Proceedings of the Royal Society B

Electronic Supplementary Material

for

Defector clustering is linked to cooperation in a pathogenic bacterium

Edward W. Tekwa[1,2]*, Dao Nguyen[3], Michel Loreau[4], Andrew Gonzalez[5]

[1]Department of Biology, McGill University, 1205 Dr. Penfield, Montreal, QC, H3A 1B1, Canada. (wong.tek.wa@gmail.com)

[2](current) Department of Ecology, Evolution, and Natural Resources, Rutgers University, 14 College Farm Road, New Brunswick, NJ, 08901, USA.

[3]Meakins Christie Laboratories, Research Institute of the McGill University Health Centre, and Department of Medicine, McGill University, 1001 Decarie Blvd, Montreal, QC, H4A 3J1, Canada. (dao.nguyen@mcgill.ca)

[4]Theoretical and Experimental Ecology Station, CNRS and Paul Sabatier University, 09200, Moulis, France. (michel.loreau@cnrs.fr)

[5]Department of Biology, McGill University, 1205 Dr. Penfield, Montreal, QC, H3A 1B1, Canada. (andrew.gonzalez@mcgill.ca)

Computing clustering

From the calibration procedure that corrects for the differences in GFP and mCherry fluorescence [1], the Imaris threshold settings for each fluorescent colour were 3.83 μm for GFP and 4 μm for mCherry. These are understood as the estimated fluorescent footprint of each bacterium. Thus, bacteria of the same colour closer than 4 μm apart were likely counted as one. This undercounting bias is weaker for between-morph clustering measurements, because the focal bacterium is of a different colour than the neighbours that are being counted. The between-morph clustering may still be underestimated because the neighbours of another colour may be clustered among themselves, but can serve as a lower limit for clustering estimates.

By comparing mono-fluorescent monocultures with mixed-fluorescent "monocultures" (either cooperators or defectors only, 7 replicates), we found that monocultures were undercounted by a factor of 0.6369 on average. Thus, we infer that a portion M=1-0.6369 comes from missed counts within 4 μm of focal individuals. As well, GFP counts on average were greater than mCherry counts by a factor of 1.1098. Let G_G=1/1.1098, and G_M=1, to account for the GFP and mCherry bias. We thus added $G_i M n_i E[A_4]/A$ to within-morph neighbour counts, where $E[A_4]$ is the expected interaction area with a radius of 4 μm, when non-habitat areas within the radius are subtracted. A is the total habitat area, n_i is the number of morph i individuals, and n_{ii} is the number of morph i neighbours around one focal individual of morph i. We also set the denominator such that C_{ii} approaches 1 as the interaction radius approaches infinity. Thus, the uncorrected raw within-morph clustering coefficient $\tilde{C}_{ii}$ is:

$$\tilde{C}_{ii} = X_{ii} / X_i = \frac{E[G_i(n_{ii}-1)/A_f]}{G_i(n_i-1)/A} \quad \text{(S1)}$$

where A_f is the interaction area at the given scale, with non-habitat areas within the scale subtracted. The corrected version C_{ii} is:

$$C_{ii} = \frac{E[(G_i(n_{ii}-1)+G_i M n_i A_4/A)/A_f]}{(G_i(n_i-1)+G_i M n_i E[A_4]/A)/A} \quad \text{(S2)}$$

Bootstrapping

Bootstrapping was used to estimate the spreads of regression coefficients for the experiment and simulations. The basic idea is to draw randomly and with replacement (resample) a pool of datapoints many times, each time computing the metric of interest (such as the mean, or regression coefficients in our case). The histogram of the computed metric is an

empirical distribution of the metric estimated from the data itself, without first assuming the shape of the distribution (normal, log-normal, etc.). In our bootstrap least-squares linear regression for the experiment, there are 24 replicates that match the dependent variables (cooperator frequency and population density) with the independent variables (clustering coefficients). For each bootstrap regression, we draw 24 replicates randomly and with replacement from the original dataset. For each such resampled dataset, each containing 24 replicates, we perform a multiple linear regression to obtain the effect sizes of clustering coefficients (C_{cc}, C_{cd}, C_{dd}) on cooperator frequency and population density. By repeating this procedure 10000 times, we obtain estimates of the spreads of all clustering effects. Bootstrap regression allows us to obtain effect spreads that are non-parametric, which is important to ascertain their statistical significance. For example, if we assumed a normal distribution for the effect of cooperator clustering on cooperator frequency (Fig. 3c in the main text), it would become less significantly negative. That is because the empirical spread of the effect skews towards extremely negative values and much less often towards positive values, and a normal function fitted to this distribution would estimate a fairly high variance and thus a tail that reaches farther towards 0 effect. This skews the estimate towards the null hypothesis of no effect. Conversely, if the empirical spread skews towards 0, a normality assumption could skew the effect estimate towards falsely rejecting the null hypothesis, which the bootstrap regression would avoid.

