Multiscale spatial segregation analysis in digital images of biofilms

Supplementary Materials

Iztok Dogsa[1*], Ines Mandic-Mulec[1]

[1]Chair of Microbiology, Department of Food Science and Technology, Biotechnical Faculty, University of Ljubljana, Večna pot 111,1000 Ljubljana, Slovenia, EU

*Correspondence to: iztok.dogsa@bf.uni-lj.si, +386-1-3203414

SI Results

The uncertainty in multiscale spatial segregation analysis

Due to the random sampling with field of view (FOV), which is a square of certain size *d*, n times across the digital image, the obtained weight average mean, $\widehat{S_d}$, (eq 5) of local segregation levels $S_{d,i}$, (eq 3) has the standard deviation:

$$SD_d = \sqrt{\frac{\sum_{i=1}^{n} w_{d,i}(\widehat{S_d} - S_{d,i})^2}{ws}}, \qquad \text{(eq S1)}$$

where $ws = \sum_{i=1}^{n} w_{d,i}$. The standard error of weight average mean can be approximately calculated as given by Gatz et al. (1995) and is used in our approach to estimate the uncertainty of segregation level ($\widehat{S_d}$):

$$SE_d = \sqrt{\frac{n}{(n-1)ws_d^2} \sum_{i=1}^{n} w_{d,i}^2 (\widehat{S_d} - S_{d,i})^2}\,, \qquad \text{(eq S2)}$$

The $\widehat{S_d}$ and SE_d can be calculated for each size of FOV, *d* and to have SE_d / $\widehat{S_d}$ relatively constant across all *d*, one could empirically determine *n* (the number of randomly placed FOV of size *d* in an image) for each *d* separately and then calculate final SE_d and $\widehat{S_d}$, which will be very time-consuming. Alternatively, one can find an expression that will approximate the scaling of *n* with *d* in order to keep SE_d / $\widehat{S_d}$ approximately constant. As the number of differently placed FOV in the image grows with the square of the difference in size of an image (*dmax*) and size of FOV *d*, we assumed that *n* necessary to keep SE_d / $\widehat{S_d}$ constant across all *d* can be calculated as follows:

$$n = \alpha(d_{max} - d)^2 + 3, \qquad \text{(eq S3)}$$

where α (sampling factor) is a non-negative tunable parameter that influences linearly the size of *n*. If uncertainty (SE_d) is for some chosen α too large, α needs to be increased (see Macro implementation, main text, for more details). The addition of number 3 keeps the minimum value of $n \geq 3$ to have at least some statistical meaning of eq S1 and eq S2.

The standard error estimate for multiscale spatial segregation level in eq 7 to 10 (main text) can be calculated by the standard error propagation formula.

Multiscale spatial segregation analysis in 3D digital images

The common approach in microscopy to obtain 3D spatial information is to use confocal microscopy, where several 2D images of different planes (the depth of the sample) are acquired and stacked together, followed by the reconstruction of the 3D image of the sample. The most obvious transformation in our approach for 2D image to 3D space would be to replace the 2D FOV with a 3D variant of the FOV; instead of a square of size *d* we have moved across the image, we could now use a cube with side of size *d* that we move across a stack. In the case of 2D FOV, the size run from some *dmin* (e.g., pixel size of dimensions *dmin* x *dmin*) to *dmax* applied typically on images of *dmax* x *dmax* dimensions.

