems) in 1 L of ultrapure Milli-Q water and adjusting the pH to 8.0 ± 0.1. The 1X ASW solution was filter-sterilised using a 0.22 µm pore-size filter (#SEGPU1145, Millipore). The resulting salinity of the pH-adjusted and sterile-filtered 1X ASW stock solution was 35 ppt. Sterility was tested by incubating a sealed 5 mL aliquot at 37°C for 24 hours and visually assessing transparency.

100% stock of SWC medium was prepared by diluting 12 g of Instant Ocean, 1.5 mL glycerol (#15523, Sigma-Aldrich), 2.5 g Bacto Peptone (#211677, BD Difco) and 1.5 g yeast extract (#212750, BD Difco) in 1 L of ultrapure Milli-Q water. The 100% SWC solution was filter-sterilised using a 0.22 µm pore-size filter. A 5 mL aliquot was incubated at 37°C for 24 hours to confirm sterility.

100% CGM3 was prepared by diluting 5 g of Cereal Grass (#IS5020, Basic Science Supplies) in 1 L of 1X ASW directly after autoclaving. The medium was filtered using a 150 mm filter paper (#1001-150, Whatman) placed in a Büchner funnel (#60246, CoorsTek) and vacuum-filtered to remove large particles. This step was repeated once. Aliquots were then filter-sterilised using a 0.22 µm filter connected to a vacuum line and stored at room temperature.

Bacteria food pellet preparation

H. oceani, *E. pacifica*, and *A. machipongonensis* food pellets were prepared by inoculating either 5 mL of an overnight pre-grown culture in 100% SWC (from a glycerol stock) or a single bacterial

colony (from an SWC-agar plate) into 500 mL of 100% SWC. The bacterial culture was then incubated at room temperature for ~2-3 days in a rocking shaker until the optical density (OD₆₀₀) reached ~0.6. Next, the 500 mL bacterial culture was distributed into 50 mL falcon tubes and harvested at 4,000-5,000xg for 30 minutes at 4°C in a tabletop centrifuge. After centrifugation, the supernatant was removed by pouring off the liquid and then using a serological pipette to remove any remaining liquid. The bacterial pellet from each tube was weighed to calculate the required resuspension volume of 1X ASW to reach a final bacterial concentration of 20 mg/mL, and the pellet was resuspended by pipetting up and down. After resuspension, 500 µL aliquots were distributed into 1.5 mL microcentrifuge tubes and centrifuged at 16,000xg for 10 minutes at 4°C. The supernatant was carefully removed with a fine transfer pipette, and the resulting food pellets were flash-frozen in liquid nitrogen and stored at -80°C. Ready-to-use food pellets were prepared by resuspending a 10 mg/mL food pellet in 1 mL of 1X ASW.

Imaging *C. flexa* flagella-in and flagella-out sheets (Figure 1b-d)

To image flagella-in sheets (**Figure 1b**), approximately 1mL of dense ChoPs7 cultures grown in CGM3 medium was transferred into a FluoroDish (#FD35-100, World Precision Instruments) and incubated for 30 minutes to allow colonies to settle at the bottom of the dish. Colonies were imaged by differential interference contrast (DIC) microscopy using a C-Apochromat 63X/1.4 Zeiss objective with a 1.6X optovar, mounted on a Zeiss Observer Z.1 inverted microscope equipped with a Hamamatsu ORCA-Flash 4.0 V2 CMOS camera (C1140-22CU). All experiments were performed in three independent biological replicates (n=56 sheets imaged). Flagella-out sheets (**Figure 1d**) were imaged as described in¹ in seven independent biological replicates, each including two technical replicates (n=30 sheets imaged).

Tracking cell division of single cells (Figure 1e-f; Video S1)

The cell concentration of an exponentially growing ChoPs7 culture grown in SWC medium was estimated using a LUNA-II automated cell counter (LogosBiosystems). The culture was then diluted to 1 cell/µL in 5% SWC supplemented with 1% *H. ocean*i resuspended food pellet. A 1 µL aliquot of the diluted culture was pipetted onto the centre of a well in a black 96-well ibiTreat µ-plate (#89626, Ibidi) and covered with 400 µL of Ibidi anti-evaporation oil (#50051, Ibidi). The sample was imaged using a Plan-Apochromat 20X/0.8 M27 Zeiss objective on a Zeiss Axio Observer Z.1 inverted microscope, employing tile scan mode to cover the entire droplet surface, Definite Focus, and a ColorBand filter (#FGL610, Thorlabs). All experiments were performed in three independent biological replicates (n=3).

Timelapse imaging of mixed clonal-aggregative multicellularity (Figures 1g and S1; Videos S2-S3)

Colonies from ChoPs7 cultures grown in CGM3 medium were transferred to a FluoroDish (#FD35-100, World Precision Instruments) and incubated for 30 minutes to allow them to settle at the bottom of the dish. Colonies were imaged every 5 minutes by DIC microscopy using a Plan-Apochromat 63X/1.4 oil-immersion Zeiss objective mounted on a Zeiss Observer Z.1 inverted microscope equipped with a Hamamatsu ORCA-Flash 4.0 V2 CMOS camera (C1140-22CU) in 'timelapse' mode. All experiments were performed in four independent replicates (two biological replicates, each including two technical replicates; n=4).

Aggregation dynamics over time (Figure 2b-d; Videos S4-S6)

Two types of timelapse movies of aggregation were generated: low-magnification overviews capturing a large number of cells imaged with a 5X objective (Figure 2b; Video S4) and high-magnification, high-frequency movies of a smaller number of cells imaged with a 63X objective to enable cell tracking (Figure 2c-d; Videos S5-S6). The experiment was performed in four independent biological replicates (n=4).

Low-magnification movies (Figure 2b; Video S4): 3 mL of dense ChoPs7 cultures grown in SWC medium were collected by centrifugation at 4,700xg for 20 minutes at 4°C, resuspended in 1 mL of supernatant, and dissociated by vortexing for 30 seconds. 200 µL of the resulting cell suspension were transferred into a well of a Corning 96-well plate (#13539050, Fisher scientific) and imaged every 5 seconds overnight by DIC microscopy using an EC Plan-Neofluar 5X/0.16 M27 Zeiss objective mounted on a Zeiss Axio Observer Z.1 inverted microscope. Settings: Define Focus and 'Subtract background' (background defined as an image acquired with the same settings but without a sample in the microscope). The resulting image stack was processed in Fiji Imaging Software version v2.16.0 (ImageJ2²) to remove background noise. Specifically, a Gaussian blur (50-pixel radius) was applied to each slice of the stack and inverted, generating an 'inverted background' stack. The original stack was also inverted, the inverted background was subtracted from it, and the resulting image was inverted again.

High-magnification movies (Figure 2c-d; Videos S5-S6): 10 mL of a mid-log phase ChoPs7 culture ($\sim 1 \times 10^4$ cells/mL) were collected by centrifugation at 4,700xg for 20 minutes at 4°C, resuspended in 200 µL of supernatant, transferred to a 1.5 mL microcentrifuge tube, centrifuged

again at 11,000xg for 15 minutes at 4°C, resuspended in 20 µL of supernatant, and dissociated by vortexing for 30 seconds. 0.2 µL of the dissociated cell suspension were pipetted onto the centre of a well in a Corning 96-well plate (#13539050, Fisher scientific) and covered with 200 µL of Ibidi anti-evaporation oil (#50051, CliniSciences). The sample was imaged every second overnight using DIC microscopy on a Zeiss Axio Observer Z.1 inverted microscope equipped with a C-Apochromat 63x/1.20 W water immersion Zeiss objective, using Define Focus and 'Subtract background' settings (background defined as above). The resulting image stacks were processed in Fiji Imaging Software version v2.16.0 (ImageJ2), and cells were tracked manually using the Manual Tracking plugin.

Fluorescence staining of aggregation dynamics over time (Figures 2e-g and S2; Videos S7-S11)

Approximately 10-20 mL of exponentially growing ChoPs7 cultures grown in SWC medium were harvested by centrifugation at 3,300xg during 15 minutes at room temperature and washed twice with 20 mL of 1X ASW. Cells were resuspended in ~1 mL of 1X ASW by pipetting up and down, dissociated by vortexing for 20 seconds, and counted using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-15 µm using the 'Histogram' option. Next, a total of 2.2×10^5 cells were seeded in 1 mL of 1X ASW in a 24-well plate and incubated for 10 minutes, 30 minutes, 2 hours, 6 hours, and 24 hours at 25°C. After incubation, 45 µL of cells were transferred using a truncated P200 tip into a black 96-well ibiTreat µ-plate (#89626, Ibidi) pre-treated with 200 µL of 100 µg/mL Poly-D-Lysine hydrobromide (#P6407, Sigma-Aldrich) (each well washed twice with 200 µL of Milli-Q water prior to cell seeding). Cells were allowed to settle for 5-10 minutes at room temperature. Cells were then fixed by adding 16% ice-cold paraformaldehyde (PFA) (#15710, Electron Microscopy Sciences) to a final concentration of 4% and incubated for 5 minutes at room temperature. Next, cells were permeabilised for 5 minutes at room temperature using 0.1% Triton X-100 (#A16046, ThermoFisher Scientific) prepared from a 10% stock in distilled water. Finally, cells were stained with a 10X dye mix in 1X ASW containing 1:100 FM 4-64X (stock at 5 mg/mL, pre-diluted in Milli-Q water; #F34653, Invitrogen) and 1:100 Alexa Fluor 488 Phalloidin (stock at 2 units/µL, pre-diluted in DMSO; #A12379, Invitrogen), yielding final concentrations of 1:1000 for each dye.

Samples were imaged using Plan-Apochromat 40X/1.3 Oil DIC (UV) VIS-IR M27 and Plan-Apochromat 63X/1.4 Oil M27 Zeiss objectives with Airyscan MPLX SR-4Y and 2X line averaging modes on a Zeiss Axio Observer Z1/7 inverted microscope equipped with an LSM900 Airyscan

2. All experiments were performed in four independent replicates (two biological replicates, each including two technical replicates; n=4). A total of n=75 sheets were imaged, distributed by timepoint as follows: 10 minutes (n=11), 30 minutes (n=14), 2 hours (n=17), 6 hours (n=20), and 24 hours (n=13).

Image processing and quantification: Images were processed using the 'Airyscan process' function in the 'Image Processing' menu of Zen Blue software and further analysed with Imaris Imaging software version 9.9.1 (build 61122 for x64). The 'Volume' tool was used to generate 3D renderings, captured using the 'Snapshot' and movie recorder tools. The number of cells with misaligned apicobasal orientation was counted manually. The 'Surfaces' tool was applied for image segmentation and automatic quantification of the number of cells per sheet, with manual corrections using the 'Cut Surface' tool when necessary. Two orthogonal cross-sections per sheet (each containing ≥ 4 cells) were visualised using the 'Oblique Slicer' tool. Angles between collar contacts of neighbouring cells were measured manually using the 'Surface of Object' and 'Intersect' options within the 'Measurement Points' tool by drawing a triangle connecting the actin signal at the base of both collars and the tips of the collar-collar contacts (see **Figures 2f** and **S2a**). Cells with the same apicobasal orientation had collar-collar angles $<120^\circ$.

In parallel, images were batch processed in Fiji Imaging Software version 2.14.0/1.54f using a custom macro script. Briefly, brightness and contrast were automatically adjusted, and a maximum-intensity z-projection was applied to all images. Image contrast was then enhanced five times using the 'Enhance contrast' command from the 'Process' menu, setting a saturation value of 0.35. The processed images were stacked into a single RGB image and converted to 8-bit. Thresholds were manually adjusted using the 'Adjust Threshold' command in the 'Image' menu. Images were then binarized using the 'Convert to Mask' and 'Fill holes' commands in the 'Process' menu. Finally, the 'Analyse Particles' tool was used to quantify the area and circularity of each sheet, with analysis parameters set to a particle size range of 20 to infinity μm^2 and a circularity range of 0 to 1.

Live imaging of dual labelling of aggregates (Figure S3a; Video S12)

Approximately 40 mL of exponentially growing ChoPs7 cultures grown in SWC medium were harvested by centrifugation at 3,300xg for 15 minutes at room temperature and washed twice with 40 mL of 1X ASW. Cells were resuspended in ~1 mL of 1X ASW by pipetting up and down, dissociated by vortexing for 20 seconds, and divided into two 1.5 mL microcentrifuge tubes. Cells

in each tube were stained 1:1000 with either CellTrace CFSE (green; #C34570, ThermoFisher Scientific) or CellTrace Far Red (magenta; #C34572, ThermoFisher Scientific), prepared from 5 mM and 1 mM stock solutions in DMSO, respectively, and incubated for 20 minutes at room temperature on a rocking shaker. After incubation, 1% Bovine Serum Albumin (BSA) (#BP9705-100, Fisher Scientific; prepared from a 50 mg/mL stock solution in distilled water) was added to quench excess unconjugated dye for 1 minute at room temperature by gently inverting each tube. Next, cells were harvested again at 3,300xg for 15 minutes at room temperature, washed once with 1 mL of 1% BSA in 1X ASW, followed by one wash with 1 mL of 1X ASW, and finally resuspended in 500 μ L of 1X ASW. Cell concentration was estimated using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-15 μ m using the 'Histogram' option. Before seeding, samples were vortexed for 30 seconds to ensure dissociation into single cells. The green- and magenta-labelled single-cell populations were mixed at a 1:1 ratio, and 2×10^5 total cells were seeded in 300 μ L of 1X ASW in a black 96-well ibiTreat μ -plate (#89626, Ibidi).

Cells were imaged every 3 minutes at five positions using DIC and 4% intensity 488 nm and 650 nm LEDs, with a Plan-Apochromat 20X/0.8 M27 Zeiss objective on a Zeiss Axio Observer Z.1 inverted microscope, using Definite Focus and a ColorBand filter (#FGL610, Thorlabs). All experiments were performed in three independent biological replicates (n=3).

Airyscan imaging of dual labelling of aggregates (Figures 2h and S3b-c; Video S13)

Green- and magenta-labelled single-cell populations were stained as described above, mixed at a 1:1 ratio, and seeded in 1 mL of SWC medium in a 24-well plate (#3526, Corning). Plates were incubated overnight at 25°C, protected from light with a layer of aluminium foil. Non-mixed single-labelled populations (2×10^5 single cells) were seeded in parallel as controls. After incubation, 100 μ L of colonies were collected from the centre of the well using a truncated P200 filter tip and transferred to a black 96-well ibiTreat μ -plate (#89626, Ibidi) coated with 200 μ L of 100 μ g/mL Poly-D-Lysine hydrobromide (#P6407, Sigma-Aldrich; stock solution in distilled water). Each well was washed twice with 200 μ L of distilled water prior to cell seeding. Cells were incubated for 15 minutes at room temperature, then fixed with 16% ice-cold PFA (#15710, Electron Microscopy Sciences) to a final concentration of 4% PFA and incubated for an additional 5 minutes at room temperature. Next, cells were permeabilised for 5 minutes at room temperature with 0.1% Triton X-100 (#A16046, ThermoFisher Scientific; prepared from a 10% stock solution in distilled water).

Finally, cells were stained with 1:1000 Alexa Fluor Plus 405 Phalloidin (stock at 2 units/ μ L, pre-diluted in DMSO; #A30104, Invitrogen).

Samples were imaged using a Plan-Apochromat 40X/1.3 Oil DIC (UV) VIS-IR M27 Zeiss objective in Airyscan MPLX SR-4Y and 2X line averaging modes on a Zeiss Axio Observer Z1/7 inverted microscope equipped with an LSM900 Airyscan 2 detector. All experiments were performed in four independent replicates (two biological replicates, each with two technical replicates; n=4). A total of n=16 sheets were imaged, distributed as follows: green-only control (n=5), magenta-only control (n=5), mixed (n=6).

Image processing and quantification: Images were processed using the 'Airyscan process' function in the 'Image Processing' menu of Zen Blue software and analysed with Imaris Imaging software version 9.9.1 (build 61122 for x64). The 'Volume' tool was used to generate 3D renderings, captured using the 'Snapshot' and movie recorder tools.

Growth curve of Aphidicolin-treated single cells and colonies (Figures 2i and S4a)

Chops7 cultures grown in CGM3 medium were treated overnight with the cell proliferation inhibitor aphidicolin (#38966-21-1, Santa Cruz Biotechnology) at 17 μ g/mL, or with dimethyl sulfoxide (DMSO; drug vehicle control), following a previously established protocol for the choanoflagellate *S. rosetta*³. To confirm the drug's efficacy in *C. flexa*, growth curves were performed using three concentrations of aphidicolin (8, 10, and 17 μ g/mL). Cells used for the growth curve were treated with aphidicolin or DMSO and incubated overnight. The following morning treated cells were seeded in 24-well plates at a density of 1×10^3 cells/mL in 2.5% CGM3 medium supplemented with 0.5% *H. oceanii* resuspended food pellet, in triplicate wells. Cells were fixed with 16% ice-cold PFA (#15710, Electron Microscopy Sciences) to a final concentration of 4% and counted every 24 hours using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 3-10 μ m using the 'Histogram' option, over a total period of 96 hours. All experiments were performed in three independent biological replicates, each including three technical replicates per condition (n=9).

Quantification of aggregation in aphidicolin-treated cells (Figures 2i and S4b-c)

Chops7 cultures grown in CGM3 medium were treated with aphidicolin or DMSO (drug vehicle control) as described above and incubated overnight. The following morning, colonies were transferred to 50 mL Falcon tubes and vortexed at the 'fast' setting on a Vortex Genie 2 for 1

minute to dissociate them into single cells. Single cells were concentrated to 3×10^5 cells/mL for each replicate, and 200 μ L of the cell suspension were transferred into an μ -Slide 8-Well chamber (#80826, Ibidi).

Cells were imaged at 6-9 different locations per replicate every 30 minutes for 2 hours using a C-Apochromat 10X/0.45 Zeiss objective with an optovar 1.6X mounted on a Zeiss Observer Z.1 inverted microscope equipped with a Hamamatsu ORCA-Flash 4.0 V2 CMOS camera (C1140-22CU) and the 'Tile Positions' option. All experiments were performed in three independent biological replicates, each including three technical replicates per condition (n=9). A minimum of 1,474 particles were imaged per condition.

Image processing and quantification: Aggregate area was quantified using ImageJ Imaging Software version 2.14. Tiled images were processed sequentially using the following commands (default settings unless specified): 'Smooth' and 'Find Edges' (from the 'Process' menu); 'Despeckle' (from the 'Process' menu), 'Make Binary', 'Dilate', 'Erode', and 'Fill Holes' (from the 'Binary' submenu of the 'Process' menu); and finally 'Analyze Particles' (from the 'Analyze' menu) to quantify particle areas.

Aggregation induction in fixed versus live cells (Figure 2j-k)

Approximately 40 mL of exponentially growing ChoPs7 cultures grown in SWC medium were harvested by centrifugation at 3,300xg for 15 minutes at room temperature and resuspended in 1 mL of 1X ASW by pipetting up and down. Cells were dissociated by vortexing for 20 seconds and counted using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-11 μ m using the 'Histogram' option. Before seeding, samples were vortexed again for 30 seconds to ensure complete dissociation into single cells. A total of 5×10^4 cells were seeded in 1 mL of SWC medium in a 24-well plate and immediately fixed with 16% ice-cold PFA (#15710, Electron Microscopy Sciences) to a final concentration of 4%. Cells were then placed on an orbital shaker (Rotamax 120, Heidolph) at 50 rpm for 24 hours at room temperature. Control plates containing live (non-fixed) cells and static (non-agitated) conditions were prepared in parallel.

After 24 hours of incubation, cells were imaged at five distinct locations per well using transmitted light (TL) brightfield microscopy with an EC Plan-Neofluar 5X/0.16 M27 Zeiss objective on a Zeiss Axio Observer Z.1 inverted microscope. The experiment was performed in three independent

biological replicates, each including two technical replicates per condition (n=6). A minimum of 2,566 particles were quantified in each condition.

Image processing and quantification: Images were batch processed in Fiji Imaging Software version 2.14.0/1.54f using a custom macro script. Briefly, images were smoothed three times using the 'Smooth' command in the 'Process' menu. A 'Subtract Background' step was applied to reduce image shadows, using a rolling ball radius of 50 pixels. Next, images were binarized using the 'Convert to Mask' command in the 'Process' menu and analysed using the 'Analyse Particles' command to calculate area and circularity for each particle. Analysis parameters were set to a particle size range of 25 to infinity μm^2 and a circularity of 0.01 to 1, excluding particles at image edges.

Curaçao sampling procedure (Figures 3 and 6; Extended Data Figures 1-4; Table S1; Videos S14-S16)

Fieldwork data were collected in Shete Boka National Park, a protected area located in Willemstad on the northwestern coast of Curaçao (12°22'5.718"N, 69°06'56.916"W), in two independent expeditions during July-August 2023 (*Exped-A* and *Exped-B*; n=150 splash pools) (Figure 3, excluding Figure 3k; Extended Data Figures 1-4) and July-August 2024 (*Exped-C*; n=12 splash pools) (Figures 3k; Extended Data Figure 7a). The park spans nearly 10 km of rocky, wave-exposed coastline and contains approximately 10 pocket bays, or '*bokas*'.