Model derivation

Here we derive the mathematical foundations of the dynamic system in equation (1) of the main text. We define local density as X_{ij}, assume that all focal individuals u of morph i

weigh their neighbours by the same function ϕ_{ij}, and note that the expectation of x_{uj} across all u that are morph i is the same as the expectation of the average local density in cell location y, $X_{ij}(y)$, across all y:

$$X_{ij}=\mathrm{E}[x_{uj}]=\frac{1}{N_i}\sum_{u\in i}\sum_{v\in j}\phi_{uj}(y_v-y_u)=\frac{1}{N_i}\sum_{u\in i}\sum_{v\in j}\phi_{ij}(y_v-y_u)$$

(S3)

$$=\mathrm{E}[X_{ij}(y)]=\sum_{q\in\Omega}\phi_{ij}(y-q)\mathrm{E}[N_i(y)N_j(q)]$$

q is the location of potential morph j neighbors, and Ω is the habitat space, which is a countable but infinite set of discrete cells. The expectation $\mathrm{E}[N_i(y)N_j(q)]$ is taken over all cells y. So line 1 of equation (S3) uses individuals u as focal points (with focal location y_u and neighbour location y_v), while line 2 uses space y as focal points (with focal location y and neighbor location q); these are equivalent Lagrangian and Eulerian perspectives. We define the spatial covariance between morph i and j at distance y-q as:

$$\mathrm{Cov}_{ij}(y-z)=\mathrm{E}[N_i(y)N_j(z)]-\mathrm{E}[N_i]\mathrm{E}[N_j]$$

(S4)

where the first expectation over all focal cells y and the second expectation over all cells. As a simple example, we assume that the interaction effect is linear and can be expressed as a_{ij}. Over an ensemble of realizations of the same system configuration, we take Δt to be small enough for only one birth or death event to occur. Then, the expected change in the number of morph i individuals (N_i) in Δt is:

$$\frac{\mathrm{E}[\Delta N_i(y)]}{\Delta t}=\mathrm{E}\left[r_iN_i(y)+\sum_{j=1}^{S}a_{ij}\sum_{q\in\Omega}\phi_{uj}(y-q)\mathrm{E}[N_i(y)N_j(q)]\right]$$

(S5)

$$=r_i\mathrm{E}[N_i]+\sum_{j=1}^{S}a_{ij}\left(\frac{1}{h}\mathrm{E}[N_i]\mathrm{E}[N_j]+\sum_{q\in\Omega}h\phi_{uj}(y-q)\frac{1}{h}\mathrm{Cov}_{ij}(y-q)\right)$$

where h is the area of a cell. Thus, $E[N_i]/h$ is the expected global density of morph i across all cells. We invoked the Eulerian form of equation (S3) to express local densities and equation (S4) to go from line 1 to 2 of equation (S5). We assume that the distribution of individuals is stationary to the second order and isotropic [2], such that the distribution is fully described by global densities and $Cov_{ij}(y\text{-}q)$. Thus, we can move the focal cell y to the origin and rewrite $Cov_{ij}(y\text{-}q)$ as $Cov_{ij}(q)$. In the limit that the cell size h is infinitely small, the point global density of morph i is $X_i=\lim_{h\to 0}N_i/h$ and $cov_{ij}=\lim_{h\to 0}Cov_{ij}/h^2$. We obtain the continuous-time, continuous-space analog of equation (S5) by dividing the equation by h:

$$\frac{dX_i}{X_i dt}=r_i+\sum_{j=1}^{S} a_{ij}\left(X_j+\frac{1}{X_i}\int_{\Omega}\phi_{uj}(q)cov_{ij}(q)dq \right) \quad \text{(S6)}$$

The bracketed term in equation (S6) is the continuous-space definition of local density X_{ij}, which is a combination of the first and second spatial moments (X_j and cov_{ij}). We did not assume that birth is associated with seed dispersal as was done for plant interactions in the original spatial moment derivation [2]; rather, we assume that movement can take place at any time, which is realistic for organisms such as bacteria. Note that movement does not affect morph densities directly because it is simply a spatial redistribution of individuals, but it affects local densities through changing cov_{ij} [3]. As well, cov_{ij} will be a function of higher moments (densities of triplets and so on) through birth and death processes, so equation (S6) and equation (1) do not constitute a closed set of equations. However, they do sufficiently establish local densities as variables of interest.