However, in typical confocal microscopy 3D images, the x and y dimensions are often 10 to 100 times bigger than the z-dimension. As the *dmax* cannot exceed the smallest of x,y,z dimensions of the 3D image, which is usually the sample thickness (z-dimension), it is not possible to run dimensions of 3D FOV with *dmax* x *dmax* x *dmax* to cover the entire image volume in single 3D FOV. Furthermore, the resolution in the z-dimension in confocal images is significantly lower than in the x,y dimension, often exceeding the dimension of the segregated particles under the observation. Therefore, our approach treats 3D images as a stack of slices, where every slice is treated as a 2D image that is scanned with 2D FOVs of maximal dimension corresponding to the image size, just as in the case of multiscale spatial segregation analysis of 2D images described through eq 1 to eq 5 (main text). The only difference is now that when obtaining $S_{d,i}$ to calculate the segregation level in eq 5, FOV is not sampled only in single 2D image but through the entire stack of images. Considering stack position, all $S_{d,i}$ contribute equally to their averages in eq 5. The total areas, *A* and *B*, correspond to the number of particles in the entire stack. With this approach, we are evaluating segregation level in the x,y direction, but in a 3D object. Importantly, although the slices in the stack are represented as 2D images, each slice contains the optical information according to the optical slice thickness; i.e., the confocality of the microscope is not perfect, therefore the emitted light from the sample comes

not only from one infestimaly thin plane but from several planes that comprise the optical thickness of the slice. This can be calculated according to pinhole diameter, the numerical aperture (NA) of the objective, light wavelength and the refractive index of immersion medium (Wilhelm, 2010). The optical slice thickness often exceeds the size of the particle (e.g., bacterium) in the image, especially when working with objectives that can capture large dimensions (*dmax*) of images as they have typically lower NA, making the optical slice thickness relatively high. For example, a typical 10x/0.3 NA objective can capture a confocal image of dimensions 600 µm x 600 µm, with a resolution in x,y of about 0.6 µm, with a typical optical slice thickness of about 13 µm, which is much thicker than the typical size of a bacterium (about 1 µm). Consequently, the two bacteria cannot physically occupy the same position in 3D space, but can be placed one on top of another and share the same pixels in the image slice, which means that even if we evaluate segregation level, $\widehat{S_d}$, at *dmin* = 1 pix, we will not get $\widehat{S_d}$ = 1. Therefore, in such cases, the simulated minimum segregation level, $\widehat{S_d^{min}}$, allows particle overlapping.

The outcome of such a simulation is a randomly mixed sample with a zero segregation level for all dimensions of FOV, which also means that minimal multiscale spatial segregation (*MSSLmin*, eq 7) is zero. The *dmin* for the evaluation of $\widehat{S_d}$ should not go below the optical slice thickness, otherwise the predominant dimension of FOV will be optical slice thickness and $\widehat{S_d}$ will be dominated by segregation in the z-axis, making the interpretation of $\widehat{S_d}$ as a function of FOV dimension *d* more complicated. Of course, if one is able to improve the optical sectioning, either by using better optics or by deconvolution, the *dmin* can be lowered; ideally, we will be able to optically resolve the two bacteria placed on top of each other and get $\widehat{S_d}$(*dmin*) = 1. As in the case of the multiscale spatial segregation analysis in 2D images, the maximal segregation level, $\widehat{S_d^{max}}$, will be for all *d* equal to 1, as two types of particles displace each other out of the

observed space. This means only one type of the particle is present in the stack of slices or the particles are separated among slices, so that in each slice only one type of particle is present.

Validating multiscale spatial segregation analysis by application to synthetic images

The black pixel-sized particles completely excluded the white pixel-sized particles from space of the image [(Fig. 1A), (d), main text]. As expected in Section 2.1. (main text), one can observe in Fig. 1B the segregation level, $\widehat{S_d}$ = 1, regardless of the spatial scale of the observation (*d* of FOV). On the other hand, in a case where none of the particles can exit the system, the segregation maximum will have a different appearance (Fig. 1A), (c). Here, white and black pixel-sized particles coexist at the spatial scale of the entire image (i.e., when the FOV is of *dmax*) and both types of particles are present in equal (relative) numbers. Correspondingly, the $\widehat{S_d}$ at *dmax* = 0 (Fig. 1B). However, when we turn to smaller spatial scales (smaller *d* of FOV), one of the two type of particles becomes quickly dominant and, as expected, segregation increases rapidly. Very soon, FOVs containing only one type of particle become dominant and segregation at this spatial scale has to approach 1, as can indeed be observed in Fig. 1B. On the contrary, the segregation level has to be significantly lower in Fig. 1A (b), where image (c) was pixel randomized and overlapping of opposite pixels was not allowed. Here, as predicted, the segregation level, $\widehat{S_d}$, is close to zero for all *d* (FOV sizes), except for *d* approaching particle size (pixel in this case), where $\widehat{S_d}$ rapidly increases to 1 (Fig. 1B). When overlapping is permitted, as is the case in image (a), the $\widehat{S_d}$ correctly drops further than is the case in image (b). At *d* = 1, where the size of the FOV is 1 x 1 pixel, half of the FOV contains either black or white pixel particle (S_d = 1) and half of the FOV contains black and white pixel particle (S_d = 0), giving the predicted segregation level, $\widehat{S_d}$ = 0.5 (Fig. 1B). The challenging tests are more