Exped-A sampling and monitoring: For *Exped-A*, at least 10 mL of seawater were collected from n=79 different splash pools (Sp1-Sp79) using 25 cm² tissue-culture treated flasks along ~2 km of coastline, spanning from Boka Wandomi to Boka Pistol (Figure 3d; Table S1). Fifteen of these splash pools (n=15) were randomly selected for daily monitoring of evaporation and refilling cycles over an 8-day time course (Figures 3d, g-h; Extended Data Figure 3; Table S1). Among these, n=7 splash pools experienced gradual evaporation without desiccation; n=4 splash pools experienced complete desiccation; and n=4 splash pools were refilled. Each splash pool was uniquely identified using a physical tagging system and its GPS coordinates recorded with an iPhone 12 Mini (Apple). A photograph of each splash pool and its surrounding environment was also taken with the same device.

Exped-A on-site measurements: The following parameters were measured *in situ*: salinity, using a refractometer (#B07FQPFJGX (ASIN), Gain Express) (Figures 3e,h; Extended Data Figures

1 and 3; **Table S1**); temperature, using a thermometer (#B07CB8JG21 (ASIN), ThermoPro); and depth, using measuring tape. Splash pool seawater or soil (in dry splash pools) were collected for rehydration experiments (**Figure 3i-j**; **Video S16**; **Table S1**). The presence of sheets was later visually assessed at the CARMABI biological station in Curaçao using a Leica DM IL LED inverted microscope equipped with a Nikon Z 50 camera (**Figure 3**; **Extended Data Figures 1-3**; **Video S15**; **Table S1**). As controls, open sea salinity and temperature were measured in the bokas of *Boka Wandomi* and *Boka Kalki*.

Exped-B sampling: For *Exped-B*, a random number generator was used to select a randomised sampling location between 150 and 250 m upstream *Boka Wandomi*, avoiding sites previous sampled during *Exped-A* (**Figure 3d**; **Extended Data Figure 1c**). The selected location selected was 204 m upstream *Boka Wandomi*, where an area measuring 10 by 4 meters was mapped and defined (**Figures 3d'** and **6a-b**; **Extended Data Figure 1c**; **Table S1**). All splash pools containing at least 5 mL of seawater within this area (n=71) were collected and analysed as described above (M1-M77). New *C. flexa* strains (Strain 2 and Strain 3) were isolated from splash pools M44 and M60, respectively (**Figure 6a-b**) (see section '*Manual isolation of sheets collected in the field*').

Exped-C sampling: For *Exped-C*, at least 10 mL of seawater were collected from n=12 different splash pools (SpA-SpL) located near *Boka Wandomi* and *Boka Pistol* and analysed as described above. To maximise the likelihood of finding *C. flexa* sheets, the salinity of each splash pool was measured before sampling. Splash pools with salinity values within the permissive range for sheet occurrence (15 to 128 ppt) were selected for collection. Samples were later inspected using an inverted microscope to estimate *C. flexa* cell density based on the number of observed sheets (**Extended Figure 7b**; **Table S1**) (see section '*Cell density measurements of splash pool samples*').

Drone orthomap imagery of *Exped-B* (Figure 6b; Extended Data Figure 1c)

Orthomap imagery for *Exped-B* was obtained through drone-based nadir photography using a DJI Mavic 2 Enterprise. Low-altitude imagery was recorded at 4 m above the study area by manual drone operation, ensuring adequate image overlap for photogrammetric processing. The final orthomap was constructed using Agisoft Metashape Professional Edition v2.1.2⁴, resulting in a photogrammetrically corrected basemap of the study area with a spatial resolution of 1 mm per pixel. Georeferencing and spatial mapping of the orthomap were conducted in ArcGIS Pro (v3.0.0; ESRI, 2023).

Soil rehydration of splash pools (Figure 3i-k; Table S1; Video S16)

A total of n=32 soil samples were rehydrated across two independent fieldwork expeditions (*Exped-A* and *Exped-C*). Soil samples of n=6 splash pools that underwent complete desiccation during *Exped-A* (Sp6, Sp12, Sp15, Sp43, Sp69 and Sp70) were scraped and collected daily for eight days using a spatula into 25 cm² tissue-culture treated flasks, ensuring that soil was sampled from multiple areas within each splash pool to maximise sample representativity (**Video S16**). Collected soil samples were then rehydrated in the laboratory with 50-100 mL of sterile-filtered seawater collected from Boka Wandomi (filtered through a 0.22 µm pore-size filter; #SLGP033RS, Millex®-GP), adjusting the rehydration volume to reach a salinity of ~40 ppt. At least three independent rehydrations were performed for each soil sample. The presence of sheets was monitored daily during the next five days using a Leica DM IL LED inverted microscope equipped with a Nikon Z 50 camera at the CARMABI biology station in Curaçao (**Figure 3i-j; Table S1**).

Soil samples of n=26 additional splash pools (Soil1-Soil26) surveyed during *Exped-C* were collected as described above as an independent replicate experiment. Soil samples were rehydrated in the laboratory with 50-100 mL of sterile-filtered seawater collected from Playa Piskado (Westpunt, Curaçao) using a 0.22 µm pore-size filter (#SLGP033RS, Millex®-GP), adjusting the rehydration volume to reach a salinity below 100 ppt. The presence of sheets was monitored daily during the next five days using a Leica DM IL LED inverted microscope equipped with a Canon EOS REBEL T6i camera. Sheets were successfully recovered from soil sample *Soil4* 72 hours post-rehydration (**Figure 3k; Table S1**).

Manual isolation of sheets collected in the field (Figures 3f and 6a-c; Extended Data Figure 2)

Establishment of single-sheet-bottlenecked cultures: To establish new *Choanoeca flexa* strains from individual sheet colonies (single-sheet-bottlenecked, or SSB), we manually isolated single sheets from splash pool samples collected from M44, M60, and M61 (*Exped-B*, n=3 independent strain isolations from three different splash pools) as follows: for each sample, a first round of sheet isolation was performed by distributing 2 mL of the splash pool sample into multiple wells of a 6-well plate. Individual sheets were then manually pipetted (collecting a 0.2 µL volume with a micropipette for each sheet) and transferred into new 6-well plates containing 2 mL of 1X ASW per well. Each isolated sheet was transferred at least twice more into new wells of a 6-well plate (same volume and medium) to dilute away potential contaminants. Multiple sheets from the same

splash pool were then manually pipetted again (0.2 μ L volume) and transferred into clean 25 cm² tissue-culture treated flasks containing 20 mL of filtered seawater from Boka Wandomi and supplemented with 0.5% *H. oceani* resuspended food pellet. Cultures were subsequently reseeded from individual sheets via a second round of sheet isolation, following the same procedure: manual pipetting of a single sheet, at least two washes in a 6-well plate, and final transfer into a new 25 cm² tissue-culture treated flask containing 10 mL of SWC medium supplemented with 1% *H. oceani* resuspended food pellet. Flasks were maintained in SWC medium in a 25°C incubator under a 12-hour light-dark cycle.

Establishment of single-cell-bottlenecked cultures: To establish single-cell-bottlenecked (or clonal) cultures of newly collected *Choanoeca flexa*, we performed Fluorescence-Activated Cell Sorting (FACS) on previously established single-sheet-bottlenecked (SSB) cultures derived from M44 (Strain 2, M44B) and M60 (Strain 3, M60B), as follows. For each SSB culture, around 40 mL of exponentially growing cells in SWC medium were harvested at 3,300xg during 15 minutes at room temperature and resuspended in ~3 mL of 1X ASW by pipetting up and down. Sheets were then dissociated by vortexing the culture for 20 seconds and by adding 2.5 mM EGTA (#11453097, ThermoFisher Scientific, pre-diluted in Milli-Q water) to the cell suspension. Single cells were sorted into individual wells using a MoFlo Astrios Cell Sorter (Beckman Coulter) into three 96-well tissue-culture treated plates (#229195, Cell Treat), each well containing 200 μ L of 1X ASW supplemented with 0.1% SWC and 0.5% *H. oceani*. An additional wash with 1X ASW was performed between samples during sorting to prevent cross-contamination. Additional clones from ChoPs7 cultures (Strain 1) were established as controls, following either the same FACS protocol (clone 1G11) or alternatively by limiting dilution from a dense ChoPs7 culture diluted to 0.5 cells/mL in a 96-well plate containing SWC medium (clones 1A1 and 1A4). Single-cell-bottlenecked clonal cultures were grown for 6 days in 96-well plates and then expanded into a 75 cm² tissue-culture treated flasks containing SWC medium supplemented with 1% *H. oceani*.

18S rDNA sequencing and phylogenetic analysis (Extended Data Figure 2d; Table S2; Supplementary Files S1-S4)

18S rDNA sequencing: Genomic DNA (gDNA) was extracted from previously established single-sheet-bottlenecked and clonal cultures as follows. Approximately 40 mL of each exponentially growing culture in SWC medium were harvested at 3,300xg during 15 minutes at room temperature and resuspended directly in 30-120 μ L of DNAzol® Direct (#DN131, Molecular Research Centre). Cell lysis was promoted by pipetting up and down, followed by incubation of

the lysates at room temperature for at least 30 minutes. The 18S ribosomal DNA (18S rDNA) locus was amplified by PCR using the Q5® High-Fidelity DNA Polymerase (#M0491L, New England Biolabs) and degenerate primers (see sequences in **Table S2**). We followed the nested PCR program described in ¹. For the first round of PCR, 2.5 µL of gDNA (directly from the DNAzol-containing tubes) were used as template. For the nested PCR, 0.5 µL of the first PCR product were used as template. Both forward and reverse primers from the nested PCR, as well as an additional forward primer (named 'Sequencing', see **Table S2**), were used for Sanger sequencing. All experiments were performed in six independent biological replicates (n=6).

Phylogenetic analysis of 18S rDNA sequences: 18S rDNA sequences from newly isolated single-sheet-bottlenecked and clonal *C. flexa* cultures, as well as from other opisthokonts (including the previously published *C. flexa* 18S rDNA sequence; see **File S1**) ¹ were analysed using a custom phylogenetic workflow implemented in NGPhylogeny.fr ^{5,6}. Sequences were first aligned with MAFFT v7.467 L-INS-i algorithm under default parameters (**File S2**), and subsequently trimmed with Gblocks v0.91b ⁷ under minimally stringent default parameters (**File S3**). A Maximum Likelihood phylogenetic tree was then generated using PhyML v3.3.20190909 ^{8,9} under the General Time Reversible (GTR) model, with empirical nucleotide equilibrium frequencies, an estimated transition/transversion ratio, optimised proportion of invariable sites, estimated gamma model with four rate categories, and optimised parameters for tree topology, branch length, and substitution model, using a combination of NNI and SPR tree search methods (**File S4**). The resulting 18S rDNA tree was visualised and edited with FigTree v1.4.4 ([http:// tree.bio.ed.ac.uk](http://tree.bio.ed.ac.uk)) and further edited with Adobe Illustrator CC 2017 to produce the final tree figure.

Artificial evaporation of splash pool samples (Extended Data Figure 4)

3 mL of splash pool samples collected from Sp64 and Sp65 (*Exped-A*) were seeded into 6-well plates (#130184, ThermoFisher Scientific) and left uncovered at room temperature (between 25-28°C) on a bench to promote gradual evaporation. The presence of sheets was monitored daily during the next six days using a Leica DM IL LED inverted microscope equipped with a Nikon Z 50 camera at the CARMABI biology station in Curaçao. The experiment was performed in four independent replicates (two technical replicates for each splash pool sample, n=4).

Calculation of natural evaporation rates (Figures 4a-b and S5a)

To calculate natural evaporation rates, we used time-resolved salinity data collected from the twelve splash pools monitored in the field. We reasoned that a constant evaporation rate

(assumed to be proportional to the surface area of each splash pool) should result in a linear decrease in water volume over time (until saturation), following the relationship:

$$V(t) = V_0 - k \cdot t$$

where $V(t)$ is volume at time t , V_0 the initial volume, and k the evaporation rate.

Thus, before reaching saturation, the molar salt concentration can be expressed as:

$$c(t) = n_{\text{salt}}/V = n_{\text{salt}}/(V_0 - kt)$$

where n_{salt} is the total quantity of salt in a splash pool. It follows that:

$$1/c(t) = (V_0)/n_{\text{salt}} - (k/n_{\text{salt}}) \cdot t$$

and

$$c_0/c(t) = 1 - k \cdot V_0 \cdot t$$

In this model, when salinity is normalised to its initial value, the inverse of normalised salinity ($1/(\text{normalised salinity})$) should decrease linearly with time before reaching saturation. Remarkably, this simplified linear model fit the data well for all 12 splash pools ($r^2 = 0.94$ to 1; **Figure S10A**). We thus used the slope of the $1/(\text{normalised salinity}) = f(\text{time})$ as a proxy for the rate of salinity increase, which was then used to calibrate our artificial evaporation experiment.

Artificial evaporation experiment (Figures 4a-d and S5b-c; Video S17)

3 mL of dense ChoPs7 cells grown in SWC medium were seeded into four separate 6-well plates (#130184, ThermoFisher Scientific) and transferred on a grid inside a 30°C incubator. One 6-well plate was kept with its lid on (low-evaporation control), while the remaining plates were left uncovered (gradual evaporation condition) over a 9-day time course. A plastic box (dimensions: 35.7 cm x 23.5 cm x 13.5 cm) was placed over the plates to allow air exchange while minimizing contamination from airborne particles (**Figure 4a**). At each timepoint, a 50 µL sample was collected to measure salinity using a refractometer (#B07FQPFJGX (ASIN), Gain Express). An additional 100 µL sample was transferred to a black 96-well ibiTreat µ-plate (#89626, Ibidi) coated with poly-D-lysine (as described above). The number of single cells and sheets was quantified from at least 16 images per timepoint, captured using a Plan-Apochromat 20X/0.8 M27 Zeiss

objective on a Zeiss Axio Observer Z.1 inverted microscope, using a ColorBand filter (#FGL610, Thorlabs).

The number of cells per sheet was estimated based on sheet area, using the following method: first, cells were manually counted and sheet area was measured in 27 sheets ranging from 2 to 96 cells. Then, the estimated average area of a cell was calculated from this dataset. Linear regression of sheet area versus cell number yielded a regression coefficient of $r^2=0.98$, indicating that cell number can be reliably approximated as a linear function of sheet area. In subsequent analyses, the number of cells per sheet was estimated from sheet area. Doublets were excluded from the final analysis because their frequency was variable and because they may represent either aggregation events or dividing cells, thus not necessarily indicating incipient multicellularity. Once complete desiccation was achieved (after ~4 days of incubation), all dry samples from the 'gradual evaporation condition' were rehydrated with 3 mL of 1X ASW, transferred to a 25°C incubator with the plate lid on, along with the low-evaporation control samples, and monitored daily as before. A negative control flask containing 3 mL of 1X ASW and 3 mL of 5% SWC (without cells) was prepared and visually inspected under a Leica DM IL LED microscope to confirm that the 1X ASW stock used for rehydration was not contaminated with ChoPs7 or other cells. Gradual evaporation experiments were performed in eleven independent replicates (n=11). Low-evaporation control experiments were performed in six independent replicates (three biological replicates, each including two technical replicates; n=6).

Live imaging of sheet dissociation during gradual evaporation (Figures 4d and S5b; Video S17)

3 mL of dense ChoPs7 cells grown in SWC medium were seeded into three separate 6-well plates (#130184, ThermoFisher Scientific) and placed on a grid inside a 30°C incubator. The plates were left uncovered (gradual evaporation condition) to allow gradual evaporation (see section '*Artificial evaporation experiment*'). Evaporation was monitored until the samples reached a salinity of 82 ppt (approximately 48 hours after the start of the experiment). At that point, 600 µL of culture were collected from the bottom of a well and transferred to a black 96-well ibiTreat µ-plate (#89626, Ibidi). The samples were imaged every minute for 24 hours with a Plan-Apochromat 20X/0.8 M27 Zeiss objective on a Zeiss Axio Observer Z.1 inverted microscope, with the tile scan option, Definite Focus, and a ColorBand filter (#FGL610, Thorlabs). The lid of the plate was kept on throughout the experiment, except for two 1-hour intervals (after 3-4 hours and 5-6 hours from

the start of the timelapse) to maintain a controlled level of evaporation. At the end of the experiment, salinity reached 105 ppt.

Image processing and quantification: The timelapse (composed of two stacks) was binarized and analysed with Fiji Imaging Software version 2.14.0/1.54g. First, the stacks were converted to 8-bit using the 'Type' option and duplicated with the 'Duplicate' command in the 'Image' menu. To homogenise image brightness, a Gaussian blur filter (setting the radius value at 100) was applied to the duplicated stacks using the 'Filter' tool in the 'Process' menu. Next, the 'Calculator Plus' function in the 'Process' menu was used to normalise image intensity, combining the original and Gaussian-filtered stacks with the following parameters: i1=original stacks, i2=blurred stacks, Operation=Divide, and k1=150 (average image intensity). Intensity thresholds were later manually adjusted to minimise background artifacts (e.g., shadows from bacterial biofilms) and retain only choanoflagellate cells. Finally, the timelapse was binarized, closed, and dilated using the 'Make Binary', 'Close', and 'Dilate' options in the 'Binary' tool from the 'Process' menu. Detailed time course quantification was performed for two biological replicates, and six additional replicates were quantified only at the post-rehydration end point (see section '*Survival of cysts and flagellates after desiccation*').

Live imaging of cell rehydration after gradual evaporation (Figure 4e; Video S18)

3 mL of dense ChoPs7 cells grown in SWC medium were seeded into three separate 6-well plates (#130184, ThermoFisher Scientific) and placed on a grid inside a 30°C incubator. The plates were left uncovered (gradual evaporation condition) to allow gradual evaporation (see section '*Artificial evaporation experiment*'). Samples were monitored until complete desiccation, which occurred approximately 72 hours after the start of the experiment. The desiccated samples were then rehydrated with 3 mL of 1X ASW per well. Cysts were collected from the bottom of the wells using a cell scraper (#08-100-241, Fisher Scientific) and 600 µL of the cyst-containing culture were transferred to a black 96-well ibiTreat µ-plate (#89626, Ibidi).

The samples were imaged every minute overnight, with the plate lid on, using a Plan-Apochromat 20X/0.8 M27 Zeiss objective mounted on a Zeiss Axio Observer Z.1 inverted microscope, using the Definite Focus mode and a ColorBand filter (#FGL610, Thorlabs). All experiments were performed in four independent replicates (two biological replicates, each including one or three technical replicates per condition; n=4).

Imaging loss of multicellularity after direct addition of salts (Extended Data Figure 5a-c)

Approximately 40 mL of exponentially growing ChoPs7 cells grown in SWC medium were harvested at 3,300xg for 15 minutes at room temperature, washed twice with 20 mL of 1X ASW, and vortexed for 30 seconds. Cells were counted to estimate cell concentration using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-15 μm using the 'Histogram' option. Next, 1×10^4 cells were seeded in 1.2 mL of SWC medium (at 1X salinity) in a 12-well plate (#130185, ThermoFisher Scientific) and incubated in a 25°C incubator for 6 days. After 3 days of growth in 1X salinity, salinity was gradually increased by 1X per day through the addition of a 10X ASW stock solution, until reaching either 3X or 4X salinity. A 1X salinity condition (no salt addition) was used as a control. Samples were imaged on the sixth day, after spending a total of 6 days at 1X salinity, 2 days at 3X salinity, or 1 day at 4X salinity following the gradual increase in salt concentration.

100 μL of each sample were transferred to a poly-D-lysine-coated 96-well plate (as described above) and imaged by DIC microscopy using a C-Apochromat 40X/1.2 W Corr M27 Zeiss objective on a Zeiss Axio Observer Z.1 inverted microscope, using a ColorBand filter (#FGL610, Thorlabs). All experiments were performed in four independent replicates (two biological replicates, each including two technical replicates per condition; $n=4$).

Growth rate after direct addition of salts (Extended Data Figure 5d-e)

Approximately 40 mL of exponentially growing ChoPs7 cells grown in SWC medium were harvested at 3,300xg for 15 minutes at room temperature, washed twice with 20 mL of 1X ASW, vortexed for 30 seconds, and imaged using a LUNA-II automated cell counter (LogosBiosystems). In this experiment, the particle detection range was not adjusted using the histogram option (as in previous assays). Instead, images generated by the cell counter were exported and analysed using batch quantification with a custom macro script in the Fiji Imaging Software version 2.14.0/1.54g. In brief, the 'Smooth' and the 'Subtract Background' tools (setting a rolling ball radius to 50) from the 'Process' menu were first applied. Images were then converted to binary with the 'Convert to Mask' command from the 'Binary' tool in the 'Process' menu using the default parameters. The resulting binary images were analysed using the 'Analyse Particles' command to measure particle area, with parameters set to a particle size range of 20-100 μm^2 and circularity of 0.01-1, excluding particles on the edges of the images.