Finally, for the cooperation-competition model in the main text, we substitute the clustering coefficient in place of local density using $X_{ij}= C_{ij} X_j$, and define the interaction effects as $a_{ic}=-k+a$, and $a_{ij}=-k$. Equation (S6) then becomes equation (1).

Based on equation (1), we can write the intrinsic growth rate r, clustering to cooperators C_c, and clustering to defector C_d of an individual of character z as:

$$\begin{aligned} r(z) &= r_d - (r_c - r_d)z \\ C_c(z) &= C_{cd} - (C_{cc} - C_{cd})z \\ C_d(z) &= C_{dd} - (C_{cd} - C_{dd})z \end{aligned} \tag{S7}$$

The substitutions in equation (S7) allow us to write the fitness of an individual as a single expression of the form $w(z)$:

$$w(z) = r_d - (r_c - r_d)z - (k-a)X_c(C_{cd} - (C_{cc} - C_{cd})z) - kX_d(C_{dd} - (C_{cd} - C_{dd})z) \tag{S8}$$

We then analyse the selection pressure that each clustering coefficient exerts on cooperation using Price's Equation [4], which states that the change in the average character of a population is $dZ/dt = \mathrm{cov}(w, z)$. By inserting equation (S8) (w) into the covariance equation, we obtain:

$$\begin{aligned} dZ/dt &= \mathrm{var}(z)\,(r_c - r_d) - \mathrm{var}(z)\,(k - a)X_c(C_{cc} - C_{cd}) - \mathrm{var}(z)\,kX_d(C_{cd} - C_{dd}) \\ &= \mathrm{var}(z)((r_c - r_d) - (k - a)X_cC_{cc} + kX_dC_{dd} + ((k - a)X_c - kX_d)C_{cd}) \end{aligned} \tag{S9}$$

Compare this result to Equation 8 in [5], which assumed discrete-space (but continuous time), and weak selection such that rare mutant cooperators M invade resident defectors R. For cost and benefit affecting survival, their selection direction is given by the inclusive fitness effect:

$$\Delta W_{IF} = b\Delta(q_{o/M} - q_{o/R}) - \Delta(C\text{-}B(q_{M/M} - q_{M/R})) \tag{S10}$$

Δ stands for the partial derivative operator with respect to helping investment at neutrality (small effect). b is the background fecundity, C is cost of helping, and B is the benefit. By equating the local densities $X_{ij}=X_jC_{ij}=q_{j/i}$, collecting terms by k and a, and changing the subscripts $M=c$ (mutant cooperators) and $R=d$ (resident defectors), Equation S9 becomes:

$$dZ/dt = \mathrm{var}(z)(\, r_M - r_R + a(q_{M/M} - q_{M/R}) - k(q_{x/M} - q_{x/R}) \tag{S11}$$

Further, given $q_{M/i} + q_{R/i} = q_{x/i}=1 - q_{o/i}$,

(S12) $$dZ/dt = \mathrm{var}(z)(r_M - r_R + a(q_{M/M} - q_{M/R}) + k(q_{o/M} - q_{o/R})$$

Finally, by equating costs $r_M - r_R = \Delta C$, $a = \Delta B$, and $k = b$, Equation S12 is identical to Equation S10. Thus, our approach generalizes and simplifies existing results by allowing for any selection strength and continuous space and time dynamics. The terminology of k instead of b also highlights the ecologically familiar competitive effect instead of the indirect effect of fecundity which in the end translates into competition [5].

Connection between Clustering Coefficients and Relatedness

To connect clustering coefficients with the more familiar relatedness metrics in inclusive fitness theory [6], we reformulate the main model as a one-locus population genetics model that accounts for interaction effects. The fitness of an individual u is the sum of its intrinsic growth probability, expected interaction effect received from each of all other individuals v ($a_{v\to u}$), and all expected costs incured upon encountering individual v in a small temporal increment Δt, taken over an ensemble of realizations of the same configuration:

(S13) $$w_u = r_u + \sum_{v \neq u} a_{v->u}$$

Ley z_u be the morph or character value of individual u. The expected changes in the number of morph i individuals (N_i) and of all individuals (N) are then

(S14) $$\frac{\Delta N_i}{\Delta t} = \sum_{u=1}^{N} z_u w_u \ , \ \frac{\Delta N}{\Delta t} = \sum_{u=1}^{N} w_u$$

The change in the morph i frequency p_i is:

(S15) $$\frac{\Delta p_i}{\Delta t} = \frac{\Delta}{\Delta t}\left(\frac{N_i}{N}\right) = \frac{\Delta N_i}{(N+\Delta N)\Delta t} - \frac{N_i \Delta N}{N(N+\Delta N)\Delta t}$$