complex chessboard images of (e)–(h) in Fig. 1A. As can be seen in the graph in Fig. 1B, our approach correctly predicts the chessboard segregation levels, $\widehat{S_d}$; with smaller chessboard squares, the size of the FOV, where $\widehat{S_d}$ of pixel-sized particles starts to rapidly increase, decreases. Accordingly, the *MSSL*, (eq 7, main text), which is the size of the FOV-averaged segregation level, decreases in the direction of the decreasing size of the chessboard squares (Table S1). As the chessboard squares consist of several basic pixel-sized particles that are clustered together, the *MSSL* does not reach the *MSSL* value for the minimal segregation case. Also, the *MSSL* of the maximal segregation case is not reached, because in image (e), i.e., the 2 x 2 chessboard, the mixing is still better than in the case of image (c), where pixel-sized particles minimized contacts in order to avoid each other. The relative *MSSL* (eq 10, main text), which takes explicitly into account the maximum and minimum segregation cases and is denoted $rMSSL$ (Table S1), tells us what the multiscale segregation level is on a scale from 0 to 1. As expected, the segregation in the 2 x 2 chessboard with *rMSSL* = 0.72 ± 0.05 is relatively close to the maximum segregation, where the black and white pixel particles are not allowed to exclude each other from the space (image (c) in Fig. 1B). Doubling the number of chessboard squares roughly halves the *rMSSL*.

Mixture of oats and raisins

Digital images depicting a real-life scene with a predictable multiscale segregation analysis outcome are shown in the middle of Fig. S1A. In this case, we begin by adding raisins to the edge of a plate of oats. This initial state (t0) presents the maximally segregated situation and its segregation level curve is indistinguishable from the modelled maximum segregation case. As a result, *rMSSL*, which indicates how strongly the two types of particles (oats and raisins) in an image segregate on the scale from 0 to 1, is 0.95 ± 0.09 (Fig. S1A, Table S2). This

indicates that the segregation is maximal in the sample image. As we mix the sample (t1 to t5), the *rMSSL* drops until it finally reaches 0.05 ± 0.03 (Fig. S1A, Table S2), indicating that the sample is well mixed (i.e., almost the minimum possible segregation of 0.00 has been reached). The plot of segregation levels versus the size of the FOV in Fig. S1B represents segregation changes with a different spatial scale of observation. For example, when the size of the FOV is 400 x 400 pixels, which corresponds to a quarter of the image, most of the FOVs will consist of either oats or raisin at t0; therefore, the segregation level, $\widehat{S_d}$, at this size of the FOV is close to 1. At t1, $\widehat{S_d}$ drops to 0.7, indicating that a notable part of the 400 x 400 pixel-sized FOVs is now occupied by oats and raisins; at t5, $\widehat{S_d}$ is only 0.2, the same as in the minimal segregation simulation, indicating the oats and raisins are now, as expected, fully mixed. However, when the size of the FOV is scaled down to the size of the raisins (about 70 pixels), the differences between segregation levels at different times of mixing become less prominent. On scales smaller than the size of the raisins, the sample is intrinsically segregated as the oats and the raisins cannot interpenetrate each other and therefore the segregation level is expectedly approaching the value of 1.