Next, 1×10^4 - 1.2×10^4 cells were seeded in 1.2 mL of SWC medium (at 1X salinity) in a 12-well plate (#130185, ThermoFisher Scientific) and incubated in a 25°C incubator for 6 days. Salinity was gradually increased to 3X or 4X by addition of a 10X ASW solution, as described above. A 1X salinity condition (no salt addition) was used as a control. The presence of sheets and single cells (flagellates or cysts) was monitored daily, and sample volume, salinity, and cell number were quantified daily as described before. All experiments were done in four independent replicates (two biological replicates, each including two technical replicates per condition; n=4).

The growth rate in each condition was calculated over a 3-day interval during which salinity was gradually increased (from day 3 until day 6), as follows:

$$growth\ rate = \frac{\log_2\left(\frac{N_t}{N_0}\right)}{T}$$

where N_t corresponds to the number of cells at the final timepoint (t); N_0 corresponds to the initial number of cells at time 0; and T corresponds to the time interval between the final timepoint (t) and time 0.

Cyst imaging under gradual evaporation (Figure 4f-i)

Cyst formation was monitored daily over a four-day period by DIC microscopy, acquiring images using C-Apochromat 40X/1.1 water, C-Apochromat 63X/1.4 oil-immersion, or C-Apochromat 100X/1.4 oil-immersion Zeiss objectives mounted on a Zeiss Observer Z.1 inverted microscope equipped with a Hamamatsu ORCA-Flash 4.0 V2 CMOS camera (C1140-22CU). Approximately 200 mL of dense Chops7 cultures grown in 1X ASW (33-35 ppt salinity) and CGM3 medium were transferred into a Bio-Assay Dish (#240845, ThermoScientific). On Day 1, the culture was placed in a 28°C incubator with the lid partially open to allow gradual evaporation. After 4-6 hours, when salinity reached 60 ppt, the lid was closed overnight to allow cells to adapt to the new salinity, and the temperature was increased to 29°C. On the morning of Day 2, the lid was partially reopened to resume gradual evaporation until salinity reached 80 ppt. The lid was closed overnight, and the temperature was increased to 30°C. On Day 3, the lid was opened until salinity reached 100-110 ppt (equivalent to 3X ASW). The culture was then maintained at 3X ASW for 24 hours to allow additional cells to transition into cysts. On the morning of Day 4, cells were imaged as before.

For imaging, 2 mL of culture were transferred each day into a FluoroDish (#FD35-100, World Precision Instruments) pre-treated with poly-D-lysine and incubated for 30 minutes to allow settling at the bottom of the plate prior to imaging. All experiments were performed in three biological replicates, each including two technical replicates per condition (n=6).

Cyst F-actin staining (Figures 4j-l and S6)

Approximately 40 mL of exponentially growing ChoPs7 cells grown in SWC medium were counted using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-15 μm using the 'Histogram' option. Next, 7×10^5 ChoPs7 cells were seeded in 3 mL of SWC medium in four separate 6-well plates (#130184, ThermoFisher Scientific) without lids and incubated at 30°C to allow gradual evaporation (see section '*Artificial evaporation experiment*'). Two additional 6-well plates were kept with lids on as low-evaporation controls. Samples were monitored daily, and salinity was measured after 4 days of incubation. Next, samples were scraped, homogenised, and collected in 1.5 mL microcentrifuge tubes. For the low-evaporation control (flagellates), 1 mL of culture was directly transferred to a microcentrifuge tube. All tubes were centrifuged at 3,300xg for 15 minutes at room temperature. The supernatants were carefully removed, and samples were resuspended with 375 μL of 1X ASW. The low-evaporation control sample was additionally vortexed for 30 seconds to dissociate sheets into single cells. All samples were then fixed by adding 16% ice-cold PFA (#15710, Electron Microscopy Sciences) to a final concentration of 4% and incubated for 5 minutes at room temperature. Next, 100 μL of each sample were seeded into a black 96-well coated with poly-D-lysine and permeabilised with 0.1% Triton X-100 (as described before). Cells were stained with a 10X dye mix solution diluted in Milli-Q water containing 1:100 FM 4-64X (from a 5 mg/mL stock solution in distilled water; #F34653, Invitrogen), 1:100 Alexa Fluor 488 Phalloidin (from a 2 units/ μL stock solution in DMSO; #A12379, Invitrogen), and 1:100 Hoechst (from a 1 $\mu\text{g}/\text{mL}$ stock solution; #H21486, Invitrogen), reaching a final concentration of 1:1000 (relative to each stock solution).

Samples were imaged using a Plan-Apochromat 40X/1.3 Oil DIC (UV) VIS-IR M27 Zeiss objective with Airyscan MPLX SR-4Y and 2X line averaging modes on a Zeiss Axio Observer Z1/7 inverted microscope equipped with LSM900 Airyscan 2. All experiments were performed in three biological replicates, each including two technical replicates (n=6). A total of n=12 cells (n=10 cysts and n=2 flagellate controls) were imaged.

Image processing and quantification: Images were processed using the 'Airyscan process' function in the 'Image Processing' menu of the Zen Blue software and later analysed with Fiji Imaging Software version 2.9.0/1.53. A straight line was drawn transversally across the equatorial plane of each cell and in the F-actin channel using the 'Straight Line' tool in the Fiji toolbar. The fluorescence intensity profile of F-actin across the cell was then displayed using the 'Plot Profile' function in the 'Analyse' menu.

Growth quantification after gradual evaporation (Figures 4m and S7)

Approximately 40 mL of ChoPs7 cells from an exponentially growing culture grown in SWC medium were counted to estimate cell concentration using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-15 μm using the 'Histogram' option. Next, 2.5×10^5 cells were seeded in 3 mL of SWC medium in four separate 6-well plates (#130184, ThermoFisher Scientific) without lid and incubated at 30°C during a 5-day time course to let samples evaporate gradually (see section '*Artificial evaporation experiment*'). Two additional 6-well plates were kept with their lid on as low-evaporation controls. Sample volume, salinity and cell number were quantified daily as before. Lids of plates in 'gradual evaporation' condition were closed to stop evaporation once samples reached either 2X or >3X salinity compared to 1X control samples.

The growth rate in each condition was calculated over the 5-day time interval as described above (see section '*Growth rate after direct addition of salts*'). All experiments were performed in three biological replicates, each including at least two technical replicates per condition (n=6).

Cyst nucleus-to-cytoplasm ratio calculation (Figure S8)

Approximately 20 mL of exponentially growing ChoPs7 cells grown in SWC medium were harvested at 3,300xg during 15 minutes at room temperature and counted using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-15 μm using the 'Histogram' option. Next, 6×10^5 ChoPs7 cells were seeded in 3 mL of SWC medium in two separate 6-well plates (#130184, ThermoFisher Scientific) without lids and incubated at 30°C to allow gradual evaporation until complete desiccation (see section '*Artificial evaporation experiment*'). An additional 6-well plate was kept with its lid on as a low-evaporation control. Samples were monitored daily until complete desiccation was reached after ~48 hours incubation, after which all 6-well plates were incubated for an additional 24 hours at 30°C with lids on. The desiccated plates were then rehydrated with 1 mL of 1X ASW, and samples were scraped,

homogenised, and collected into microcentrifuge tubes. For the low-evaporation control (flagellates), 1 mL of culture was directly transferred to a microcentrifuge tube. All tubes were centrifuged at 3,300xg for 15 minutes at room temperature. The supernatants were carefully removed, and samples were resuspended with 500 µL of 1X ASW. The low-evaporation control sample was additionally vortexed for 30 seconds to dissociate sheets into single cells. All samples were fixed by adding 16% ice-cold PFA (#15710, Electron Microscopy Sciences) to a final concentration of 4% and incubated for 5 minutes at room temperature. Next, 100 µL of each sample were mounted into a poly-D-lysine-coated 96-well plate, permeabilised, and stained with FM 4-64FX, Alexa 488 Phalloidin, and Hoechst, as described above.

At least 10 cells per condition were imaged using a Plan-Apochromat 40X/1.3 Oil DIC (UV) VIS-IR M27 Zeiss objective with Airyscan MPLX SR-4Y and 2X line averaging modes in a Zeiss Axio Observer Z1/7 inverted microscope equipped with LSM900 Airyscan 2. All experiments were performed in three biological replicates, each including two technical replicates per condition (n=6). A total of n=65 cells were imaged (n=27 cysts and n=38 flagellate controls).

Image processing and quantification: Images were processed using the 'Airyscan process' function in the 'Image Processing' menu of the Zen Blue software and subsequently analysed with Imaris Imaging software version 9.9.1 (build 61122 for x64). The 'Volume' tool was used to generate 3D renderings of the images. The 'Surfaces' tool was used for image segmentation and quantification of the volume and sphericity of both the cell body (excluding the collar and flagellum) and the nucleus of each individual cell.

The nucleus-to-cytoplasm ratio (N:C) was calculated as follows:

$$N:C\ ratio = \frac{Volume_{Nucleus}}{Volume_{Cell\ body} - Volume_{Nucleus}}$$

Latrunculin B effect on sheet dissociation (Extended Data Figure 6a-d)

Approximately 20 mL of ChoPs7 cells from an exponentially growing culture in SWC medium were harvested at 3,300xg for 15 minutes at room temperature, washed twice with 40 mL of 1X ASW, and resuspended in 1 mL of 1X ASW by pipetting up and down. Cells were dissociated by vortexing for 20 seconds and counted to estimate cell concentration using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-11 µm using the 'Histogram' option. Before seeding, samples were vortexed again for 30 seconds to ensure

complete dissociation into single cells. Next, 5×10^4 cells were seeded in 1 mL of 1X ASW supplemented with 1% SWC (without additional bacteria) in a 24-well plate and incubated at 30°C for 24 hours. After incubation, cells were treated with 2 μ M Latrunculin B (#L5288-1MG, Sigma-Aldrich; from a 253 μ M stock solution pre-diluted in 100% ethanol). The equivalent volume of 100% ethanol and 1X ASW were added as vehicle and negative controls.

Samples were imaged after 4 hours of treatment at 6 distinct locations throughout each well by TL brightfield microscopy using an EC Plan-Neofluar 5X/0.16 M27 Zeiss objective on a Zeiss Axio Observer Z.1 inverted microscope. The experiment was performed in three biological replicates, each including two technical replicates per condition (n=6). A minimum of n=465 particles were analysed in each condition.

Image processing and quantification: Images were batch processed in Fiji Imaging Software version 2.14.0/1.54f using a custom macro script. In brief, images were smoothed three times using the 'Smooth' command in the 'Process' menu. A 'Subtract Background' step was applied to reduce image shadows, setting a rolling ball radius of 50 pixels. Next, images were binarized using "Process > Convert to Mask" and analysed using 'Analyse Particles' to quantify area and circularity. Parameters were set at particle size range $>25 \mu\text{m}^2$ and circularity from 0.01 to 1, excluding particles on the edges of the images.

A sample from each condition in one of the replicates was fixed and stained with FM 4-64X and Alexa Fluor 488 Phalloidin, as described above, and imaged using a Plan-Apochromat 40X/1.3 Oil DIC (UV) VIS-IR M27 Zeiss objective with Airyscan MPLX SR-4Y and 2X line averaging modes on a Zeiss Axio Observer Z1/7 inverted microscope equipped with LSM900 Airyscan 2. Images were then processed using the Zen Blue software. A maximum intensity z-projection was applied with Fiji Imaging Software version 2.14.0/1.54f.

Latrunculin B effect on cell aggregation (Extended Data Figure 6e-h)

Approximately 40 mL of ChoPs7 cells from an exponentially growing culture grown in SWC medium were harvested at 3,300xg for 15 minutes at room temperature, washed twice with 40 mL of 1X ASW, and resuspended in 1 mL of 1X ASW by pipetting up and down. Cells were then dissociated by vortexing for 20 seconds and counted to estimate cell concentration using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-11 μ m using the 'Histogram' option. Before seeding, samples were vortexed again for 30 seconds to

ensure complete dissociation into single cells. Next, 5×10^4 cells were seeded in 1 mL of SWC medium in a 24-well plate and immediately treated with 2 μ M Latrunculin B (#L5288-1MG, Sigma-Aldrich; from a 253 μ M stock solution pre-diluted in 100% ethanol). The equivalent volume of 100% ethanol and 1X ASW were respectively added to the vehicle control and negative control conditions. Cells were incubated at 30°C for 4 hours.

Samples were imaged after 4 hours of treatment at 6 distinct locations throughout each well by TL brightfield microscopy using an EC Plan-Neofluar 5X/0.16 M27 Zeiss objective on a Zeiss Axio Observer Z.1 inverted microscope. The experiment was performed in three biological replicates, each including two technical replicates per condition (n=6). A minimum of n=2900 particles were analysed in each condition.

Image processing and quantification: Images were batch processed in Fiji Imaging Software version 2.14.0/1.54f using a custom macro script as before (see section '*Latrunculin B effect on sheet dissociation*'). A sample of each condition in one of the replicates was fixed, stained, and imaged as before.

Survival of cysts and flagellates after desiccation (Figure 4o-p)

Survival of cysts after desiccation: 4 mL of a dense ChoPs7 culture grown in SWC medium were seeded in each well of a 6-well plate (#130184, ThermoFisher Scientific) and incubated without lid in a 30°C incubator (see section '*Artificial evaporation experiment*'). All sheets were dissociated and cells had switched to a cyst-like morphology after 48 hours. Complete evaporation occurred after 72 hours. Cysts were then rehydrated with 4 mL of 1X ASW and transferred to a 25°C incubator maintained at 60% humidity. 8 days post-rehydration, all wells contained abundant, actively swimming sheet colonies. Cells were counted using the standard protocol with a LUNA-II automated cell counter (LogosBiosystems, as described before).

Survival of flagellate cells after desiccation: 0.5 mL of a dense ChoPs7 culture grown in SWC medium were seeded in each well of a 6-well plate (#130184, ThermoFisher Scientific) and incubated without lid in a 30°C incubator. Evaporation was complete after 21 hours. Desiccated cells retained a visible flagellum and collar and remained interconnected through their microvilli, indicating that they had not differentiated into cysts. The desiccated flagellates were rehydrated with 0.5 mL of 1X ASW and transferred to a 25°C incubator at 60% humidity. 8 days post-

rehydration, no living cells were observed in any of the wells, and the LUNA-II automated cell counter (LogosBiosystems) did not detect any cell either.

To image desiccated sheets, 1 mL of a dense Chops7 culture was seeded onto a FluoroDish (#15159112, Fisher Scientific), transferred to a 30°C incubator without lid, and incubated overnight to promote rapid evaporation. After 24 hours, the sample was completely desiccated. For imaging, the sample was rehydrated with 1 mL of 1X ASW and immediately imaged by DIC microscopy using a C-Apochromat 63X/1.20 W corr UV-VIS-IR Zeiss objective (#421787-9970-000, Zeiss) on a Zeiss Axio Observer Z.1 inverted microscope. All experiments were performed in six independent technical replicates (n=6).

Quantification of prey capture (Figure 4q-r)

Bacterial staining: A *H. oceanii* food pellet (20 mg) was resuspended in 1 mL of 1X ASW. Bacteria were stained with BactoView-Live Green [FITC] (#40102, Biotium) at 2X concentration (4 µL per mL) and incubated for 30 minutes at room temperature in the dark. To remove unincorporated dye, the stained bacteria were centrifuged at 2,750x g for 5 minutes, and the supernatant was carefully removed using a transfer pipette. The stained bacterial pellet was then resuspended in 1mL of 1X ASW.

Prey capture assay: ChoPs7 cultures were grown to high density in CGM3 medium. To assess colony capture efficiency, 200 µL of colonies were transferred into µ-slide 8-well chambers (#80826, Ibidi) and incubated for 30 minutes to let sheets settle. In parallel, to obtain single cells, part of the culture was transferred into a 15 mL falcon tube and vortexed at high speed ('fast' setting) on a Vortex Genie 2 for 1 minute to dissociate colonies into single cells. Cells were allowed to recover for 5 minutes before being transferred into µ-slide 8-well chambers (#80826, Ibidi). Stained bacteria were diluted 1:20 in 1X ASW, and 100 µL were added to each well containing colonies or single cells. *C. flexa* was incubated with bacteria for 1 minute before fixation with 16% ice-cold PFA (#15710, Electron Microscopy Sciences) to a final concentration of 4%. After fixation, samples were left to settle – 30 minutes for colonies and 2 hours for single cells – before imaging. All experiments were performed in three independent biological replicates, each including two technical replicates per condition (n=6).

Imaging and quantification: Colonies and single cells were imaged by DIC and green fluorescence microscopy to visualise stained bacteria, using a C-Apochromat 40X/1.1 optovar 1.6X Water

Zeiss objective on a Zeiss Observer Z.1 inverted microscope equipped with a Hamamatsu ORCA-Flash 4.0 V2 CMOS camera (C1140-22CU). To quantify capture efficiency, we calculated the ratio between the number of bacteria attached to the collars of choanoflagellate cells to the total number of choanoflagellate cells in each image. A minimum of n=163 particles were quantified in each condition.

Normalised growth and efficiency of aggregation over 2 hours at different salinities (Figure 5a; Extended Data Figure 7a)

Normalised growth: To calculate normalised growth under 1X and 2X salinity conditions, the growth rates calculated in the experiment shown in **Figure 4m** (see section ‘*Growth rate quantification after gradual evaporation*’) were normalised (for each experimental condition C) by the maximal growth rate measured in the experiment (across all conditions):

$$normalised\ growth_{condition\ C} = \frac{growth\ rate_{condition\ C}}{max.\ growth\ rate_{all\ conditions}}$$

Efficiency of aggregation: Chops7 cultures were grown to maximal density in 1X ASW and CGM3 medium. To obtain cultures in 2X ASW, cells were concentrated by centrifugation the day before the experiment, resuspended in 2X ASW, and incubated for 24 hours. On the day of the experiment, colonies from both culture conditions (1X and 2X ASW) were transferred into 50 mL Falcon tubes and vortexed at high speed (‘fast’ setting) on a Vortex Genie 2 for one minute to dissociate colonies into single cells. Single cells were concentrated by centrifugation to 1×10^6 cells/mL, and 200 μ L were transferred into a 96-well plate (#165305, Nunc). Cells were imaged by widefield microscopy at 4 different locations per technical replicate every 10 minutes for 2 hours using the ‘Tile positions’ and ‘Time points’ options. The microscope used was a Nikon Eclipse Ti2-E inverted microscope equipped with a D-LED light source, Chroma 8901 Quad Filter Cube, a 20X CFI Plan Apo VC NA1.2 objective, and a Nikon Digital Sight 50 M Camera. Aggregate area was quantified using ImageJ Imaging Software version 2.14. Tiled images were processed using the following commands with default settings in the following order: ‘Smooth’ and ‘Find Edges’ (from the ‘Process’ menu); ‘Despeckle’ (in the ‘Process’ menu), ‘Make Binary’, ‘Dilate’, ‘Erode’, and ‘Fill Holes’ (from the ‘Binary’ tool in the ‘Process’ menu); and ‘Analyse Particles’ (in the ‘Analyse’ menu). All experiments were performed in three independent biological replicates, each including two technical replicates per condition (n=6). A minimum of 1832 particles were quantified per timepoint and condition.

The *Efficiency of Aggregation* (E_a) in each experimental replicate R was calculated at the 2-hour timepoint as follows (assuming negligible contribution of cell division over 2 hours, which is only 15-20% of the duration of a cell cycle):

$$E_a = \frac{Area_{replicate\ R}}{max. Area_{2X}}$$

Cell density measurements in splash pools (Extended Data Figure 7b; Table S1)

To estimate *C. flexa* cell density in field samples, 500 μ L of splash pool seawater from n=12 different splash pools collected during *Exped-C* (SpA-SpL) were transferred into a 24-well plate using a cut P1000 tip. Samples were incubated for at least 10 minutes at room temperature to allow sheets to settle at the bottom of the wells. Within each well, sheets were manually counted on a Leica DM IL LED inverted microscope equipped with a Canon EOS REBEL T6i camera. The *C. flexa* cell concentration range was estimated from microscopic observations by considering the observed minimum and maximum number of cells per sheet (**Table S1**). All experiments were performed in twelve independent replicates, quantifying each sample in two technical replicates (n=12).

Beforehand, the accuracy of the manual counting method was assessed using ChoPs7 laboratory cultures. To do so, we quantified the cell concentration of a ChoPs7 culture using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-15 μ m using the 'Histogram' option, and compared it with the cell concentration estimated using the manual counting method described above. The automated cell counter quantified an average cell concentration of 1.0×10^5 cells/mL, which fell within the estimated range of $6.0 \times 10^3 - 4.0 \times 10^5$ cells/mL, confirming the accuracy of the manual counting method. All experiments were performed in three technical replicates (n=3).