Using equations S13-15, we obtain:

$$\frac{\Delta p_i}{\Delta t} = \frac{1}{N+\Delta N}\sum_{u=1}^{N} z_u\left(r_u + \sum_{v\neq u} a_{v\to u}\right) - \frac{p_i}{N+\Delta N}\sum_{u=1}^{N}\left(r_u + \sum_{v\neq u} a_{v\to u}\right) \quad \text{(S16)}$$

$$= \frac{1}{N+\Delta N}\sum_{u=1}^{N}(z_u - p_i)\left(r_u + \sum_{v\neq u} a_{v\to u}\right)$$

We can convert equation S16 to an inclusive fitness formulation. We switch the index of the interaction effect between pairs from $a_{v\to u}$ to $a_{u\to v}$, such that the effect direction switches from what the focal individual u receives to what u initiates.

$$\frac{\Delta p_i}{\Delta t} = \frac{1}{N+\Delta N}\sum_{u=1}^{N}(z_u - p_i)\left(r_u + \sum_{v\neq u}\frac{z_v - p_i}{z_u - p_i} a_{u\to v}\right) \quad \text{(S17)}$$

The term $(z_v - p_i)/(z_u - p_i)$ is the correlation form of relatedness, defined for every pair of individuals [7].

We take the limit of $\Delta t \to 0$ in equation (S17), where $N+\Delta N \approx N$, to arrive at a continuous-time analogue for the change in morph frequency. For a 2-morph population (with cooperators c and defectors d), we associate cost, benefit, and relatedness terms with morph, such that the indices now refer to the morph instead of the individual. We now assume that all individuals of a morph i provide the same effect ($a_{u\to\bullet}|_{u\in i} = a_{i\bullet}$) to each interacting neighbour without discrimination in the small time interval dt. Specifically, $a_{c\bullet} = a - k$ and $a_{d\bullet} = -k$. Subsituting $X_{ij} = C_{ij} X_j$, from equation (S17) we get:

$$\frac{dp}{dt} = \frac{1}{N}\sum^{N_c}(1-p)(r_c + X_c R_c a_{c\bullet}) + \frac{1}{N}\sum^{N_d}(0-p)(r_d + X_d R_d a_{d\bullet}) \quad \text{(S18)}$$

$$= p(1-p)(r_c + (C_{cc}X_c + C_{cd}X_d)R_c a_{c\bullet} - r_d - (C_{dc}X_c + C_{dd}X_d)R_d a_{d\bullet})$$

where the relatedness terms are:

(S19) $$R_c = \frac{\frac{C_{cc}X_c}{C_{cc}X_c+C_{cd}X_d} - p}{1-p}, R_d = \frac{\frac{C_{dc}X_c}{C_{dc}X_c+C_{dd}X_d} - p}{-p}$$

The $\Sigma_{v\neq u}$... summations from equation (S17) are replaced in equation (S18) by $X_{i\bullet}$ (average total local density around morph *i*) because both represent the average sum of effects on neighbours by one individual. $(1/N)\Sigma^{Ni}$... is replaced by p_i times the average of the term in the summation. There are two relatedness terms, because in this model both cooperators and defectors initiate interactions. The fractions involving clustering coefficients in equation (S19) are the fractions of interaction partners that are cooperators. Relatedness thus effectively normalize clustering effects so that there is no *a priori* dependence of clustering on global densities. The difference between the two relatedness terms and their multiplicands (the total interaction effects), plus the difference between the intrinsic growth rates, which are the last bracketed expression in equation (S18), is analogous to the inclusive fitness effect [8,9].

R_c is equivalent to nearest-neighbour relatedness under neutrality (with $q_{x/M} = q_{x/R} = q_{x/x}$), which can be expressed as $(q_{M/M} - q_{M/R})/q_{x/x}$ [10]. We can rearrange equations (S18), plug in R_c in terms of *q*, and substitute the indices *M*=*c* and *R*=*d* to resemble the the inclusive fitness result of equation (S10):

(S20) $$dp/dt = p(1-p)((r_M - r_R + (q_{M/M} - q_{M/R})a_{c\bullet} + q_{x/R}R_d a_{d\bullet})$$

Again, $a_{c\bullet}=a - k$ and $a_{d\bullet}=-k$, so equation (S20) can be rearranged as:

(S21) $$dp/dt = p(1-p)((r_M - r_R + a(q_{M/M} - q_{M/R}) + k(q_{M/R} - q_{M/M} + q_{x/R}R_d))$$

Comparing equations (S10) and (S21), we can deduce that R_d is equivalent to $(q_{R/R} - q_{R/M})/q_{x/x}$, which is exactly nearest-neighbour relatedness with the *R* and *M* indices switched. This equation for R_d also matches the definition we independently derived in equation (S19). Note that R_d includes the defector clustering term ($C_{dd} = q_{R/R}$), which is in the inclusive fitness formulation

(equations (S10) or (S21)) embedded in the term associated with k or b, the competition or "ecological cost" term [5]. Thus, defector clustering already appeared distinctively from cooperator clustering in previous literature, but it was not isolated on its own.