As expected, the mixing of the oats with raisins decreased their segregation over time on all spatial scales, except at the spatial scales of raisins. As the proposed multiscale spatial segregation analysis of the real-life images with clearly visible segregation level returned expected results, we focused on segregation cases with less obvious outcomes.

SI Methods

Macro implementation

For the conversion of the original two channel images or stacks (each channel for one type of particle) into two binary images especially suitable for microscopy, the user can run a macro *Convert_to_bin* (Fig. S3). Alternatively, different well-established approaches can be taken by the user in the FIJI environment to obtain the two binary images (e.g., thresholding colour images and converting them to binary images). The purpose of the next macro, *Sim_seg_extremes*, is to simulate minimally and maximally segregated cases of the particles of two types based on two binary images of size (or stack of images) *dmax x dmax* derived from the original image (or stack of images), each representing one type of the particle. The outputs of the macro *Sim_seg_extremes* are images displaying randomly placed circular particles of each type with set size (i.e., minimal segregation images), and images where particles are set farthest away or are even excluded from the image (i.e., maximal segregation images). The user can also set the parameter defining whether the overlapping of the particles is allowed.

Images representing the segregation extremes serve together with the original binary images as an input for the second macro, *MSS_calc*, which calculates segregation measures in eq 5 and 7 to 10 (main text). The macro *MSS_calc* further accepts 8-bit binary images.

Besides input images representing two particle types, the *MSS_calc* macro requires a user to enter input parameters required to calculate segregation level (eq 5, main text) as a function of size *d* of the field of view (FOV). The FOV is a square of size *d* x *d*, which is randomly moved across the image. The pixels inside this square are used to calculate the local segregation level $S_{d,i}$ according to eq 3 (min text). The input parameters include *dmin* and *dmax*, corresponding to the smallest and largest FOV used to calculate $S_{d,i}$ and cannot be less than a single pixel or size of the image, respectively. One also has to select the resolution factor, which determines the number of different *d* at which $S_{d,i}$ will be calculated and by which factor the values of *d* will be spaced, in the direction from *dmax* to *dmin*. Finally, the sampling factor α, which determines how many local segregation levels, $S_{d,i}$, will be calculated from randomly placed

FOVs of size d in the image (eq S3), is entered. After obtaining sufficient local segregation levels, $S_{d,i}$, the macro first calculates the segregation levels as a function of size d of the FOV: $\widehat{S_d}$, $\widehat{S_d^{max}}$ $\widehat{S_d^{min}}$ and then determines $MSSL_{max}$, $MSSL_{min}$, $MSSL$ and $rMSSL$ (eq 5 and 7 to 10, main text). The segregation measures obtained are then stored in the table and saved to the file. For more information about the macros, please see user manual and examples available at https://github.com/IztokD/MSSegregation-package).

Acquisition of example images and macro settings

Synthetic images for testing

The corresponding minimum and maximum segregation images were obtained by our *Sim_seg_extremes* macro. The parameters in the macro *MSS_calc* were *dmin* = 1 pixel, corresponding to the modelled pixel sized particle and *dmax* = 246 pixel, corresponding to the size of the image. To obtain at least 10,000 randomly placed FOVs for segregation level evaluation at *dmin* and standard error < 2%, the sampling factor was set to 0.2. The resolution factor was set to 0.95, which means that from *dmax* to *dmin* of FOV, 107 points will be plotted on a graph of segregation level vs. size *d* of FOV; each point is plotted at *d* is smaller by a factor of 0.95 from the preceding *d* in the direction from *dmax* to *dmin*.