Efficiency of clonality and aggregation after 24 hours at different cell densities (Figure 5b; Extended Data Figure 7c)

Approximately 20 mL of ChoPs7 cells from an exponentially growing culture grown in SWC medium were harvested at 3,300xg for 15 minutes at room temperature and resuspended in 2 mL of 1X ASW by pipetting up and down. Cells were then dissociated by vortexing for 20 seconds and counted to estimate cell concentration using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-11 μ m using the 'Histogram' option.

Before seeding, samples were vortexed again for 30 seconds to ensure complete dissociation into single cells. Next, 10^2 , 10^3 , 10^4 , and 10^5 cells were seeded in 1 mL of either 1X ASW or 2X ASW supplemented with 1% SWC and 0.5% *H. oceanii* in a 24-well plate. Cells were incubated at 30°C and imaged after 24h, 48h and 72h incubation at 5 distinct locations throughout each well by TL brightfield microscopy using an EC Plan-Neofluar 5X/0.16 M27 Zeiss objective on a Zeiss Axio Observer Z.1 inverted microscope. The experiment was performed in three biological replicates, each including two technical replicates per condition (n=6). A minimum of n=390 particles were analysed in each condition (except for the lowest cell density, where at least 50 particle areas were quantified).

Image processing and quantification: Images were batch processed in Fiji Imaging Software version 2.14.0/1.54f using a custom macro script. In brief, images were first smoothed five times using the 'Smooth' command from the 'Process' menu. A 'Subtract Background' step was applied to reduce shadows, using a rolling ball radius of 5 pixels and the sliding paraboloid option. Then, images were converted to binary using the 'Convert to Mask' command in the 'Process' menu. Images were then analysed using the 'Analyse Particles' command to calculate the area of each particle, setting the analysis parameters to particle size $>10 \mu\text{m}^2$ and circularity from 0.05 to 1, excluding particles on the edges of the images.

We defined the efficiency of aggregation (E_a) as the maximal size reached by aggregation alone at 2X salinity, normalised by the largest size measured at 2X salinity at the 24-hour timepoint. We defined the efficiency of clonality (E_c) as the additional growth allowed by cell division at 1X salinity relative to the aggregative baseline measured at 2X salinity.

The *Efficiency of Aggregation* (E_a) and the *Efficiency of Clonality* (E_c) in each condition were thus calculated based on areas measured after 24 hours with the following formulas:

$$E_a = \frac{Area_{2X}}{max.Area_{2X}} ; \text{ and } E_c = \frac{Area_{1X}}{Area_{2X}} - 1 = \frac{Area_{1X} - Area_{2X}}{Area_{2X}}$$

where $Area_{2X}$ corresponds to the area of particles in aggregative conditions (salinity = 2X), $max.Area_{2X}$ corresponds to the maximum area value of particles measured under aggregative conditions (salinity = 2X), and $Area_{1X}$ corresponds to the area of particles in clonal-aggregative conditions (salinity = 1X). This definition relies on the fact that aggregation operates at comparable

speed under both salinities, as demonstrated by live imaging over 2 hours (see section ‘Normalised growth and efficiency of aggregation over 2 hours at different salinities’).

As an important control, we first verified that our metrics for efficiency of aggregation and clonality were not saturated. This concern arises because sheet size plateaus at an approximately maximal size of $\sim 1,800 \mu\text{m}^2$, after which further growth is limited by fragmentation. In principle, saturation could thus bias the ‘efficiency of clonality’ metric: if sheets at 2X salinity already reached maximal size via aggregation alone, colonies at 1X salinity could not become larger – even if clonal division occurred – thereby artifactually yielding an E_c to be 0. To avoid this artifact, we selected the 24-hour timepoint, when sheets had not yet reached their maximal size (typically attained between 48 and 72 hours, depending on conditions; sheets at 24 hours did not exceed $\sim 1,200 \mu\text{m}^2$). Thus, the low efficiency of clonality observed at 10^5 cells/mL most likely reflects true inhibition of cell division at high density, rather than a measurement saturation effect.

Airyscan imaging comparing aggregative sheets versus control (Figure 5d-f; Extended Data Figure 7d-j)

For control sheets: Approximately 20 mL of ChoPs7 cells from an exponentially growing culture in SWC medium were harvested at 3,300xg for 15 minutes at room temperature and resuspended in 1 mL of 1X ASW by pipetting up and down. Cells were manually counted using a Neubauer-Improved hemocytometer (#0640010, Marienfeld) to estimate cell concentration. Next, 5×10^4 cells were seeded in 1 mL of SWC medium in a 24-well plate and serially diluted 1:3 five times (mixing by pipetting up and down 10 times per well), resulting in a final low cell concentration of ~ 200 cells/mL to ensure clonal expansion of sheets. Cells were incubated for 3 days in a 30°C incubator.

For aggregative sheets: Approximately 20 mL of ChoPs7 cells from an exponentially growing culture in SWC medium were harvested at 3,300xg during 15 minutes at room temperature and resuspended in 1 mL of 1X ASW by pipetting up and down. Cells were dissociated by vortexing for 30 seconds and counted using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-11 μm using the ‘Histogram’ option. Before seeding, samples were vortexed again for 30 seconds to ensure complete dissociation into single cells. Next, 1×10^5 cells were seeded in 1 mL of 2X ASW and 1% SWC supplemented with 1% *H. ocean*i resuspended food pellet in a 24-well plate and incubated for 24 hours in a 30°C incubator.

Fixation and imaging protocol: After incubation, 100 μ L of cells were pipetted using a truncated P200 tip into a black 96-well ibiTreat μ -plate (#89626, Ibidi) pre-treated with 200 μ L of 100 μ g/mL Poly-D-Lysine hydrobromide (#P6407, Sigma-Aldrich), allowing cells to settle for 10 minutes at room temperature. Then, cells were fixed by adding 16% ice-cold PFA (#15710, Electron Microscopy Sciences) to a final concentration of 4% and incubated for 5 minutes at room temperature. Next, cells were permeabilised for 5 minutes at room temperature using a 0.1% Triton X-100 solution (#A16046, ThermoFisher Scientific) prepared from a 10% stock solution in distilled water. Finally, cells were stained with a 10X dye mix solution diluted in Milli-Q water containing 1:100 FM 4-64X (5 mg/mL stock in Milli-Q water; #F34653, Invitrogen) and 1:100 Alexa Fluor 488 Phalloidin (2 units/ μ L stock in DMSO; #A12379, Invitrogen), reaching a final concentration of 1:1000 for each dye. Cells were then imaged using a Plan-Apochromat 40X/1.3 Oil DIC (UV) VIS-IR M27 Zeiss objective with Airyscan MPLX SR-4Y and 2X line averaging modes in a Zeiss Axio Observer Z1/7 inverted microscope equipped with LSM900 Airyscan 2. All experiments were performed in three biological replicates, each including two technical replicates per condition (n=6). A total of n=87 sheets (n=41 aggregative; n=46 control) were quantified.

Image processing and quantification: Images were processed using the Zen Blue software and later analysed with Imaris Imaging software version 9.9.1 (build 61122 for x64). A 3D volume rendering of each sheet was created using the 'Volume' tool, and the number of cells with misaligned apicobasal orientation was manually counted. Next, the 'Surfaces volume' option was used to automatically quantify the number of cells of each sheet, with manual corrections applied using the 'Cut Surface' option when needed to refine segmentation. Two different cross-sections per sheet (each containing at least 4 cells in a row) were visualised with the 'Oblique Slicer' tool. Angles between collar contacts of neighbouring cells were manually measured using the 'Surface of Object' intersect function from the 'Measurement Points' tool by drawing a polygon between the intersection of the actin signal at the base of both cell bodies and the tip of the collar-collar contacts. Cells with the same apicobasal orientation had collar-collar angles $<120^\circ$.

In parallel, images were batch processed in Fiji Imaging Software version 2.14.0/1.54f using a custom macro script. In brief, the brightness and contrast were automatically adjusted, and a maximum intensity z-projection was applied to all images. Then, image contrast was enhanced five times using the 'Enhance contrast' command from the 'Process' menu, setting a saturation value of 0.35. Images were then stacked into a single RGB composite and converted to 8-bit. Thresholds were manually adjusted using the 'Adjust Threshold' in the 'Image' menu. Next,

images were binarized using the 'Convert to Mask' and 'Fill holes' commands in the 'Process' menu. The 'Wand (tracing)' tool was used to select the colony outline, and the 'Convex Hull' tool from the 'Selection' option (inside the 'Edit' menu) was used to draw the outline of each colony, which was then added to the 'ROI Manager'. All colonies were analysed using the 'Measure' command at the 'ROI Manager' to calculate the area and circularity of each sheet.

Adding green cells to preformed sheets in aggregative conditions (Figure 5g-j)

Approximately 30 mL of ChoPs7 cells from an exponentially growing culture in SWC medium were harvested at 3,300xg for 15 minutes at room temperature and washed twice with 40 mL of 1X ASW. Cells were then resuspended in 500 μ L of 1X ASW by pipetting up and down, dissociated by vortexing for 20 seconds, and counted to estimate cell concentration using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-15 μ m using the 'Histogram' option. Before seeding, samples were vortexed again for 30 seconds to ensure complete dissociation into single cells. Next, 5×10^4 cells were seeded in 1 mL of 2X ASW and 1% SWC supplemented with 1% *H. ocean*i resuspended food pellet in a 24-well plate and incubated for 24 hours in a 30°C incubator. The next day, another 30 mL of ChoPs7 cells from an exponentially growing culture in SWC medium were harvested, washed and stained with CellTrace CFSE (green) as described before (see section '*Live imaging of dual labelling of aggregates*'). Stained cells were resuspended in 1 mL of 1X ASW and counted using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-11 μ m using the 'Histogram' option. Before seeding, samples were vortexed again for 30 seconds to ensure complete dissociation into single cells. Next, 5×10^4 CFSE-stained single cells were added to the preformed sheets in the previously prepared 24-well plate and incubated in a 30°C incubator.

After 2 hours and 24 hours of incubation, 100 μ L of cells were pipetted using a truncated P200 tip in a black 96-well ibiTreat μ -plate (#89626, ibidi) pre-treated with 200 μ L of 100 μ g/mL Poly-D-Lysine hydrobromide (#P6407, Sigma-Aldrich) and allowed to settle for 10 minutes at room temperature. Then, cells were fixed by adding 16% ice-cold PFA (#15710, Electron Microscopy Sciences) to a final concentration of 4% and incubated for 5 minutes at room temperature. Next, cells were permeabilised for 5 minutes at room temperature using a 0.1% Triton X-100 solution (#A16046, ThermoFisher Scientific) added from a 10% stock solution in distilled water. Finally, cells were stained with a 10X dye mix solution diluted in Milli-Q water containing 1:100 FM 4-64X (5 mg/mL stock pre-diluted in Milli-Q water; #F34653, Invitrogen) and 1:100 Alexa Fluor Plus 405

Phalloidin (2 units/ μ L stock pre-diluted in DMSO; #A30104, Invitrogen), reaching a final concentration of 1:1000 for each dye.

Cells were then imaged with a Plan-Apochromat 40X/1.3 Oil DIC (UV) VIS-IR M27 Zeiss objective using Airyscan MPLX SR-4Y and 2X line averaging modes in a Zeiss Axio Observer Z1/7 inverted microscope equipped with LSM900 Airyscan 2. All experiments were performed in three biological replicates, each including two technical replicates per condition (n=6). A total of n=48 sheets were imaged and quantified (n=22 after 2 hours; n=26 after 24 hours).

Image processing and quantification: Images were then processed using the Zen Blue software and later analysed with Imaris Imaging software version 9.9.1 (build 61122 for x64) as before.

Light-to-dark-induced inversion in aggregative sheets (Figure 5k; Extended Data Figure 8; Video S19)

Inversion behaviour in aggregative sheets (unlabelled): Approximately 20 mL of ChoPs7 cells from an exponentially growing culture, grown in SWC medium supplemented with 0.1% *A. machipongonensis*, were harvested at 3,300xg during 15 minutes at room temperature and resuspended in 1 mL of 1X ASW by pipetting up and down, and dissociated by vortexing for 30 seconds. Cells were counted using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-11 μ m using the 'Histogram' option. Before seeding, samples were vortexed again for 30 seconds to ensure complete dissociation into single cells. Next, 1×10^5 cells were seeded in 1 mL of 2X ASW and 1% SWC supplemented with 1% *H. ocean*i resuspended food pellet in a 24-well plate and incubated for 24 hours in a 30°C incubator. For control conditions, 100 cells were seeded in 1 mL of SWC medium supplemented with 0.5% *H. ocean*i and 0.1% *A. machipongonensis* in a 24-well plate and incubated 3-5 days in a 30°C incubator.

Inversion behaviour in aggregatively produced sheets (dual-labelled): Approximately 20 mL of ChoPs7 cells from an exponentially growing culture, grown in SWC medium supplemented with 0.1% *A. machipongonensis*, were harvested, washed and stained with CellTrace Far Red (magenta) and CFSE (green) as before (see section 'Live imaging of dual labelling of aggregates'). Cells were resuspended in 1 mL of 1X ASW and counted using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-15 μ m using the 'Histogram' option. Before seeding, samples were vortexed again for 30 seconds to

ensure complete dissociation into single cells. Green and magenta single-cell populations were mixed in a 1:1 ratio. A final number of 1×10^5 cells (5×10^4 cells of each colour) were seeded in 1 mL of 2X ASW and 1% SWC supplemented with 1% *H. oceani* and 0.1% *A. machipongonensis* resuspended food pellets in a 24-well plate and incubated overnight in a 30°C incubator (protected from light by a layer of aluminium foil). All experiments were performed in three independent biological replicates, each including two technical replicates per condition (n=6). A total of n=60 sheets were imaged per condition.

Imaging acquisition and quantification: Sheets were transferred to a black 96-well ibiTreat μ -plate (#89626, Ibidi), pre-coated with 150 μ L of 1 μ g/mL poly-D-lysine (#P6407-5MG, Sigma-Aldrich) for around 1 minute, then washed twice with 500 μ L of 1X ASW, and imaged on a Leica Stellaris 5 confocal microscope with a HC PL FLUOTAR 10X/0.30 DRY objective in the FRAP mode. A white light laser was used to create a 488 nm laser line with 1% intensity and a 633 nm laser line with 1% intensity. For light-to-dark stimulation, the 488 nm laser line was turned on during the pre- and post-bleaching frames and turned off during bleaching frames, while the 633 nm laser line remained continuously on. Brightfield images were acquired with Trans PMT detector, and fluorescent signals were detected using two HyD type S, covering 495 nm - 605 nm spectra for 488 nm excitation and 635 nm - 745 nm spectra for 633 excitation, respectively. The imaging interval was around 0.19 seconds. Acquired images were exported in 'ImageJ TIFF' format using Leica Application Suite X.

Ten colonies from each replicate were subjected to light-to-dark transition to assess light-sensitive inversion behaviour (total of n=60 sheets). For quantitative analysis, five representative colonies per condition were analysed to measure changes in colony area during inversion. Cell segmentation was performed on brightfield images using a custom-trained model of Cellpose (v2.2.3). Gaps between cells in the colony were filled using 'binary_dilation' and 'binary_erosion' functions from the morphology module of the scikit image package (v0.25.1)^{10,11}. The resulting colony masks were quantified using the 'label' and 'regionprops_table' functions from the measure module of the same package. The quantified area was normalised to the mean area of the first 20 frames before the light-to-dark transition and plotted using tidyverse (v2.0.0) in R (v4.1.1) and RStudio (v2021.9.0.351).

Quantification of prey capture of aggregative versus control sheets (Figure 5I)

For control sheets: Approximately 20 mL of ChoPs7 cells from an exponentially growing culture, grown in 5% CGM3 medium, were harvested by centrifugation at 3,000xg for 15 minutes at room temperature, and resuspended in 1 mL of 1X ASW by pipetting up and down. Cells were counted using a LUNA-II automated cell counter (LogosBiosystems). Next, 5×10^4 cells were seeded in 1 mL of CGM3 medium in a 24-well plate and serially diluted 1:3 five times (pipetting up and down each well 10 times), resulting in a final low cell concentration of ~200 cells/mL to promote clonal expansion of sheets. Cells were incubated for 3 days at 30°C.

For aggregative sheets: Approximately 20 mL of ChoPs7 cells from an exponentially growing culture, grown in CGM3 medium, were harvested at 3,000xg during 15 minutes at room temperature and resuspended in 1 mL of 1X ASW by pipetting up and down, dissociated by vortexing for 30 seconds, and counted using a LUNA-II automated cell counter (LogosBiosystems). Before seeding, samples were vortexed again for 30 seconds to ensure complete dissociation into single cells. Next, cells were seeded at 1×10^5 cells/mL in 2X ASW and 5% CGM3 in a culture flask and incubated for 24 hours at 30°C.

Bacterial staining: Two separate *H. oceanii* food pellets (10 mg each) were resuspended in 500 μ L of 1X ASW and 500 μ L of 2X ASW, respectively. Bacteria were stained with BactoView-Live Green [FITC] (#40102, Biotium) at a final concentration of 2X (2 μ L of dye per 500 μ L suspension) and incubated at room temperature for 30 minutes in the dark. To remove unincorporated dye, stained bacteria were centrifuged at 2,750xg for 5 minutes, and the supernatant was carefully removed with a transfer pipette. The stained bacterial pellet was resuspended in 500 μ L of 1X ASW or 2X ASW.

Prey capture assay: To assess colony prey capture efficiency, 150 μ L of colonies were transferred into a 96-well plate (#165305, Nunc) and incubated for 5 minutes to let colonies settle. In parallel, to obtain single cells, 5 mL of the same culture were transferred into a 15 mL falcon tube and vortexed with at the 'fast' setting on a Vortex Genie 2 for 1 minute to dissociate colonies into single cells. In addition, a 30G needle was used to ensure complete dissociation of residual small colonies. Similarly, single cells were allowed to recover from mechanical stimulation for 5 minutes before transferring them into a 96-well plate (#165305, Nunc). Stained bacteria were diluted 1:1000 in 1X ASW or 2X ASW, and 50 μ L of the diluted bacterial suspension were added to each well containing colonies or single cells. *C. flexa* was incubated in the presence of bacteria for 1 minute before samples were fixed with 16% ice-cold PFA (#15710, Electron Microscopy

Sciences) to a final concentration of 4%. After fixation, colonies and single cells were incubated in the dark and at room temperature for 30 minutes and 1 hour, respectively, to allow them to settle at the bottom of the plate before imaging.

Imaging and quantification: Colonies and single cells were imaged by DIC microscopy and green fluorescence to visualise the stained bacteria. Imaging was performed on a Nikon Eclipse Ti2-E inverted microscope equipped with a D-LED light source, Chroma 8901 Quad Filter Cube, and either a 10X CFI Plan Apo VC NA1.2 objective (for single cells) or a 40X CFI Plan Apo VC NA1.2 objective (for colonies), using a Nikon Digital Sight 50 M Camera. To quantify prey capture efficiency, we calculated the ratio between the number of bacteria attached to the collar of choanoflagellate cells and the total number of choanoflagellate cells in each image. All experiments were performed in three independent biological replicates, each including two technical replicates per condition (n=6).

Kin recognition between *C. flexa* and *S. rosetta* (Extended Data Figure 9)

Approximately 40 mL of ChoPs7 cells from an exponentially growing culture and 20 mL of SrEpac cells from a mixed culture containing both fast and slow swimmers grown in SWC medium were harvested, washed and stained with CellTrace Far Red (magenta) and CFSE (green), respectively, as described before (see section '*Live imaging of dual labelling of aggregates*'). Cells were resuspended in 1 mL of 1X ASW and counted using a LUNA-II automated cell counter (LogosBiosystems), adjusting the particle detection range to 1-11 μm using the 'Histogram' option. Before seeding, samples were vortexed again for 30 seconds to ensure complete dissociation into single cells. Green and magenta single cell populations were then mixed in a 1:1 ratio, and a final number of 5×10^4 cells (2.5×10^4 cells of each colour) were seeded in 1 mL of SWC medium supplemented with 0.5% *H. ocean*i and 0.5% *E. pacifica* resuspended food pellets in a 24-well plate and incubated overnight in a 22°C incubator (protected from light by a layer of aluminium foil). Immediately after seeding (time 0h) and after overnight incubation (time 20-22h), 100 μL of cells were transferred using a cut P200 tip into a black 96-well ibiTreat μ -plate pre-coated with 200 μL of 100 $\mu\text{g/mL}$ Poly-D-Lysine hydrobromide (#P6407, Sigma-Aldrich; stock solution in distilled water). Cells were incubated for 10 minutes at room temperature, fixed with 16% ice-cold PFA (PFA; #15710, Electron Microscopy Sciences) to a final concentration of 4%, and further incubated for 5 minutes at room temperature.