Weak Selection and Large Populations

If we assume that selection is weak, a neutral dimorphic finite population consisting of any two morphs (such as cooperators and defectors) will take on any possible configuration given the population size N (1 cooperators vs. N-1 defectors, 2 cooperators vs. N-2 defectors, etc.). In a mutation-selection process (with any mutation rate), it is known that selection is expected to favour the morph that, averaged over all possible configurations, has a frequency above ½ [11]. Simple Monte-Carlo simulations (1E5 samples for each population size) shows that as the population size increases, the portion of times that the frequency of one morph is different from ½ (defined as less than 0.475 or more than 0.525, or outside the 95% confidence interval of 1/2) decreases. At around a population size of 2000, the neutral dimorphic population drifts to a morph frequency of 1/2 almost all of the time. Thus, for populations bigger than 2000, it is sufficient to infer the direction of selection from investigating how morph frequency changes around 1/2; equivalently, this is when $X_c = X_d = X_i$ in a population of cooperators and defectors.

Table 1. Simulation of fraction of times that the morph frequency in a neutral dimorphic population is different from ½ (<0.475 or >0.525) as a function of population size (N). For each population size, 1E5 random configurations were used to obtain the fractions.

population size (N)	2	4	8	16	32	64	128	256	512	1024	2048	4096	8192	16384
fraction of morph frequency different from 1/2	0.500	0.625	0.727	0.803	0.860	0.708	0.535	0.417	0.269	0.111	0.0233	0.00123	1.00E-5	0

Simulation

The simulation model, implemented in Matlab, is a discretization of the local interaction model of equation (1). At each time step, cooperation and competition from neighbours in each of the 36x36 patches affects birth probabilities, while death and movement events occur at constant probabilities. We set r_d=0.1, r_d-r_c=0.01 (cost), a=0.05 (benefit), and k=0.1 (competition), each being an expected change in density per observation time T in each cell of the habitat. Growth was implemented as the sum of binomial random birth and death events for all individuals at each update. Cost, benefit, and competition affected the birth rate, and the death rate was set at 0.1. The cost-free local movement rate was set at 0.2, 0.3, or 0.6 crossings between patch per observation time. These movement rate variations effectively change the potential interaction scale and test the robustness of clustering effect measurements made at a fixed (patch) scale. The movement direction was random - unless the chosen direction was a boundary, in which case the individual stayed. The simulation process is illustrated in Fig. S3, and a time series is presented in Fig. S4. For movement rate of 0.3, the correlation between cooperator frequency and population density, and the correlation between cooperator clustering and defector clustering are plotted in Fig. S5. Similar results were obtained from movement rates of 0.2 and 0.6 (Fig. S6 & S7).

Figures

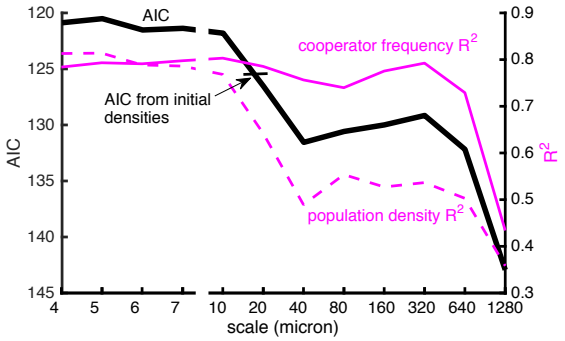

Figure S1. AIC (scale inverted with smallest values at the top) and portions of variance in cooperator frequency and total population density explained by clustering coefficients measured at different scales. The short horizontal line crossing the AIC curve represents the AIC obtained from using initial cooperator and defector densities as predictors.