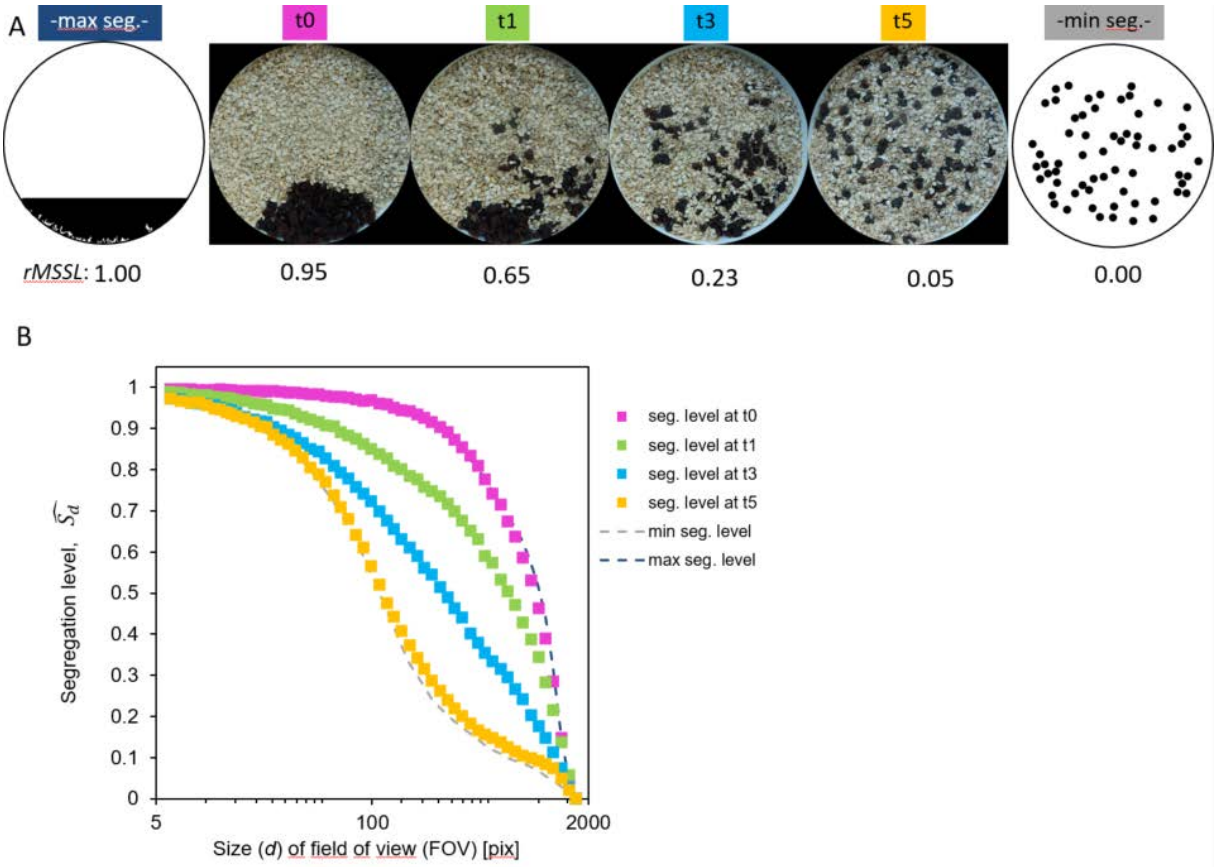
Oat and raisins images

The ceramic plate was filled with dry oats and raisins were added and manually separated from the oats to represent the situation before mixing. The photographs (1672 x 1672 pixels) of the mixture were than periodically taken with a Nikon D40 camera, while the mixture was manually

mixed with a spoon. The settings for macro package *MSSegregation* are given in Supplementary information.

The corresponding minimum and maximum segregation images were obtained by our *Sim_seg_extremes* macro. For multiscale spatial segregation analysis (macro *MSS_calc*), the sampling factor was set to 0.005, which gave around 10,000 randomly placed FOVs for segregation level evaluation at *dmin*. The standard error of the segregation level was in all sizes of FOV < 2%. The resolution factor was set to 0.9, giving a range of 55 points on a graph of segregation level vs. size (*d*) of FOV. For simulation of the minimal and maximal segregation, a raisin was assumed to be a circle 72 pixels in diameter. The simulated raisins were randomly distributed in the circle representing the original plate, and overlap was allowed. For the maximal segregation case, the placing of the raisins was limited to the edge of the circle.