The wells were scanned and imaged in at least 15 different positions per condition using TL brightfield and 40% intensity LEDs 488 nm and 650 nm, with a Plan-Apochromat 20X/0.8 M27 Zeiss objective for quantification on a Zeiss Axio Observer Z.1 inverted microscope. DIC microscopy using a C-Apochromat 40X/1.2 W Corr M27 Zeiss objective was used for figure illustration on a Zeiss Axio Observer Z.1 inverted microscope. The experiment was performed in three biological replicates, each including two technical replicates per condition and single-colour unmixed controls for each strain (n=6). A total of n=211 colonies were quantified.

Whole-genome sequencing: assembling the *C. flexa* reference genome (Figure 6)

C. flexa culture and snap freezing: ChoPs cultures for genome sequencing and assembly were established by thawing a low-passage, polyxenic, light-responsive *C. flexa* culture originally isolated in 2018 and reported in a previous study¹. This culture was not isolated from Shete Boka National Park, but rather from another site along the Curaçao coast where *C. flexa* was first discovered (12°13'38.9' N 69°00'47.0' W). Monoxenicity was established as previously described¹ by antibiotic treatment and addition of live *H. oceanii* bacteria. The resulting ChoPs strain (referred to as 'ChoPs8') was grown to maximal density ($\sim 1 \times 10^6$ cells/mL) in 5% CGM3 medium. Cultures were then transferred to 50 mL Falcon tubes, and to wash away bacteria, colonies were harvested at 3,000xg for 15 minutes at room temperature and washed three times with 45 mL of 1X ASW. After the final wash, the remaining pellet was snap frozen in liquid nitrogen and sent to Dovetail Genomics (Scotts Valley, California, United States) for genomic DNA extraction and sequencing, which were performed by Dovetail Genomics staff according to the protocols described in the following sections.

Genomic DNA Extraction: CTAB (Hexadecyltrimethylammonium bromide) buffer was prepared as follows (recipe for 1 L): 280 mL of 5M NaCl, 100 mL of 1M Tris-HCl (pH 9.5), 40 mL of 0.5 M EDTA, 20 g of CTAB, 10 g of PVP-40, 10 g of PEG 8000, and molecular biology-grade water added to a final volume of 1 L. The resulting solution was heated to 65°C to dissolve all components.

The CTAB buffer was pre-warmed in a water bath at 68°C. Approximately 10 mL of CTAB buffer was prepared per gram of cell pellet. A small fraction of the cell pellet was reserved for Omni-C sequencing (see following section). The remaining pellet was ground into a fine powder using a pre-chilled mortar and pestle in liquid nitrogen. The resulting powder was then suspended in warm CTAB buffer supplemented with 0.5% β -mercaptoethanol and incubated at 68°C for 15 minutes

with occasional agitation. 1 μ L of RNase A (100 mg/ μ L) and 10 μ L of Proteinase K (20 mg/mL) were added, and the solution was incubated at 60°C for an additional 30 minutes.

For DNA extraction, an equal volume of phenol:chloroform was added to the suspended ground cell lysate, and the solution was gently mixed by manual inversion, ensuring that the phases were completely mixed. The resulting solution was then centrifuged at 5,200xg for 10 minutes at room temperature, and the aqueous phase was then transferred into a new tube. The phenol:chloroform extraction was repeated until no debris or interphase remained. An equal volume of chloroform: isoamyl alcohol was added and mixed by manual inversion. The solution was again centrifuged at 5,200xg for 10 minutes at room temperature using a tabletop centrifuge, and the resulting aqueous phase was transferred to a new tube.

Next, 0.7 volumes of isopropanol were added to the aqueous phase to precipitate DNA, and the tube was inverted several times to mix the solutions. The resulting DNA was spooled onto a glass rod and transferred into a tube containing 9.5 mL of G2 buffer pre-mixed with 19 μ L of RNase A (100 mg/ μ L). The solution was vortexed briefly, followed by the addition of 200 μ L of Qiagen Protease, and vortexed again to homogenise. If no visible DNA could be spooled, the sample was centrifuged for 20 minutes at 4°C to pellet DNA. The supernatant was discarded, and the pellet was washed with approximately 10 mL of ice-cold 70% ethanol, followed by centrifugation and removal of the supernatant. Residual ethanol was washed off, and the DNA pellet was then resuspended in 9.5 mL of G2 buffer pre-mixed with 19 μ L of RNase A (100 mg/ μ L). The solution was then vortexed, 200 μ L of Qiagen Protease were added, and the sample was vortexed again. The resulting samples were incubated again in G2 at 50°C for 30-60 minutes, with occasional gentle agitation.

For each tube used for DNA extraction, DNA concentration was quantified using a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA) with 10 μ L of solution. Based on total DNA yield, different Qiagen Genomic-tip protocols were selected: the Mini tip protocol was used for samples with less than 25 μ g; a Midi tip protocol was used for samples with more than 25 μ g of DNA; and a Maxi column was used if the total DNA was estimated to be 100 μ g or more. Before loading the Qiagen column, sample were vortexed for at least 10 seconds to ensure homogeneity. Highly viscous samples were further diluted using G2 buffer as needed.

Genomic DNA Library Preparation and Sequencing: Genomic DNA libraries were prepared by Dovetail Genomics/Cantata Bio for Omni-C sequencing following the protocol described in¹². Briefly, chromatin was crosslinked by adding formaldehyde to the cell pellet, after which crosslinked chromatin was extracted and subsequently digested with DNase I. Chromatin ends were repaired and ligated to a biotinylated bridge adapter, followed by proximity ligation of adapter-containing ends. After that, crosslinks were reversed, and the DNA was purified. Non-ligated or external biotin residues were removed, and sequencing libraries were generated using NEBNext Ultra enzymes and Illumina-compatible adapters. Biotinylated DNA fragments were isolated using streptavidin beads, followed by PCR enrichment of each library. The resulting libraries were sequenced on an Illumina HiSeqX platform, generating approximately 30x sequence coverage. High-quality reads (MQ>50) were used for scaffolding with the HiRise pipeline¹².

gDNA samples were quantified using a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). A PacBio SMRTbell library (~20kb) for PacBio Sequel was constructed using SMRTbell Express Template Prep Kit 2.0 (PacBio, Menlo Park, CA, USA) following the manufacturer's recommended protocol. The library was bound to polymerase using the Sequel II Binding Kit 2.0 (PacBio) and loaded onto a PacBio Sequel II. Sequencing was performed on PacBio Sequel II 8M SMRT cells.

Genome Assembly: HiFi and Omni-C reads were assembled with hifiasm v0.16.1-r375¹³ using default parameters with Hi-C integration, producing a primary contig assembly. The resulting contig assembly was used as a reference for aligning the Omni-C reads. Briefly, reads were aligned using BWA-MEM v0.7.10¹⁴ with the -5SP and -t 8 options specified, and all other options set as default. SAMBLASTER v0.1.22¹⁵ was used to flag PCR duplicates, which were subsequently excluded from downstream analysis. Alignments were then filtered with Samtools v0.1.6¹⁶ using the -F 2304 filtering flag to remove non-primary and secondary alignments.

To improve genome assembly quality, we filtered the Hi-C contact data to remove low-contact read pairs based on statistical analysis of Hi-C contact scores. In brief, we generated a link_counts.tsv file containing read pair identifiers and Hi-C contact scores. We then generated histograms to assess the range and distribution of contact scores. We identified a suitable cutoff for low-contact scores using percentile analysis, selecting the 70th percentile (contact score of 376) based on data distribution. Next, we applied the determined cutoff to the link_counts.tsv

data, creating a filtered list of read pair identifiers in filtered_link_counts.tsv. A custom python script was used to extract the filtered read pair identifiers from filtered_link_counts.tsv and parse and filter the corresponding FASTA file, retaining only sequences associated to the filtered identifiers. The resulting filtered_output.fasta file represents the target choanoflagellate genome of interest, which was subsequently used for scaffolding.

Hi-C reads were aligned to filtered_output.fasta following the Phase Genomics Proximo Hi-C Kit recommendations¹⁷. Briefly, reads were aligned using BWA-MEM¹⁴ with the -5SP and -t 8 options specified, and all other options set to default. SAMBLASTER¹⁵ was used to flag PCR duplicates, which were later excluded from downstream analysis. Alignments were then filtered with Samtools¹⁶ using the -F 2304 filtering flag to remove non-primary and secondary alignments.

Chromosome-scale scaffolds were generated using Phase Genomics' Proximo Hi-C genome scaffolding platform, following the same single-phase scaffolding procedure described in¹⁸. Similar to the LACHESIS method¹⁹, this approach computes a contact frequency matrix from the aligned Hi-C read pairs, normalised by the number of restriction sites each contig, and constructs scaffolds by optimizing the expected contact frequency and other statistical patterns in Hi-C data. Approximately 20,000 independent Proximo runs were performed on the assembly to optimise the number of scaffolds and scaffold construction, maximizing concordance with the observed Hi-C data. Manual inspection and correction of scaffolding errors were carried out using Juicebox v1.0^{20,21}.

Genome Annotation: The genome of *C. flexa* was annotated using an integrated pipeline combining the results from several gene-calling algorithms with clues from transcriptomic data and protein evidence from closely related species. To generate transcriptomic evidence of gene models, publicly available *C. flexa* transcriptomic reads (sequence read archive accession: SRR900977¹) were mapped to the genome assembly using hisat2 v2.2.1²². Prior to mapping, reads were trimmed using Trimmomatic v0.39²³ using the arguments LEADING:3 TRAILING:3 MINLEN:36, and Illumina adapters were also removed. The resulting file was converted to sorted BAM format using Samtools, v1.10¹⁶ and subsequently used to provide transcriptomic evidence for gene annotation. We also used the publicly available *C. flexa* transcriptome for additional independent evidence. Transcript evidence was aligned to the genome using minimap2 2.26-r1175²⁴, using the cs argument, splice site recognition from both strands, and the -G parameter set to 5,000 following developer recommendation. This analysis resulted in 70,882 alignments.

We used all publicly available choanoflagellates proteomes for protein evidence from closely related species. Two proteomes – from the species *Monosiga brevicollis*²⁵ and *Salpingoeca rosetta*³ – were derived from genomic data, while 19 additional proteomes were translated from transcriptome assemblies of diverse choanoflagellates²⁶. In total, 688,088 protein sequences were used as homology-based evidence for gene annotation. Protein sequences were aligned to the *C. flexa* genome using DIAMOND v2.1.8²⁷ and exonerate v2.4.0²⁸. DIAMOND identified 269,722 putative alignments and exonerate identified 2,679.

Next, three gene prediction softwares were used to identify 45,273 putative gene models. First, Augustus v3.3.2²⁹ was run using *Toxoplasma* parameters, resulting in 6,822 high-quality predictions (>90% exon evidence) and 9,196 gene models without quality thresholding (15,676 total Augustus gene predictions). Second, SNAP³⁰, was trained using 194 eukaryotic BUSCO genes v2.0³¹ identified in the genome, yielding 15,083 gene models. Third, GlimmerHMM³², trained on the same set of eukaryotic BUSCO genes, identified 14,172 gene models.

All putative gene models were combined using a weighted consensus approach with the EvidenceModeler software³³. Proteins, transcripts, and each of the gene calling algorithms were given the same weight, except high-quality annotations from Augustus, which were given twice the weight. This resulting set of 16,832 gene models was further filtered to remove sequences shorter than 50 amino acids in length, repetitive elements like transposons or spanned gaps. This filtering resulted in the removal of 186 gene models, and 16,646 total gene models remained. Workflow orchestration was performed using funannotate v1.8.16³⁴. The completeness of the gene models were assessed using BUSCO v4.0.4, with the Eukaryota database of near-universally single-copy orthologous genes from OrthoDB v10^{35,36}. This analysis revealed the genome annotation is 83.6% complete, a substantial improvement over the 72.6% completeness estimated from the genome alone. Lastly, tRNAs were identified in the genome assembly using tRNAscan-SE v2.0.9³⁷, resulting in 237 tRNA models.

A final decontamination step was performed to remove putative bacterial contaminants stemming from *Halopseudomonas oceanii* (NCBI accession: GCF_963677335.1), the bacterial food source used in *C. flexa* cultures. To identify contaminant sequences, a reciprocal best BLAST hit analysis was conducted using BLAST v2.3.0+³⁸, with an expectation value threshold of $1e^{-3}$. This analysis identified 2,562 reciprocal best BLAST hits with at least 90% identity to the bacterial co-culture.

Of these, 2,518 hits mapped to a single scaffold (scaffold_530), representing 93.78% (2,518/2,685) of all gene models on that scaffold. Similarly, two additional scaffolds (scaffold_41 and scaffold_87) had numerous bacterial hits, with 36.36% (12/33) and 91.43% (32/35) of gene models showing *H. oceani*-like similarity, respectively. While these may represent putative horizontal gene transfer events, contamination was considered the most parsimonious explanation; therefore, these scaffolds were removed from the final assembly.

The final *C. flexa* genome assembly spanned 56,404,751 base pairs (bp) across 528 scaffolds, with a GC content of 50.83%. The assembly N50 and L50 values were 1,302,044 bp and 16 scaffolds, respectively, and the largest scaffold measured 3,795,244 bp, indicating high assembly contiguity. The final genome annotation encoded 14,084 genes with an overall BUSCO completeness of 82.8%. Together, these results underscore that our workflow successfully produced a high-quality genome assembly and annotation. The annotated *C. flexa* reference genome was deposited on Zenodo under doi: 10.5281/zenodo.13837466 and on NCBI (with associated raw reads) under BioProject ID PRJNA1381315.

Whole-genome sequencing of *C. flexa* strains isolated in the field (Figures 6 and S9, S11)

Genomic DNA extraction and sequencing of Strains 1-3 (SSB and clonal lines): Approximately 50 mL of exponentially growing cultures from Strain 1 (ChoPs7, clones 1-3), Strain 2 (M44B, clones 1-3) and Strain 3 (M60B, clones 1-3) grown in SWC medium were harvested, concentrated down to 2 mL and incubated 4 hours at room temperature. After incubation, cells were pelleted by centrifugation at 3,300xg for 15 min at room temperature. A total of 5×10^6 cells were further processed for gDNA extraction using either a lysis buffer coupled with ethanol and sodium acetate precipitation (see detailed protocol below; for Strain 1 clone 1A1; Strain 2 clone 1C5; and Strain 3 clone 3D7), or the Blood & Cell Culture DNA Mini Kit (#13323, Qiagen) (see detailed protocol below; remaining clones and single-sheet-bottleneck cultures).

Lysis buffer and ethanol and sodium acetate precipitation: Approximately 1×10^7 cells were harvested by centrifuging ~30-50 mL of a dense culture for 15 minutes at 3,300xg at 4°C, resuspended in 1 mL of 1X ASW, and transferred to a 2 mL microcentrifuge tube. Cells were pelleted again at 17,000xg for 5 minutes, and the pellet was finally resuspended in 100 µL of lysis buffer (20 mM Tris HCl pH 8, 150 mM KCl, 5 mM MgCl₂, 250 mM Sucrose, 1 mM DTT, 10 mM Digitonin, 1 mg/mL Sodium heparin, 1 mM Pefabloc, 100 µg/mL Cycloheximide). Samples were

incubated on ice for 10 minutes and then passed 10 times through a 26G needle (Fisher Scientific 15301557) to facilitate cell lysis. Lysates were centrifuged at 6,000xg for 10 minutes at 4°C, and the pellet was discarded. gDNA was precipitated from the supernatant by adding 0.1 volumes of 3 M sodium acetate, mixing by pipetting up and down, 2.5 volumes of cold (–20°C) ethanol, mixing by vortexing, and incubated overnight at –20°C to allow gDNA precipitation. Samples were then centrifuged for 30 minutes at 16,000xg (maximal speed) at 4°C, washed twice with ice-cold 70% ethanol (in Milli-Q water), and centrifuged again for 15 minutes at 16,000xg (maximal speed) at 4°C. After carefully removing the supernatant, the gDNA pellet was air-dried for 5 minutes at 56°C and resuspended in 30 µL of 5 mM Tris-HCl (pH 8) ('NE buffer' from Macherey gel purification kit, reference 740609.50). DNA concentration and purity were measured using a NanoDrop spectrophotometer.

Qiagen Blood & Cell Culture DNA Mini Kit extraction: 50 mL of dense cell cultures ($>1 \times 10^6$ cells/mL) were concentrated to 2 mL in 1% SWC medium by centrifugation at 3,300xg for 10 minutes, and incubated overnight to graze residual bacteria. Samples were then centrifuged again for 10 minutes at 3,300xg, the supernatant was discarded, and the cell pellet was stored at –20°C until gDNA extraction. gDNA was then purified using the Qiagen Blood & Cell Culture DNA Mini Kit according to the manufacturer's protocol. The final gDNA pellet was resuspended in 30 µL of 5 mM Tris-HCl (pH 8) ('NE buffer' from Macherey NucleoSpin® Gel and PCR clean-up kit, #740609.50). DNA concentration and purity were measured using a NanoDrop spectrophotometer.

All gDNA samples were shipped to Eurofins Genomics for INVIEWS Resequencing (10M paired-end reads, Illumina 150 bp sequencing). Raw short reads for Strains 1, 2 and 3 were deposited on Zenodo under doi: 10.5281/zenodo.13837614 and on the NCBI Sequence Read Archive (BioProject ID PRJNA1381491).

Short reads assembly and analysis: Paired-end reads were quality assessed and trimmed using FASTP (v0.20.1)³⁹. The first and the last base of each read were removed, and adaptors were trimmed using the '--detect_adapter_for_pe' option. The quality threshold was set to 30, and all other parameters were kept as default values. Qualified reads were aligned to the *C. flexa* reference genome using BWA-MEM v0.7.17⁴⁰. The mapped reads were converted to bam format and sorted using Samtools v1.18¹⁶. Duplicate reads were marked using GATK v4.1.9.0⁴¹, and BAM files were indexed using Samtools v1.18. Variant calling was performed using GATK

HaploTypeCaller, applying different ploidy assumptions (ploidy = 1, 2, and 4) to detect potential polymorphisms across samples. Resulting variants were jointly genotyped for each ploidy condition using the GATK GenotypeGVCF function⁴².

Phylogenomic tree construction (Figure 6c; Supplementary Files S5-S6)

We realised that, although Strains 2 and 3 seemed to be haploid, Strain 1 (ChoPs7) exhibited a diploid-like pattern based on the distribution of allelic frequencies in scaffold_1, as assessed using ploidyNGS⁴³. To accommodate samples with potentially differing ploidy levels, we used SNPs called under a diploid assumption and homozygous across all samples for downstream analysis. Variants were filtered using BCFtools in Samtools v1.18⁴⁴ with the following criteria: quality score > 30, filtered read depth > 4, variant type = 'SNP', minimum and maximum allowed alleles = 2, and homozygous genotypes across all samples. The resulting VCF files were converted to PHYLIP format as input for IQ-TREE using the vcfR package⁴⁵ in R (v4.1.1). A phylogenomic tree of all samples was constructed based on the identified SNPs using IQ-TREE (v2.3.2) with the General Time Reversible (GTR) substitution model, gamma-distributed rate variation, and ascertainment bias correction^{46,47}. The output tree structure was visualised using iTOL (v6.9.1)⁴⁸.

Clonality analysis: To test if single-sheet bottlenecked samples contained genetically diverse backgrounds, we quantified the degree of clonality between single-sheet bottlenecked samples and their corresponding single-cell bottlenecked derivatives. We reasoned that if single-sheet bottlenecked samples were less clonal than single-cell bottlenecked samples, they would exhibit a greater number of variants unique to single-sheet bottlenecked samples. To maximize the detection of minor alleles, we used variants called with ploidy = 4 assumption for all samples. Identified variants were filtered using GATK (v4.1.9.0) with the following parameters as recommended by GATK team of the Broad Institute (<https://gatk.broadinstitute.org/hc/en-us/articles/360037499012-I-am-unable-to-use-VQSR-recalibration-to-filter-variants>):

- for SNPs: QD < 2.0, MQ < 40.0, FS > 60.0, SOR > 3.0, MQRankSum < -12.5, ReadPosRankSum < -8.0.
- for INDELs: QD < 2.0, ReadPosRankSum < -20.0, FS > 200.0, SOR > 10.0.

The number and ratio of unique alleles (either SNPs or INDELs) for all pairwise sample combinations were computed using the vcfR package for vcf file manipulation⁴⁹, custom python (v3.10.15) scripts for data manipulation, and the tidyverse package (v2.0.0) for visualization⁵⁰ in R (v4.1.1)⁵¹ and RStudio (v2021.9.0.351)⁵².