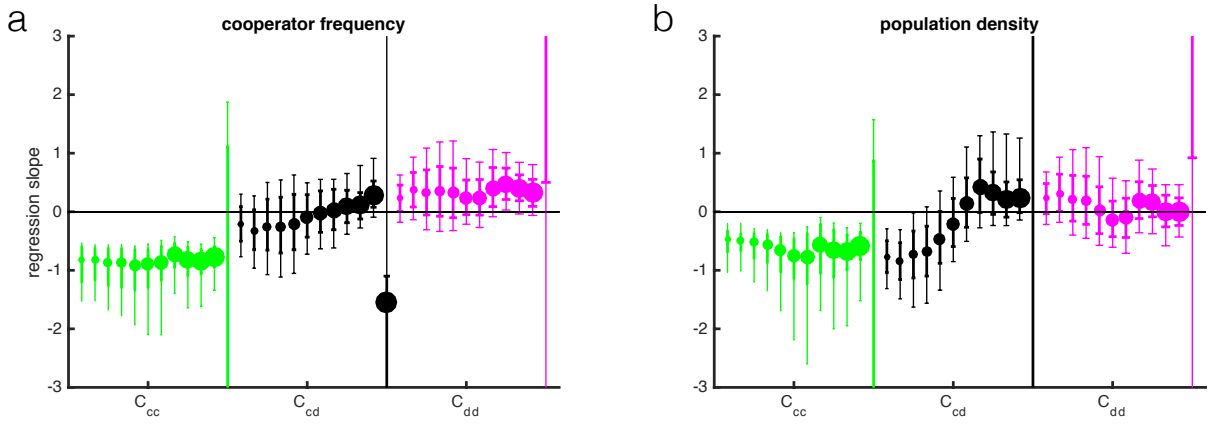

Figure S2. Bootstrapped linear regression slopes (effects) of standardized cooperator frequency on standardized clustering coefficients (C_{ij}/std(C_{ij})) measured at different spatial scales. C_{cc}, C_{cd}, C_{dd} are clustering coefficients among cooperators, between cooperators and defectors, and among defectors. The circles indicate the effect means, and the size of the circles indicate the scales in microns (4, 5, 6, 7, 10, 20, 40, 80, 160, 320, 640, 1280). The thick bars indicate the 25th and 75th percentiles, and the whiskers indicate the 2.5th and 97.5th percentiles of the effect estimates according to 10000 bootstrap resamples of the data.

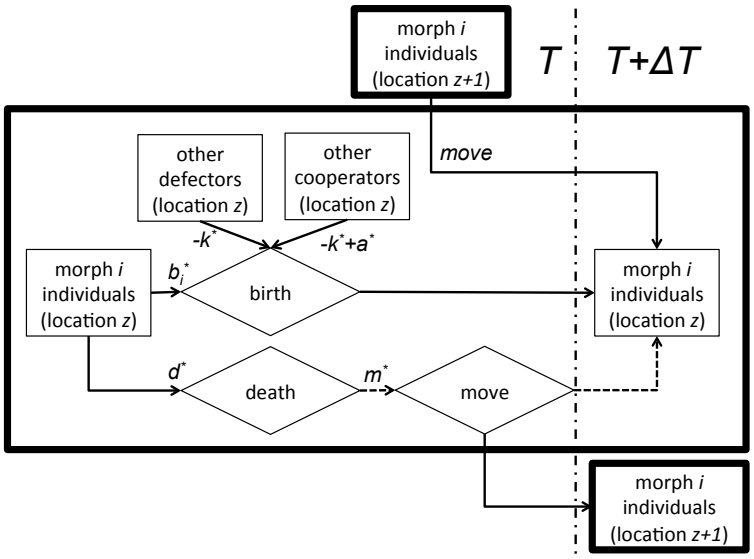

Figure S3. Simulation process chart for each location *z* in the 36x36 habitat. Thick boxes indicate distinct locations (*z*+1 is any neighbouring location of *z*). Thin boxes are state variables, and diamonds are events. Connectors flowing out are modifiers to the rates (binomial probabilities) that the events they point to occur. Rates are subscripted * to indicate that they are 1/100 of the model parameters as part of the discretization procedure. Solid connectors out of events indicate that the process continues if the events occured, whereas dashed connectors indicate the process continues if the events did not occur. Each update uses state variables from time *T* to project their values at *T+ΔT*.