SI Figures, Tables



220

221 Fig. S1: The segregation analysis of oats and raisins. In (A): raisins were added to oats at t0 and then mixed. At
222 indicated time points (t0 to t5), a 1670x1670 pix photograph was taken; minimum segregation (-min seg.-) and
223 maximum segregation (-max seg.-) was simulated with black representing raisins and white representing oats; the
224 corresponding *rMSSL* (eq 10, main text) indicating the distance averaged segregation level relative to min and max
225 seg. level were calculated from segregation levels displayed in (B) and are given below images. The sampling
226 error of algorithm estimated by eq S2 is within the symbol size.

227

228

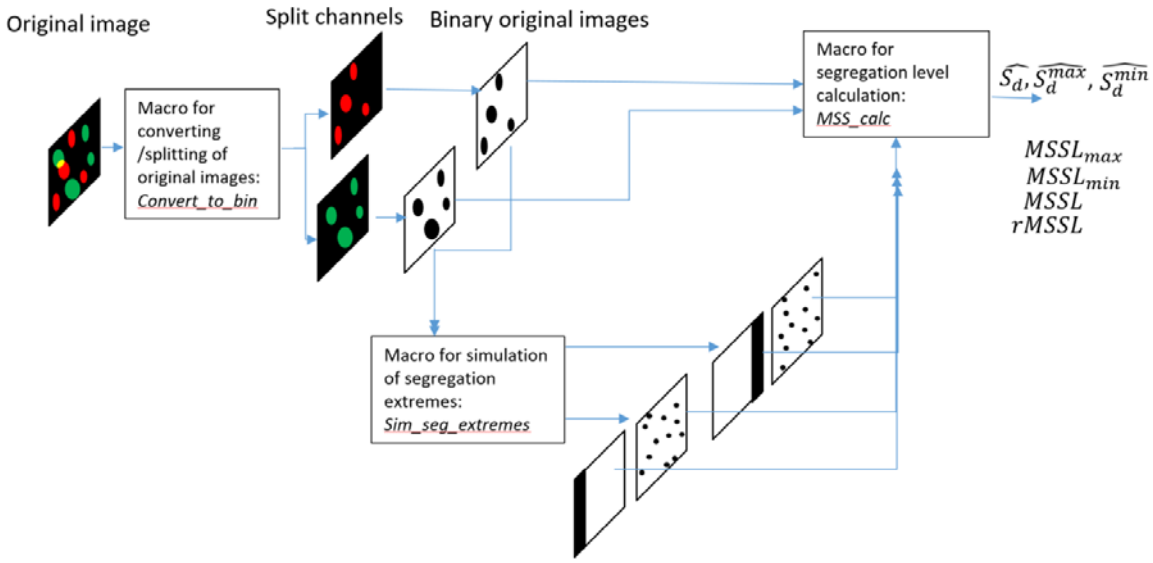


Figure S3: Workflow of multiscale spatial segregation analysis of digital images. The original image containing two types of particles (for example, microscopy image displaying red and green bacteria) is split into single type particle images, following conversion into binary images (i.e. black & white). Binary original images serve as input for Macro that generates maximum and minimum segregation images for each type of the particle. In such images the number of displayed particles is typically proportional to the area occupied by the pixels representing them. From binary original images the segregation level ($\widehat{S_d}$) as a function of size (d) of field of view (FOV) is calculated; maximum and minimum segregation images are used for calculation of maximum and minimum segregation level ($\widehat{S_d^{min}}$, $\widehat{S_d^{max}}$). Finally, multiscale spatial segregation levels (*MSSL*), relative multiscale spatial segregation level (*rMSSL*) are calculated from segregation levels (eq 7–10, main text).

	MSSL	SE (MSSL)	rMSSL	SE (rMSSL)
t0	0.53	0.03	0.95	0.09
t1	0.42	0.03	0.65	0.07
t3	0.27	0.02	0.23	0.06
t5	0.16	0.01	0.05	0.03
min seg.	0.14	0.01	0.00	0.00
max seg.	0.55	0.03	1.00	0.00

Table S1: The segregation statistics of analysis in Fig. 1; *MSSL*-multiscale spatial segregation level (eq 7, main text) with standard error, SE; *rMSSL* - relative multiscale spatial segregation level (eq 10, main text) with standard error, SE.