Analysis of polymorphic sites with putative positive selection (Figures 6h-j and S9, S11)

Coding sequences for each strain were inferred from the variant data using vcf2fasta (<https://github.com/yeeus/vcf2fasta>). Inferred coding sequences were programmatically checked across all reference sequences, and genes with incorrect inferences were removed from the analysis. Variants were jointly called for all strains using the GATK (v4.1.9.0) and filtered following the recommended parameters from GATK team described in the clonality analysis section above. The inferred coding sequences of each gene were checked for length integrity, retaining only those that were multiple of 3, of equal length across strains, and containing no gaps. Genes that did not satisfy those criteria were aligned using Clustal-omega (v1.2.4), and those genes that introduced gaps that not in multiples of 3 (causing frameshifts) were removed from downstream analyses. To compare predicted protein sequences across strains, we computed the number of non-synonymous substitutions across the strains using Biopython (v1.85) (Cock et al. 2009). The Ka/Ks ratio (*i.e.*, dN/dS ratio) was calculated between strains to identify genes under putative positive selection. Since most sites of coding sequences are typically under negative selection, and only a small fraction is often under positive selection, gene-wise Ka/Ks ratio could average out potential regions with positive selection (Wang et al. 2010). Therefore, we additionally computed sliding-window Ka/Ks ratios to capture localised signals of selection for subregions of each sequence. We split sequences using functions in KaKs_Calculator 2.0 with a window size = 114 bp and step size = 6 bp, and computed Ka/Ks ratio using KaKs_Calculator 3.0 with MYN method (Wang et al. 2010, Zhang 2022). We defined regions with Ka/Ks ratio > 2 for at least 30 base pair as 'high Ka/Ks region.'

To explore functional enrichment, we computed InterPro signatures overlapping high Ka/Ks regions. InterPro signatures of the *C. fllexa* predicted proteome were obtained using InterProScan (v5.50-84.0)^{53,54}. The InterPro signature enrichment was performed using Fisher's exact test in the base R stats package, comparing the frequency of each InterPro signature within versus outside high Ka/Ks regions. Enrichment analysis results were visualised using tidyverse (v2.0.0)⁵⁰.

Kin recognition experiments between *C. fllexa* Strains 1-3 (Figures 6d-g and S10; Extended Data Table 1)

Approximately 40 mL of ChoPs7, M44B (clones 1C5 or 1C7) and M60B (clone 3D7) cells from exponentially growing cultures grown in SWC medium were harvested, washed, and stained with

CellTrace CFSE (green) and CellTrace Far Red (magenta) as before (see section ‘Live imaging of dual labelling of aggregates’).

To assess kin recognition between *C. flexa* strains, green- and magenta-labelled single cell populations from each strain were mixed in a 1:1 ratio. A final number of 1×10^4 cells (5×10^3 cells of each colour) was seeded in 1 mL of SWC medium in a 24-well plate and incubated overnight in a 25°C incubator (protected from light by a layer of aluminium foil). After incubation, 100 µL of each sample were transferred using a cut P200 tip into a black 96-well ibiTreat µ-plate. The wells were scanned and all relaxed colonies were imaged in at least 10 different positions by DIC microscopy and 4% intensity LEDs (488 nm and 650 nm) with a C-Apochromat 40X/1.2 W Corr M27 Zeiss objective on a Zeiss Axio Observer Z.1 inverted microscope. All experiments were performed in four independent replicates (two biological replicates, each including two technical replicates, encompassing all strains and colour permutations per condition). A total of n=530 sheets were quantified.

Quantification of the segregation index: Images were analysed with Fiji Imaging Software version 2.9.0/1.53t. In brief, the relative proportion of green and magenta cell populations were quantified by manually counting the number of cells of each colour relative to the total number of cells within each colony. A *segregation index* (*s index*) between every pairwise strain combination was defined following⁵⁵, treating each colony as an individual ‘local environment’, with the following modifications:

1. Local segregation for a colony

For every colony i , we quantified the number of cells for each strain:

- $n_{A,i}$ cells of ‘Strain A’
- $n_{B,i}$ cells of ‘Strain B’
- total $N_i = n_{A,i} + n_{B,i}$

The segregation of ‘Strain A’ cells in that colony (considering self-exclusion) is:

$$seg_A(i) = \frac{n_{A,i} - 1}{N_i - 1} \text{ (if } n_{A,i} > 0 ; N_i > 1)$$

and for ‘Strain B’:

$$seg_B(i) = \frac{n_{B,i} - 1}{N_i - 1} \text{ (if } n_{B,i} > 0 ; N_i > 1)$$

2. Average segregation between colonies (per replicate)

Next, to calculate the overall seg_A and seg_B values, we averaged across all the cells of each strain, weighting by how many cells of that strain are in each colony:

$$seg_A = \frac{1}{N_A} \sum_i n_{A,i} * seg_A(i)$$

$$seg_B = \frac{1}{N_B} \sum_i n_{B,i} * seg_B(i)$$

Where:

$$N_A = \sum_i n_{A,i} ; \text{ and } N_B = \sum_i n_{B,i}$$

3. Normalise by global frequencies

Finally, we normalised by global strain proportions:

$$P_A = \frac{N_A}{N_A + N_B} ; \text{ and } P_B = \frac{N_B}{N_A + N_B}$$

And calculated the final *segregation index* (s) as:

$$S_A = \frac{seg_A - P_A}{1 - P_A} ; \text{ and } S_B = \frac{seg_B - P_B}{1 - P_B}$$

Statistical analyses

The significance of differences in pairwise comparisons was tested using the non-parametric Mann-Whitney U test. Shapiro-Wilk normality test and F-test were used to evaluate data normality and the differences in variances between conditions, respectively. All statistical analyses were performed in R using the base *stats* package version 3.6.3⁵¹.

Supplementary Figures

Figure S1

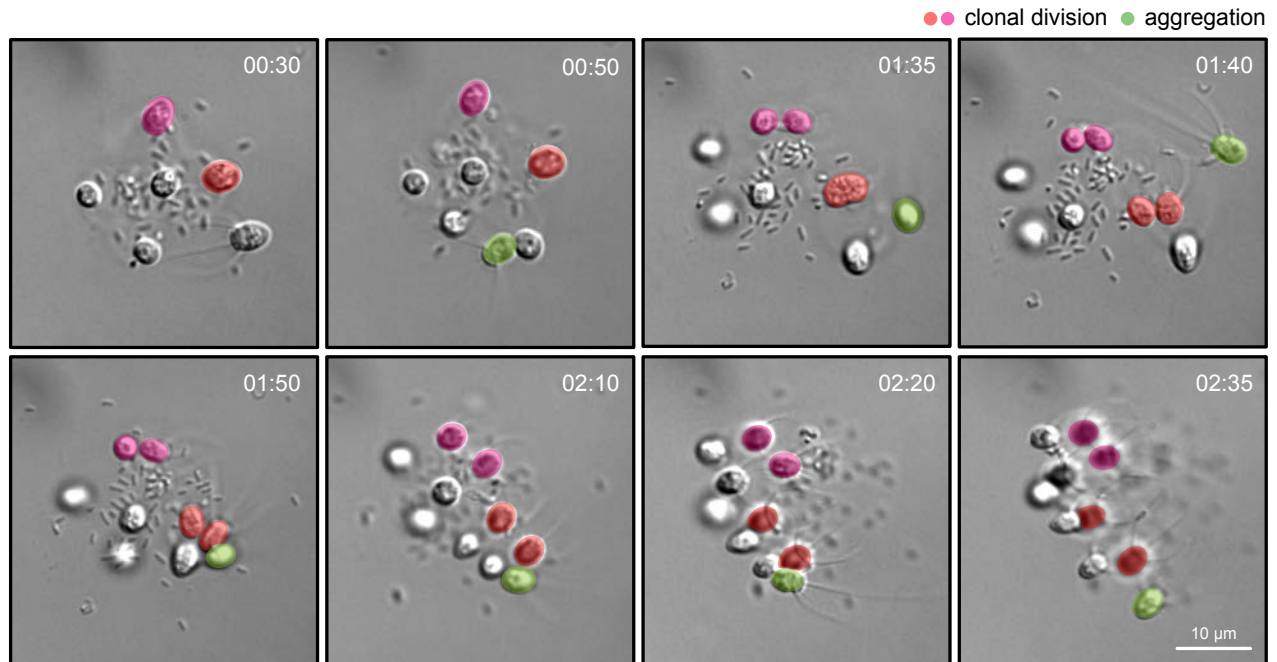
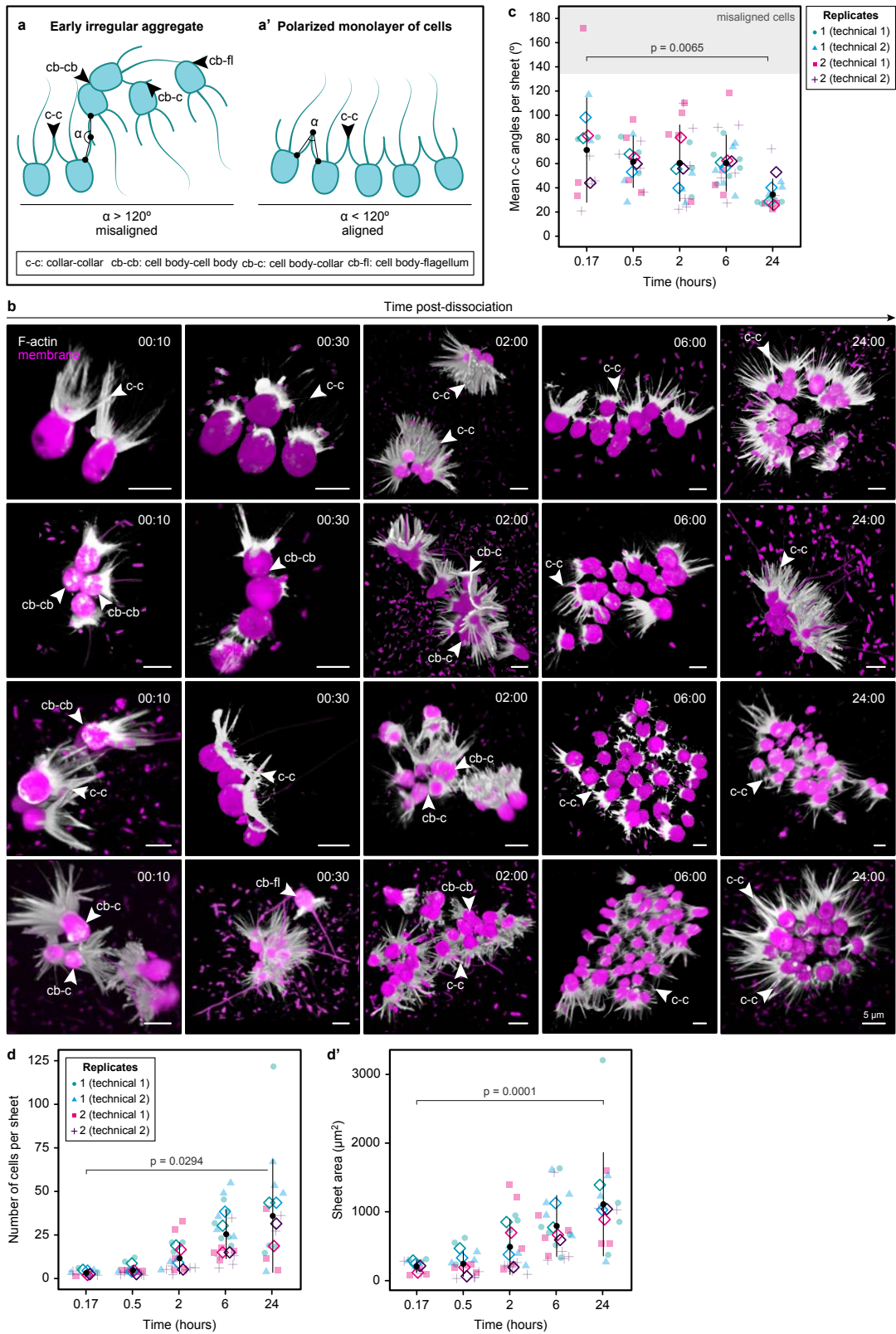


Figure S1. *C. flexa* sheets can form clonally but display aggregative features. Stills of a brightfield timelapse movie of a small-sized *C. flexa* colony expanding in cell number. Two cells within the sheet divide clonally (orange and pink pseudocolour), and a swimmer flagellate cell joins the colony by cellular aggregation (green pseudocolour). N=2 biological replicates, two technical replicates each (n=4). Figure related to **Figure 1g** and **Video S3**.

1574 **Figure S2**
1575



1576 (Figure S2 legend on the next page)

Figure S2. Aggregation is a multi-step process that allows fast establishment of multicellularity. (a) Schematics of the morphological metrics quantified from the cells within sheets in **Figures 2e** and **S2c**: adhesion angle between collar-collar contacts of neighbouring cells (α) and proportion of cells with aligned apico-basal polarity. Note the different kinds of contacts between cells (black arrowheads, see legend). (b) 3D reconstructions of Airyscan confocal z-stacks of dissociated cells fixed 10 minutes, 30 minutes, 2 hours, 6 hours, and 24 hours post-dissociation (hpd), with a membrane staining labelling the cell body and flagella (FM 4-64FX, magenta) and F-actin staining labelling the microvilli (Phalloidin 488, white). Shown are four different representative colonies for each timepoint. Note that cells within sheets frequently show unaligned apico-basal polarity and diverse orientations in early aggregation timepoints, and reorient to form polarised monolayers of cells with aligned collar-collar contacts at later timepoints (white arrowheads, see legend in a). Time scale hh:mm. (c) Quantification of mean collar-collar angles per sheet during aggregation in **Figures 2e** and **S2c**. Grey area corresponds to angles between misaligned cells ($\alpha > 120^\circ$). (d) Quantification of cell number per sheet and (d') sheet area during aggregation in **Figures 2e** and **S2c**. Black circles: mean; error bars: standard deviation; diamonds: mean values of each independent replicate. N=2 biological replicates, two technical replicates each (n=75 sheets). Statistics in c-d (comparing t=10 minutes and t=24 hours timepoints) by the Mann-Whitney U test (p-value). Figure related to **Figure 2e-g** and **Videos S7-S11**.

Figure S3

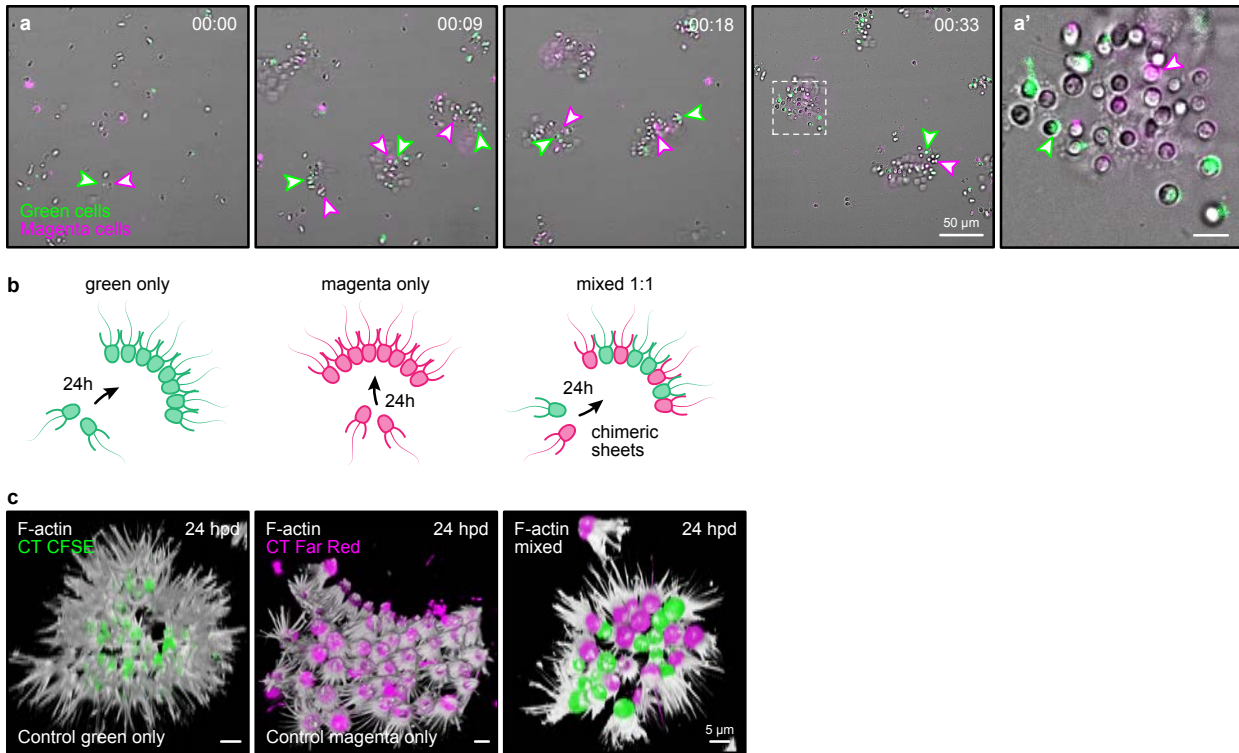


Figure S3. *C. flexa* sheets can form by aggregation. (a) Stills from a timelapse movie showing that two dissociated single cell populations labelled with either CellTrace CFSE (green) or CellTrace Far Red (magenta) aggregate into dual-labelled chimeric sheets with cells of both colours (green and magenta arrowheads). Time scale hh:mm. (a') Zoom-in of dashed square in a. Scale bar: 10 μm. N=3 independent biological replicates. (b) Schematics of the experimental design of sheet formation by aggregation of two dissociated single cell populations and unmixed, single-labelled controls for 24 hours. (c) 3D reconstructions of Airyscan confocal z-stacks of sheets formed from dissociated single cell populations labelled with green or magenta dyes as before and fixed 24 hours post-dissociation (hpd), with additional filamentous actin (F-actin) staining labelling the microvilli (Phalloidin 405, white). When both single cell populations are mixed in a 1:1 ratio they form dual-labelled chimeric colonies (mixed). Shown is one representative colony per condition. N=2 biological replicates, two technical replicates each (n=16 sheets). Figure related to **Figure 2h** and **Videos S12-S13**.