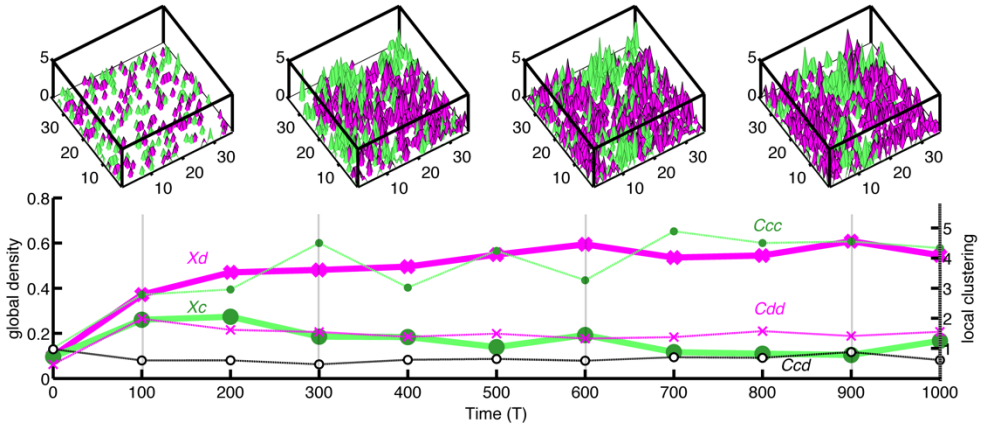

Figure S4. Four snapshots from a simulation time series with a movement rate of 0.3. Green (light) indicates locations where cooperator clusters dominate, and magenta (dark) indicates where defector clusters dominate. Global densities of cooperators (X_c) and defectors (X_d) are plotted as thick lines (scaled to the left axis), while local clustering coefficients (among cooperator clustering C_{cc}, between-morph clustering C_{cd}, and among defector clustering C_{dd}) are plotted as thin lines (scaled to the right axis).

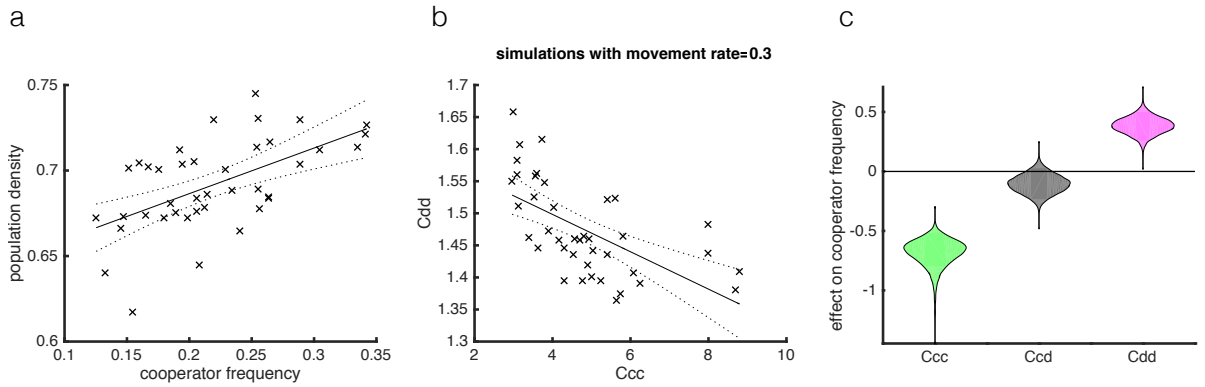

Figure S5. a, Cooperator frequencies versus total population densities in an individual-based simulation with movement rate=0.3 (n=40). The relationship was positive (slope=0.27, S.E.=0.063, $F_{2,38}$=19, p=1.2e-4, R^2=0.33). b, The relationship between cooperator and defector clustering was negative (slope=-0.0308, S.E.=0.0065, $F_{2,38}$=22, p=3.5e-5, R^2=0.37). The dotted lines represent the 95% confidence bounds. c, Bootstrapped linear regression slopes (effects) of cooperator frequency on standardized clustering coefficients (C_{ij}/std(C_{ij})) were obtained by measuring clustering within patch. The regression model predicts (R^2) 0.86 of cooperator frequency outcomes, and 0.62 of population density outcomes.