	MSSL	SE (MSSL)	rMSSL	SE (rMSSL)
t0	0.53	0.03	0.95	0.09
t1	0.42	0.03	0.65	0.07
t3	0.27	0.02	0.23	0.06
t5	0.16	0.01	0.05	0.03
min seg.	0.14	0.01	0.00	0.00
max seg.	0.55	0.03	1.00	0.00

Table S2: The segregation statistics of analysis in Fig. S1; MSSL - multiscale spatial segregation level (eq 7, main text) with standard error, SE; rMSSL - relative multiscale spatial segregation level (eq 10, main text) with standard error, SE.

	MSSL	SE (MSSL)	rMSSL	SE (rMSSL)
PS-218 YFP + PS-68 mKate2	0.046	0.007	0.026	0.010
PS-68 YFP + PS-68 mKate2	0.028	0.005	0.005	0.007
min seg. (PS-218 YFP + PS-68 mKate2)	0.024	0.004	0	0
min seg. (PS-68 YFP + PS-68 mKate2)	0.024	0.004	0	0
max seg.	0.875	0.125	1	0

Table S3: The segregation statistics of analysis in Fig. 2; *MSSL* - multiscale spatial segregation level (eq 7, main text) with standard error, SE; *rMSSL* - relative multiscale spatial segregation level (eq 10, main text) with standard error, SE.

	MSSL	SE (MSSL)	rMSSL	SE (rMSSL)
PS-218 YFP + PS-68 mKate2	0.19	0.02	0.22	0.02
PS-216 mKate2 + PS-196 YFP	0.24	0.01	0.24	0.01
PS-68 YFP + PS-68 mKate2	0.24	0.01	0.24	0.01
min seg.	0.00	0.00	0.00	0.00
max seg.	1.00	0.00	1.00	0.§0034d

Table S4: The segregation statistics of analysis in Fig. 3; *MSSL* - multiscale spatial segregation level (eq 7, main text) with standard error, SE; *rMSSL* - relative multiscale spatial segregation level (eq 10, main text) with standard error, SE.

Strain	Background	Name in the text	Genome description	Reference
BM1097	PS-216	*PS-216 mKate2*	*amyE*::P*hyperspank-mKate2* (Cm)	Stefanic et al., 2015
BM1313	PS-196	*PS-196 YFP*	*amyE*::P*hyperclo3-yfp* (Sp)	Kraigher et al., 2022
BM1125	PS-218	*PS-218 YFP*	*amyE*::P*hyperclo3-yfp* (Sp)	Kraigher et al., 2022
BM1477	PS-68	*PS-68 YFP*	*amyE*::P*hyperclo3-yfp* (Sp)	Kraigher et al., 2022
BM1451	PS-68	*PS-68 mKate2*	*amyE*::P*hyperspank-mKate2* (Cm)	Kraigher et al., 2022

Table S5: *Bacillus subtilis* strains used in this study.

SI References

Gatz, D. F., Smith, L. The standard error of a weighted mean concentration—I. Bootstrapping vs other methods. Atmospheric Environment 29, 1185–1193 (1995).

Kraigher, B., Butolen, M., Stefanic, P., Mandic Mulec, I. Kin discrimination drives territorial exclusion during Bacillus subtilis swarming and restrains exploitation of surfactin. ISME J 16, 833–841 (2022).

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

Stefanic, P., Kraigher, B., Lyons, N. A., Kolter, R., Mandice-Mulec, I. Kin discrimination between sympatric Bacillus subtilis isolates. Proceedings of the National Academy of Sciences of the United States of America 112, 14042–14047 (2015).

Wilhelm, S. (2010) Confocal laser scanning microscopy. Carl Zeiss Microscopy GmbH, Jena. Brochure 60-1-0030/e, ISBN 978-3-940885-02-9