Figure S4

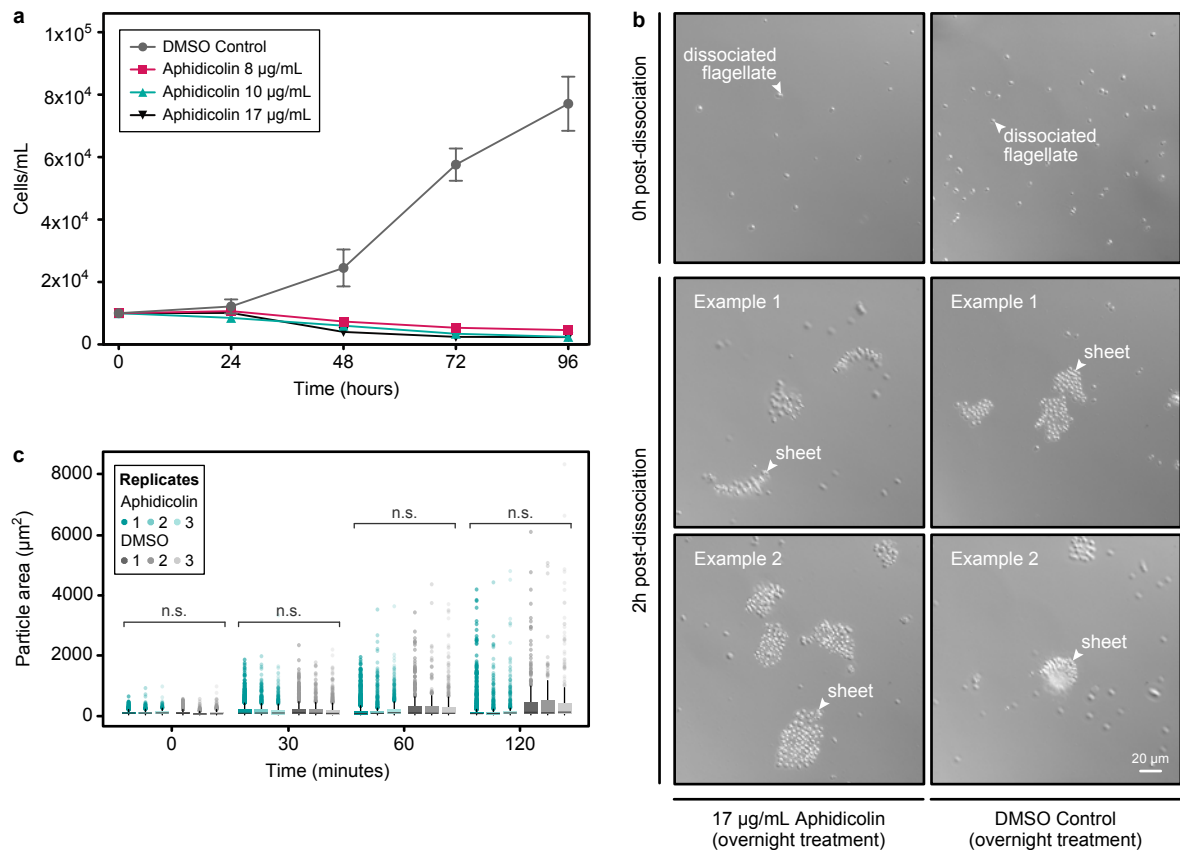
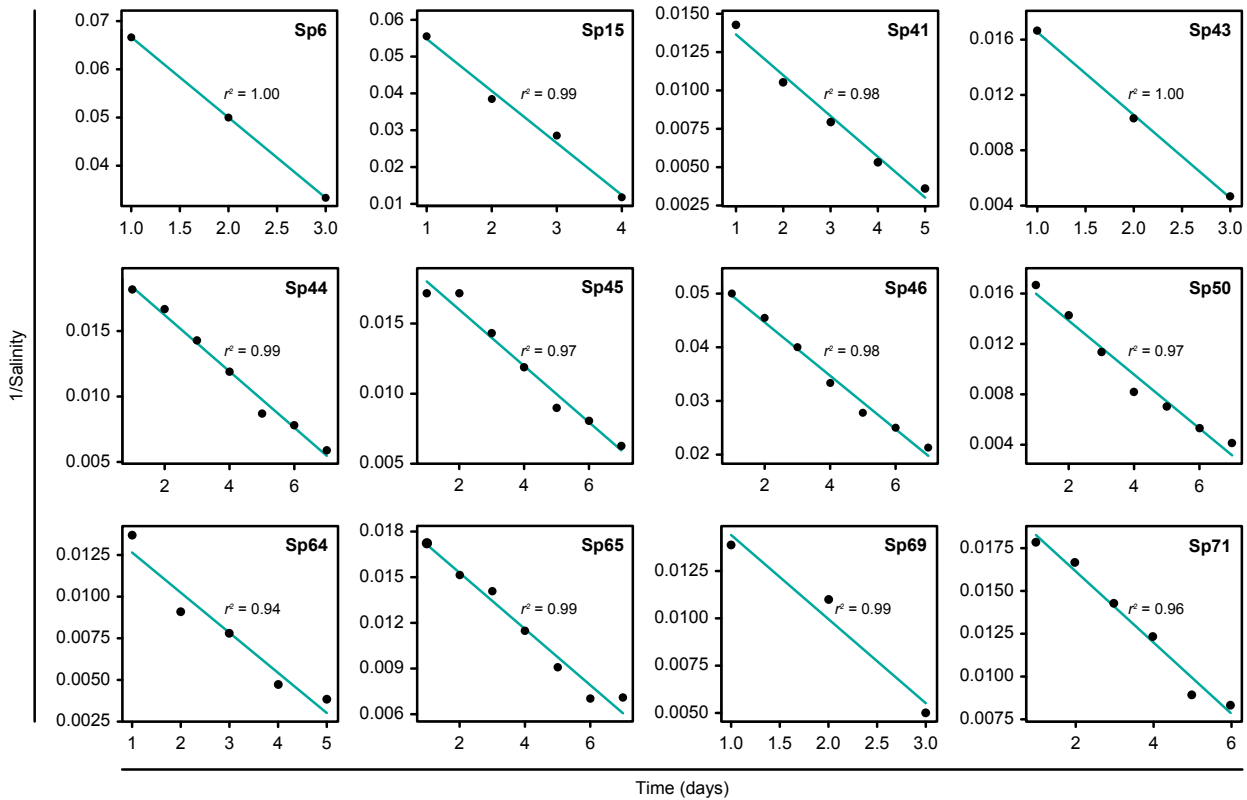


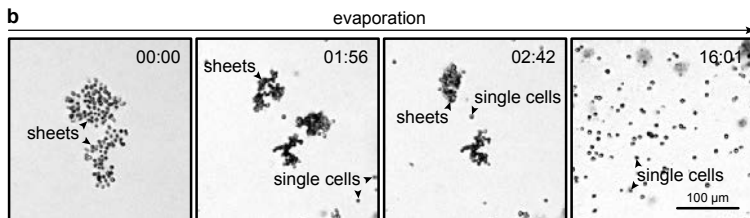
Figure S4. Aggregation is independent of cell division. (a) Dose-response growth curve of dissociated single cells treated with distinct concentrations of Aphidicolin shows that it effectively blocks cell division at concentrations equal or greater than 8 µg/mL. An equivalent volume of DMSO (dimethyl sulfoxide) was used as a negative control. N=3 independent biological replicates, including three technical replicates per condition (n=9). (b) Snapshots of cells after 0 hours (flagellate cells, white arrowheads) and 2 hours post-dissociation (sheets, white arrowheads). Cells were treated overnight with 17 µg/mL Aphidicolin cell cycle inhibitor prior dissociation. (c) Quantification of particle area during an aggregation time course of dissociated single cells pre-treated with 17 µg/mL aphidicolin overnight or with the equivalent volume of DMSO. N=3 independent biological replicates, with three technical replicates each (n=9). Statistics in c comparing average area of Aphidicolin-treated versus DMSO control at each timepoint by the Mann-Whitney U test (p-value; n.s., non-significant). Figure related to **Figure 2i**.

Figure S5

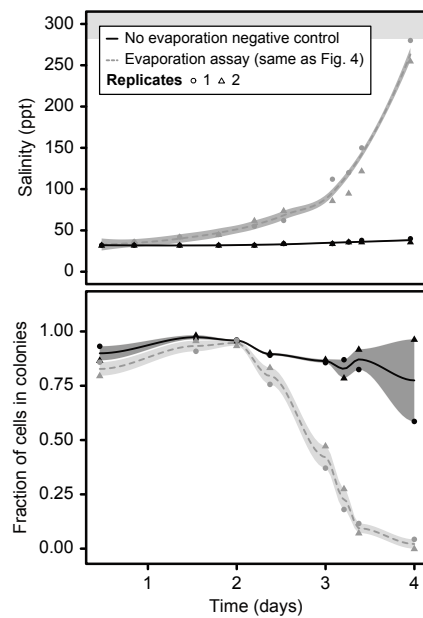
a



b



c



(Figure S5 legend on the next page)

Figure S5. Modelling splash pool gradual evaporation enables design of a laboratory experimental setup with similar evaporation dynamics. (a) Inverse of salinity over time of followed-up splash pools during an 8-day time course from Exped-A expedition (n=12 splash pools). Splash pools during the first 6 days of the time course gradually evaporated, and hence they were used to model splash pool evaporation dynamics. Splash pools with less than 3 points outside saturation range were not included in the modelling (*i.e.*, Sp12, Sp42, and Sp70; see **Extended Data Figure 3**). Panel related to **Figure 4a-b** and **Table S1**. (b) Snapshots of a brightfield timelapse movie of *C. flexa* colonies in 1X ASW salinity experiencing gradual evaporation shows sheet dissociation into single cells (black arrowheads). (c) (Upper panel) Quantification of salinity during a 4-day time course comparing the gradual evaporation laboratory experimental set up and a no evaporation negative control laboratory set up. Values are represented as mean (lines) $\pm$ s.d. (shadowed area). Grey area depicts salinity saturation outside the refractometer measuring range. N=11 independent replicates. Shown is quantification of two independent biological replicates. (Lower panel) Quantification of fraction of cells in colonies. Panels b-c related to **Figure 4c-d** and **Video S17**.

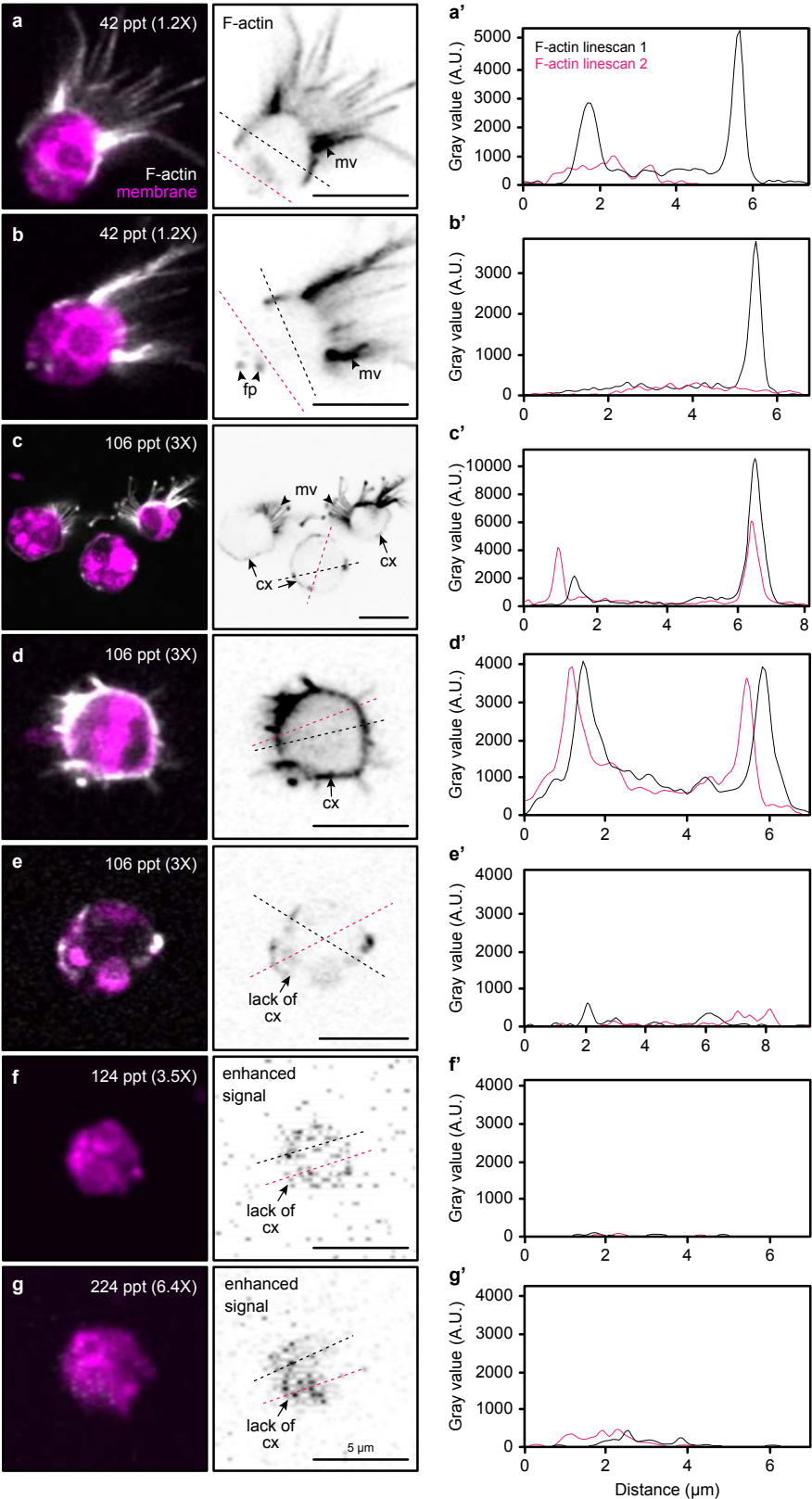


Figure S6. Hypersalinity induces encystation of dissociated flagellates and a temporary actin cortex
. (a-g) Mid z-sections of Airyscan confocal z-stacks of *C. flexa* cells during gradual evaporation fixed and stained with a membrane (FM 4-64FX, magenta) and F-actin (Phalloidin 488, white) dyes confirmed that cysts lack a collar and a flagellum at later stages of evaporation. Flagellate cells (a-b, low-evaporation control) exhibit the stereotypical filamentous actin fibres distribution in the collar microvilli (mv, black arrowhead), as described in ¹. Gradual evaporation triggers a morphological change from flagellates to cysts, showing a temporary actin cortex (cx, black arrow) during early evaporation (c-d) and later actin-free cysts closer to saturation (e-g). (a'-g') Linescans of F-actin fluorescence intensity along two dashed lines of interest in a-g, showing cortical actin as two peaks where the lines intersect the cell cortex. N=3 independent biological replicates, with two technical replicates per condition (n=12 cells). Figure related to **Figure 4j-l**.

Figure S7

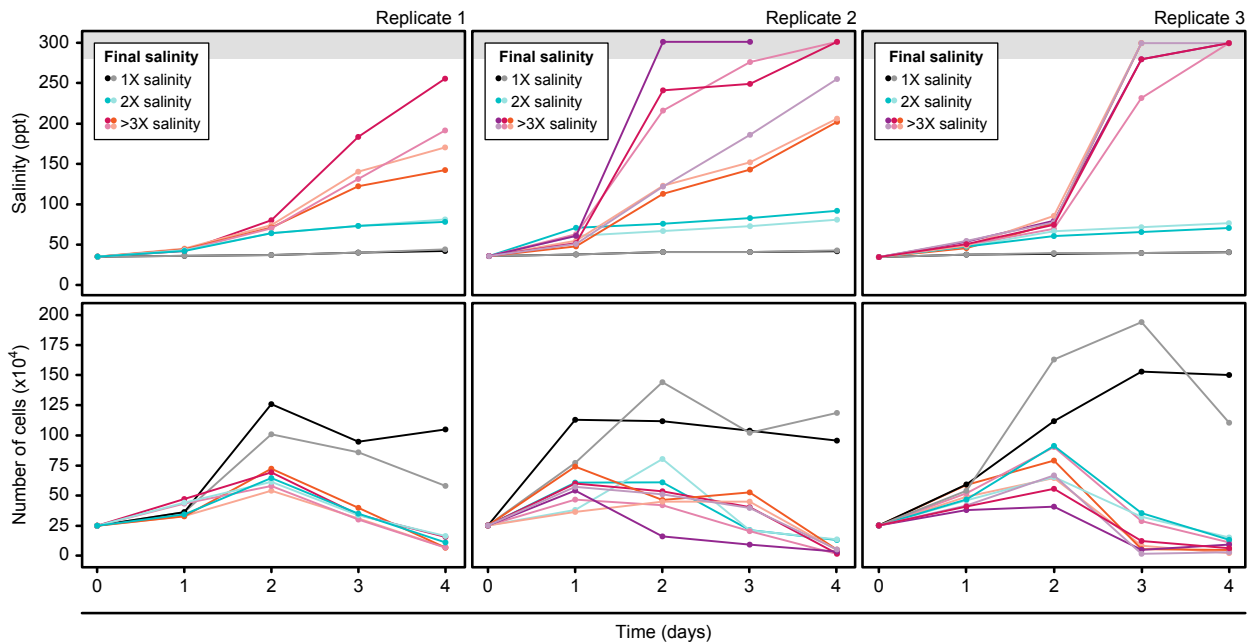


Figure S7. Salinity increase during gradual evaporation arrests cell growth. Quantification of salinity (upper panels) and cell number (lower panels) over a 5-day time course experiment in the gradual evaporation experimental set up. Grey area depicts salinity saturation outside the refractometer measuring range. 1X salinity corresponds to 35 ppt. N=3 independent biological replicates, with two technical replicates per condition (n=6). Figure related to **Figure 4m**.

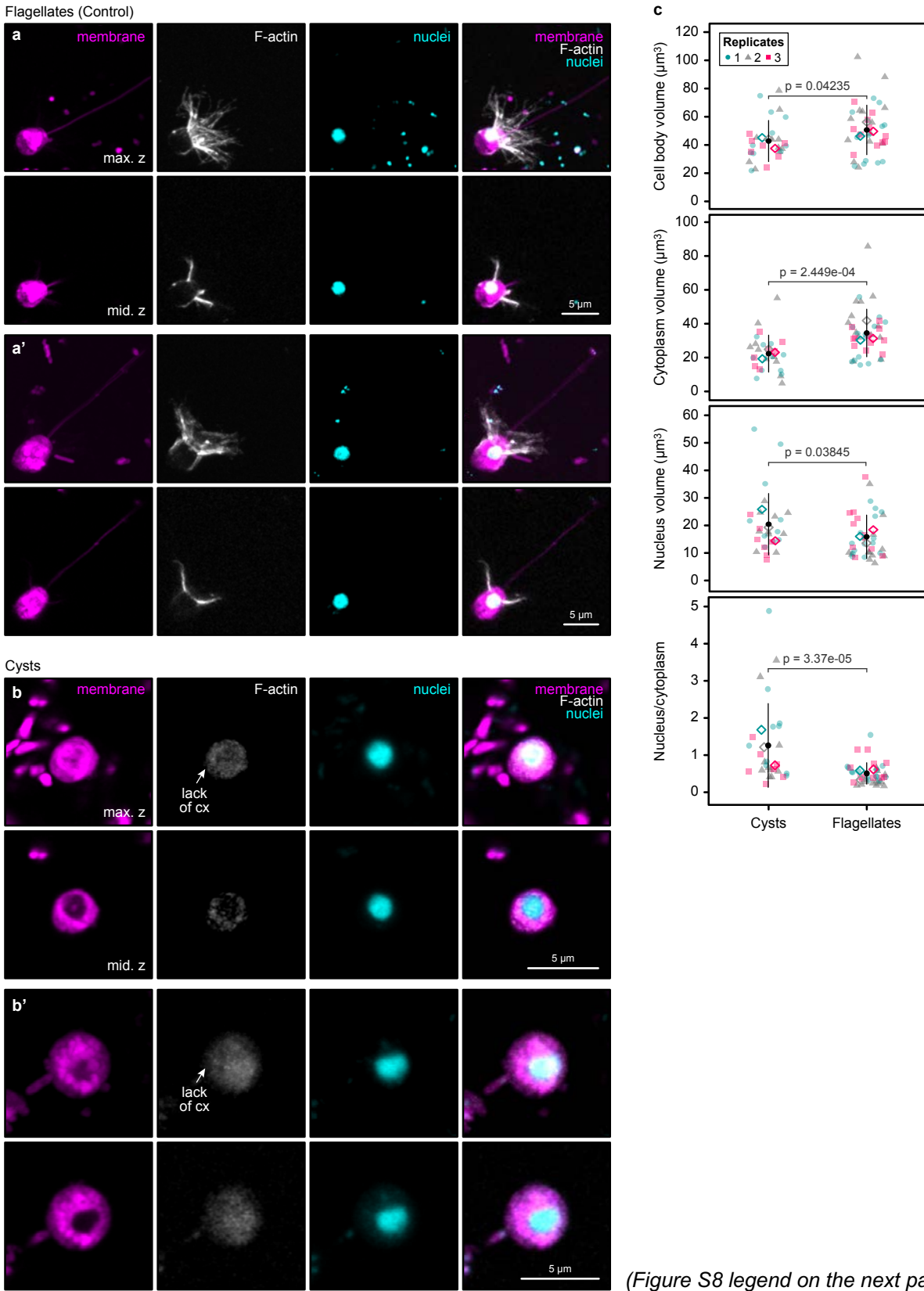


Figure S8. Cysts have a greater nucleus-to-cytoplasm ratio than flagellate cells. (a-b) Maximum z-projections (max. z, upper panels) and mid z-sections (mid. z, lower panels) of Airyscan confocal images showing two representative flagellate cells (a-a') (low-evaporation control) and two representative cysts (b-b') stained with membrane (FM 4-64FX, magenta), F-actin (Phalloidin 488, white) and nucleus (Hoechst, cyan) dyes. (c) Quantification of cell body, cytoplasm, nucleus volume, and nucleus-to-cytoplasm ratio of individual cysts (n=27) and flagellate cells (n=38) in a-b. N=3 independent biological replicates, two technical replicates each. Error bars in c are represented as mean (black circles) $\pm$ s.d. Mean values of each independent replicate are represented as diamonds. Statistics in c by the Mann-Whitney U test (p-value). Figure related to **Figure 4n**.

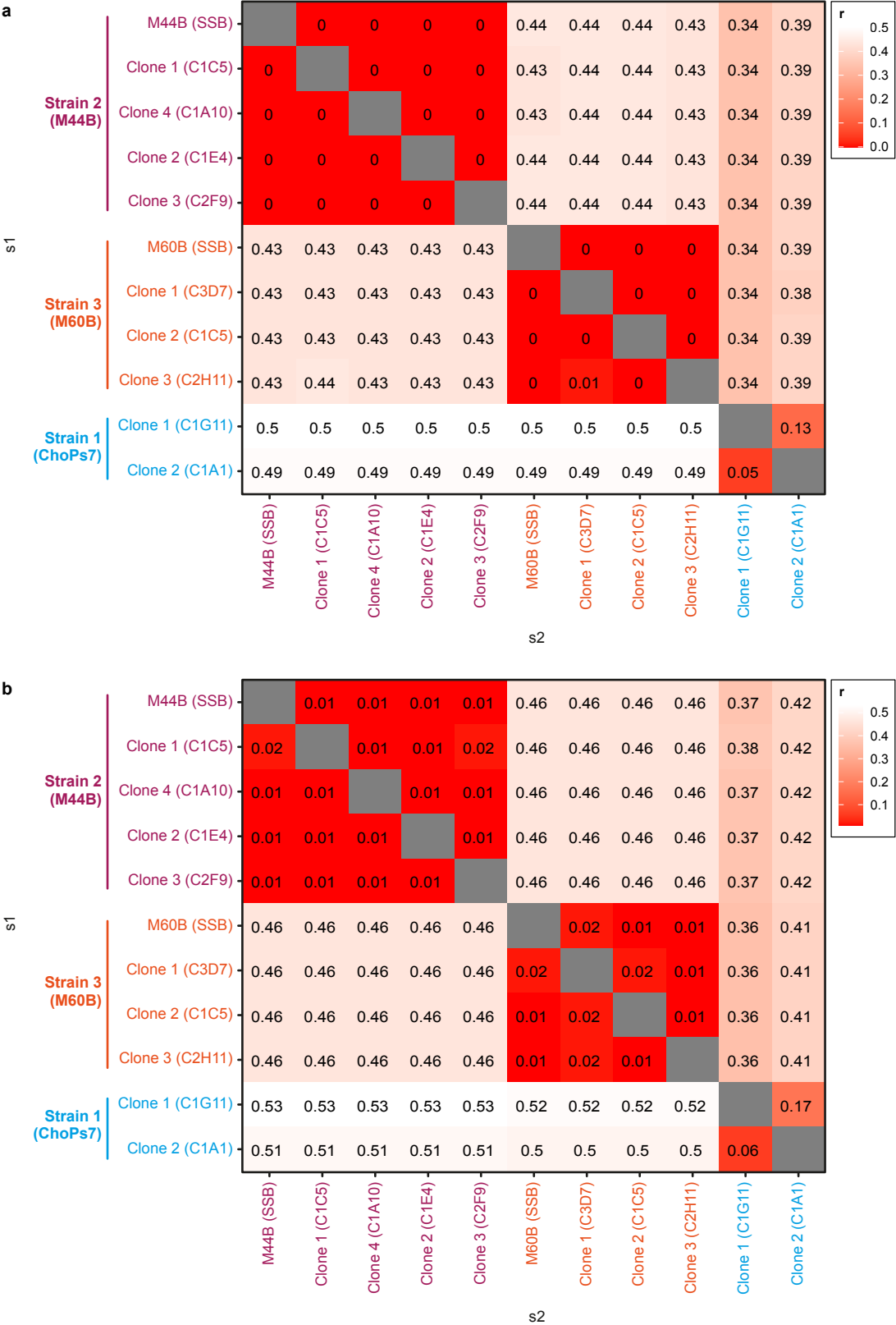


Figure S9. Genetic variation is comparable within single-sheet-bottlenecked and single-cell-bottlenecked strains. (a) The ratio of SNPs present in s1 sample but not in s2 sample, to all SNPs in s1 sample. The ratio of SNPs unique to a single-sheet-bottleneck strain (compared with one of its single-cell-bottleneck strains) is comparable to the ratio of SNPs unique to the single-cell-bottleneck strain (compared with the single-sheet-bottleneck strain). For example, the ratio of unique SNPs of Strain 2 (M44B SSB) compared with Clone 4 (C1A10) is 0, and the ratio of unique SNPs of Clone 4 (C1A10) compared with Strain 2 (M44B SSB) is 0. This suggests similar genetic variation within Strain 2 (M44B SSB) and within Clone 4 (C1A10). (b) The ratio of indels present in s1 sample but not in s2 sample, to all SNPs in s1 sample. Similar to (a), the ratio of indels unique to the single-sheet-bottleneck strains compared with its single-cell-bottleneck strain is comparable to the ratio of indels unique to single-cell-bottleneck strain compared with the single-sheet-bottleneck strain. Figure related to **Figure 6c**.

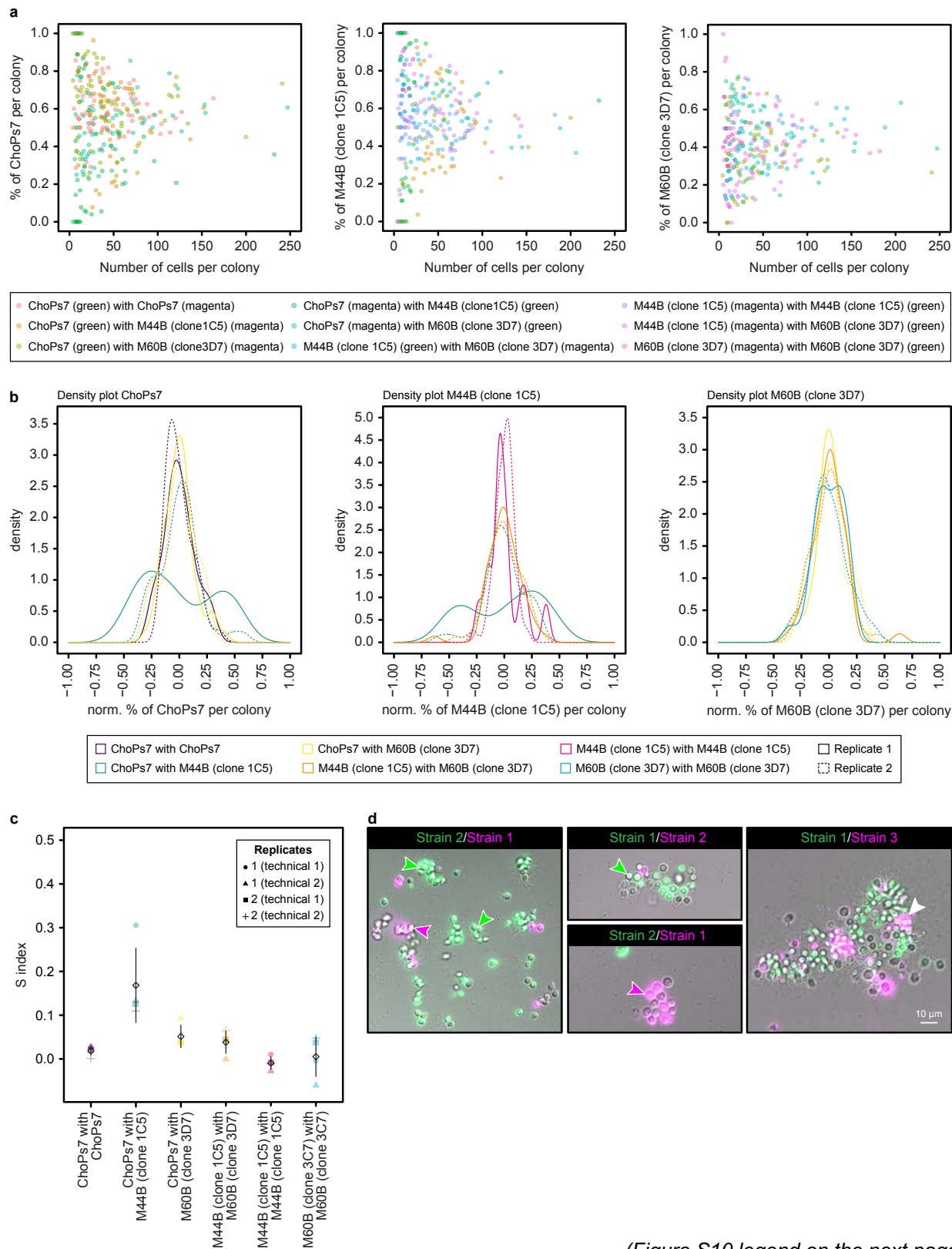


Figure S10. Aggregation is constrained by kin recognition between distinct *C. flexa* strains. (a) Distribution of number of cells per colony in all pairwise strain combinations, showing similar colony size regardless of strain. (b) Normalised density plot of relative percentage of cells of a given strain in all pairwise combinations. (c) Quantification of a segregation index (s index) in different pairwise strain combinations (see *Materials and Methods*). (d) Cells of Strains 1 and 2 (clone 1C5), marked with distinct fluorophores, preferentially aggregate with their own strain, resulting in colonies with little or no chimerism (magenta and green arrowheads). This indicates kin recognition. Cells of Strains 1 and 3 (clone 3D7), marked with distinct fluorophores, readily aggregate into chimeric sheets (white arrowhead). N=2 independent biological replicates, two technical replicates each (n=530 sheets). Figure related to **Figure 6d-g**.

Figure S11

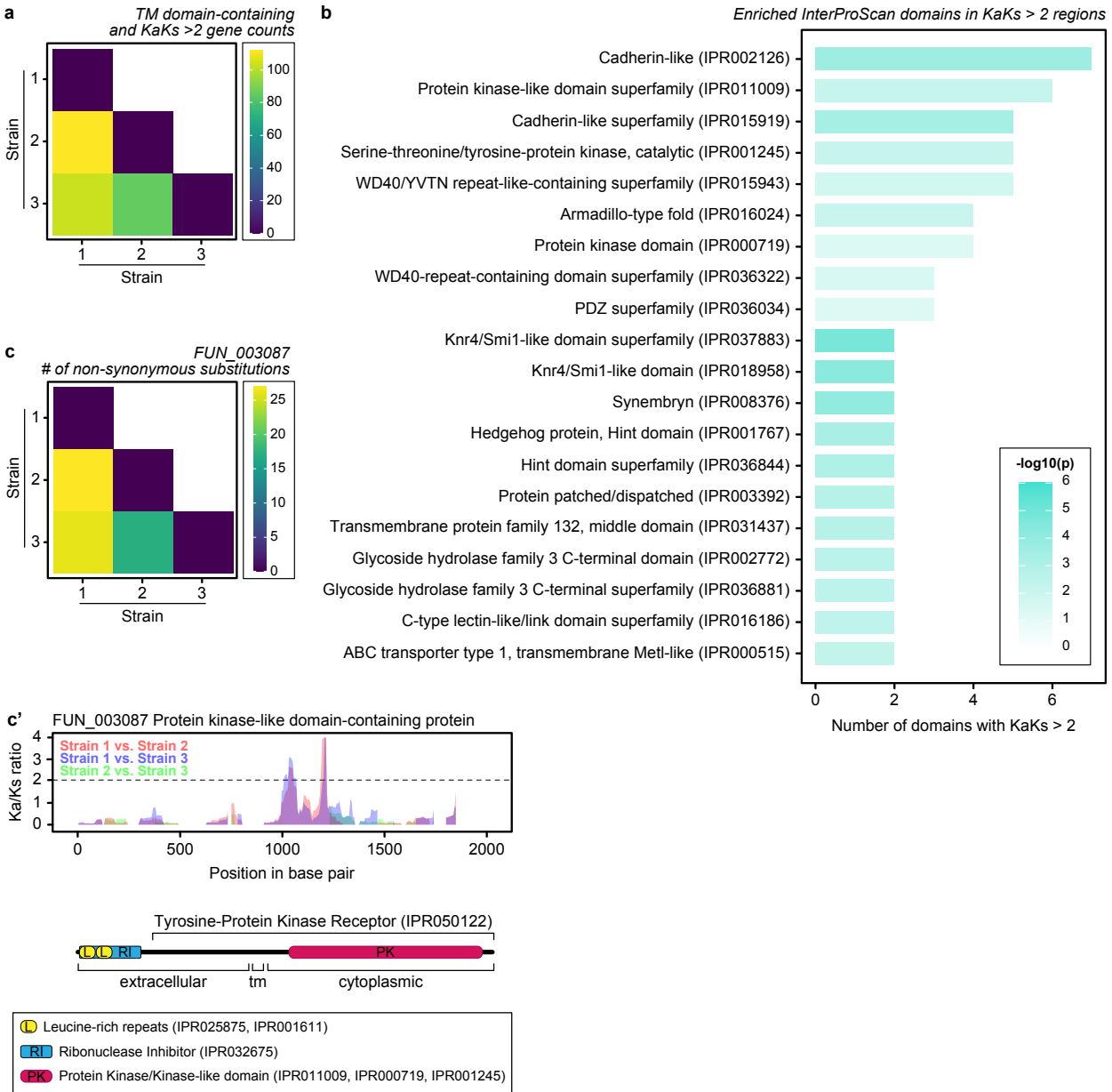


Figure S11. Candidate kin recognition loci with high Ka/Ks ratios. (a) Count of genes with a transmembrane domain and high Ka/Ks ratio regions (Ka/Ks > 2, see *Materials and Methods* section for details) for each combination of strains. (b) Top 20 InterProScan domain annotations enriched in high Ka/Ks ratio regions (Ka/Ks > 2) between Strain 1 and Strain 2. Raw list before removing redundant or similar domains (p-value, Fisher's Exact Test). (c) Number of non-synonymous substitutions in a candidate kin recognition gene encoding a protein kinase-like domain-containing transmembrane protein (FUN_003087) across all pairwise strain combinations. (c') (Upper row) Ka/Ks ratio along the gene coding sequence for each pairwise strain combination (red: Strain 1 vs. Strain 2; blue: Strain 1 vs. Strain 3; green: Strain 2 vs. Strain 3). (Lower row) InterProScan domain architecture schematics. Tm: transmembrane domain. Figure related to **Figure 6h-j**.

Supplementary Tables

Table S1 provided in a separate file.

Table S2. Primers used for 18S rDNA amplification and sequencing. Table related to **Extended Data Figure 2d.**

	Forward Primer	Reverse Primer
First PCR	CTCAARGAYTAAGCCATGCA	CCGCCCCAGYCAAACCTCCC
Nested PCR	GAAACTGCGAATGGCTC	ACCTACGGAAACCTTGTTACG
Sequencing	CGACTTTACGGAAGAGTTGTA	

Supplementary references

1. Brunet, T. *et al.* Light-regulated collective contractility in a multicellular choanoflagellate. *Science* **366**, 326–334 (2019).
2. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682 (2012).
3. Fairclough, S. R., Dayel, M. J. & King, N. Multicellular development in a choanoflagellate. *Current Biology* **20**, R875–R876 (2010).
4. AgiSoft Metashape. AgiSoft Metashape Profesional Edition. (2024).
5. Lemoine, F. *et al.* NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Research* **47**, W260–W265 (2019).
6. Dereeper, A. *et al.* Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**, W465–469 (2008).
7. Castresana, J. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution* **17**, 540–552 (2000).
8. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**, 696–704 (2003).
9. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307–321 (2010).
10. Pachitariu, M. & Stringer, C. Cellpose 2.0: how to train your own model. *Nat Methods* **19**, 1634–1641 (2022).
11. Van Der Walt, S. *et al.* scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014).

1746 12. Putnam, N. H. *et al.* Chromosome-scale shotgun assembly using an in vitro method for long-
1747 range linkage. *Genome Res.* **26**, 342–350 (2016).

1748 13. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo
1749 assembly using phased assembly graphs with hifiasm. *Nat Methods* **18**, 170–175 (2021).

1750 14. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform.
1751 *Bioinformatics* **26**, 589–595 (2010).

1752 15. Faust, G. G. & Hall, I. M. *SAMBLASTER* : fast duplicate marking and structural variant read
1753 extraction. *Bioinformatics* **30**, 2503–2505 (2014).

1754 16. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–
1755 2079 (2009).

1756 17. Phase Genomics. Aligning and QCing Phase Genomics Hi-C Data. (2019).

1757 18. Bickhart, D. M. *et al.* Single-molecule sequencing and chromatin conformation capture
1758 enable de novo reference assembly of the domestic goat genome. *Nat Genet* **49**, 643–650
1759 (2017).

1760 19. Burton, J. N. *et al.* Chromosome-scale scaffolding of de novo genome assemblies based on
1761 chromatin interactions. *Nat Biotechnol* **31**, 1119–1125 (2013).

1762 20. Durand, N. C. *et al.* Juicebox Provides a Visualization System for Hi-C Contact Maps with
1763 Unlimited Zoom. *Cell Systems* **3**, 99–101 (2016).

1764 21. Rao, S. S. P. *et al.* A 3D map of the human genome at kilobase resolution reveals principles
1765 of chromatin looping. *Cell* **159**, 1665–1680 (2014).

1766 22. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome
1767 alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907–915
1768 (2019).

- 1769 23. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina
1770 sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
- 1771 24. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–
1772 3100 (2018).
- 1773 25. King, N. *et al.* The genome of the choanoflagellate *Monosiga brevicollis* and the origin of
1774 metazoans. *Nature* **451**, 783–788 (2008).
- 1775 26. Richter, D. J., Fozouni, P., Eisen, M. B. & King, N. Gene family innovation, conservation
1776 and loss on the animal stem lineage. *eLife* **7**, e34226 (2018).
- 1777 27. Buchfink, B., Reuter, K. & Drost, H.-G. Sensitive protein alignments at tree-of-life scale
1778 using DIAMOND. *Nat Methods* **18**, 366–368 (2021).
- 1779 28. Slater, G. S. C. & Birney, E. Automated generation of heuristics for biological sequence
1780 comparison. *BMC Bioinformatics* **6**, 31 (2005).
- 1781 29. Stanke, M. *et al.* AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids*
1782 *Research* **34**, W435–W439 (2006).
- 1783 30. Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).
- 1784 31. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M.
1785 BUSCO: assessing genome assembly and annotation completeness with single-copy
1786 orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
- 1787 32. Majoros, W. H., Pertea, M. & Salzberg, S. L. TigrScan and GlimmerHMM: two open source
1788 *ab initio* eukaryotic gene-finders. *Bioinformatics* **20**, 2878–2879 (2004).
- 1789 33. Haas, B. J. *et al.* Automated eukaryotic gene structure annotation using EVIDENCEModeler
1790 and the Program to Assemble Spliced Alignments. *Genome Biol* **9**, R7 (2008).
- 1791 34. Palmer, J. M. & Stajich, J. E. Funannotate. (2023).

- 1792 35. Kuznetsov, D. *et al.* OrthoDB v11: annotation of orthologs in the widest sampling of
1793 organismal diversity. *Nucleic Acids Res* **51**, D445–D451 (2023).
- 1794 36. Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A. & Zdobnov, E. M. BUSCO Update:
1795 Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage
1796 for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Molecular Biology and*
1797 *Evolution* **38**, 4647–4654 (2021).
- 1798 37. Chan, P. P. & Lowe, T. M. tRNAscan-SE: Searching for tRNA Genes in Genomic
1799 Sequences. in *Gene Prediction* (ed. Kollmar, M.) vol. 1962 1–14 (Springer New York, New
1800 York, NY, 2019).
- 1801 38. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421
1802 (2009).
- 1803 39. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor.
1804 *Bioinformatics* **34**, i884–i890 (2018).
- 1805 40. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
1806 Preprint at <https://doi.org/10.48550/ARXIV.1303.3997> (2013).
- 1807 41. Auwera, G. A. V. de & O'Connor, B. D. *Genomics in the Cloud: Using Docker, GATK, and*
1808 *WDL in Terra*. (O'Reilly, Beijing Boston Farnham Sebastopol Tokyo, 2020).
- 1809 42. Poplin, R. *et al.* Scaling accurate genetic variant discovery to tens of thousands of samples.
1810 Preprint at <https://doi.org/10.1101/201178> (2017).
- 1811 43. Augusto Corrêa Dos Santos, R., Goldman, G. H. & Riaño-Pachón, D. M. ploidyNGS:
1812 visually exploring ploidy with Next Generation Sequencing data. *Bioinformatics* **33**, 2575–
1813 2576 (2017).

- 1814 44. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008
1815 (2021).
- 1816 45. Knaus, B. J. & Grünwald, N. J. *vcfr* : a package to manipulate and visualize variant call
1817 format data in R. *Molecular Ecology Resources* **17**, 44–53 (2017).
- 1818 46. Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: A Fast and
1819 Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular*
1820 *Biology and Evolution* **32**, 268–274 (2015).
- 1821 47. Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2:
1822 Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* **35**,
1823 518–522 (2018).
- 1824 48. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic
1825 tree display and annotation tool. *Nucleic Acids Research* **52**, W78–W82 (2024).
- 1826 49. Knaus, B. J. & Grünwald, N. J. VCFR : a package to manipulate and visualize variant call
1827 format data in R. *Molecular Ecology Resources* **17**, 44–53 (2017).
- 1828 50. Wickham, H. *et al.* Welcome to the Tidyverse. *JOSS* **4**, 1686 (2019).
- 1829 51. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria:
1830 R Foundation for Statistical Computing. *R Foundation for Statistical Computing*
1831 <https://www.R-project.org/>.
- 1832 52. Posit team. RStudio: Integrated Development Environment for R. Posit Software, PBC,
1833 Boston, MA (2025).
- 1834 53. Blum, M. *et al.* The InterPro protein families and domains database: 20 years on. *Nucleic*
1835 *Acids Research* **49**, D344–D354 (2021).

- 1836 54. Jones, P. *et al.* InterProScan 5: genome-scale protein function classification. *Bioinformatics*
1837 **30**, 1236–1240 (2014).
- 1838 55. Estrela, S. & Brown, S. P. Metabolic and Demographic Feedbacks Shape the Emergent
1839 Spatial Structure and Function of Microbial Communities. *PLoS Comput Biol* **9**, e1003398
1840 (2013).
- 1841