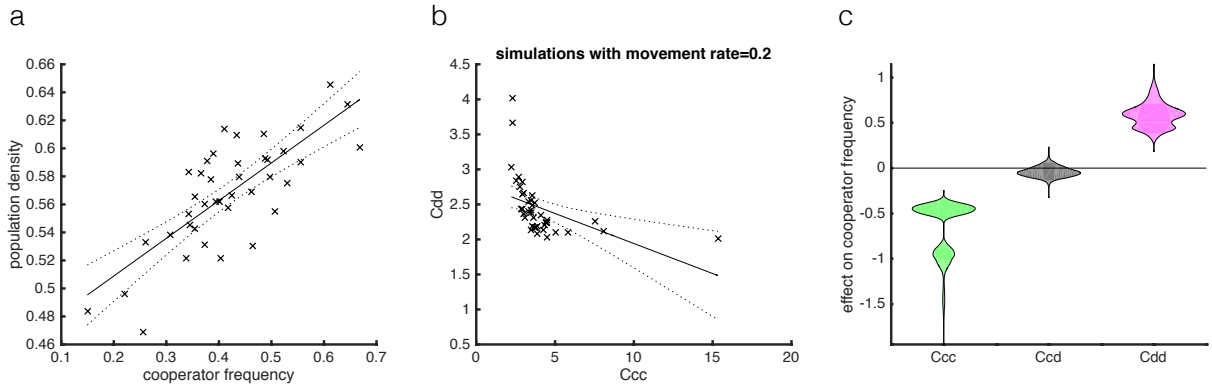

Figure S6. a, Cooperator frequencies versus total population densities in an individual-based simulation with movement rate =0.2 (n=40). The relationship was positive (slope=0.27, S.E.=0.036, $F_{2,38}$=55, p=7.0e-9, R^2=0.59). b, The relationship between cooperator and defector clustering was negative (slope=-0.086, S.E.=0.027, $F_{2,38}$=10, p=0.0030, R^2=0.21). The dotted lines represent the 95% confidence bounds. c, Bootstrapped linear regression slopes (effects) of cooperator frequency on standardized clustering coefficients (C_{ij}/std(C_{ij})) were obtained by measuring clustering within patch. The regression model predicts (R^2) 0.86 of cooperator frequency outcomes, and 0.72 of population density outcomes. In all bootstraps, cooperator clustering is negatively associated with cooperator frequency, while defector clustering is positively associated.

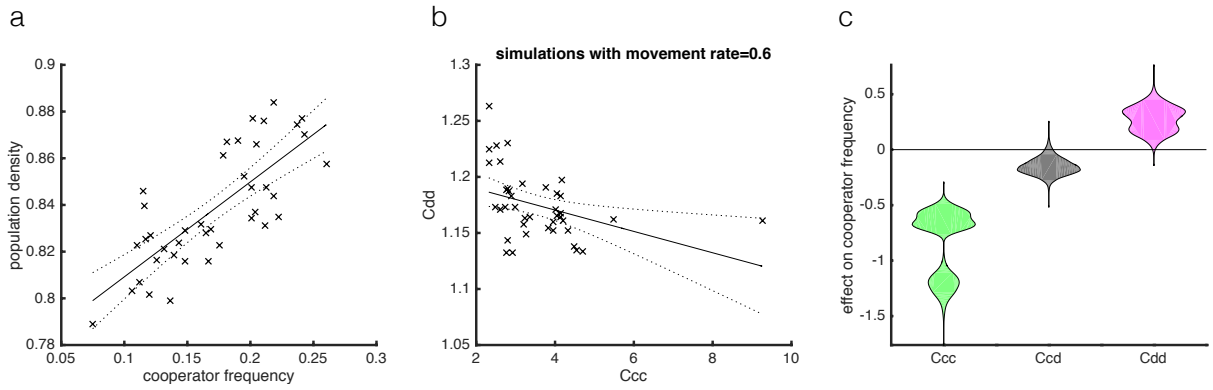

Figure S7. a, Cooperator frequencies versus total population densities in an individual-based simulation with movement rate =0.6 (n=40). The relationship was positive (slope=0.41, S.E.=0.056, $F_{2,38}$=53, p=9.9e-9, R^2=0.58). b, The relationship between cooperator and defector clustering was negative (slope=-0.0095, S.E.=0.0036, $F_{2,38}$=6.9, p=0.013, R^2=0.15). The dotted lines represent the 95% confidence bounds. c, Bootstrapped linear regression slopes (effects) of cooperator frequency on standardized clustering coefficients (C_{ij}/std(C_{ij})) were obtained by measuring clustering within patch. The The regression model predicts (R^2) 0.81 of cooperator frequency outcomes, and 0.66 of population density outcomes. In all bootstraps, cooperator clustering is negatively associated with cooperator frequency. Defector clustering is positively associated with cooperator frequency (p=0.0032, one-sided test).

Supplementary References

1. Tekwa, E. W., Nguyen, D., Juncker, D., Loreau, M. & Gonzalez, A. 2015 Patchiness in a microhabitat chip affects evolutionary dynamics of bacterial cooperation. *Lab Chip* **15**, 3723–3729. (doi:10.1039/C5LC00576K)
2. Bolker, B. M. & Pacala, S. W. 1999 Spatial Moment Equations for Plant Competition: Understanding Spatial Strategies and the Advantages of Short Dispersal. *Am. Nat.* **153**, 575–602. (doi:10.1086/303199)
3. Law, R. & Dieckmann, U. 2000 A dynamcial system for neighborhoods in plant communities. *Ecology* **81**, 2137–2148. (doi:10.1890/0012-9658(2000)081[2137:ADSFNI]2.0.CO;2)
4. Price, G. R. 1970 Selection and covariance. *Nature* **227**, 520–521.
5. Lion, S. & Gandon, S. 2010 Life history, habitat saturation and the evolution of fecundity and survival altruism. *Evolution* **64**, 1594–1606. (doi:10.1111/j.1558-5646.2009.00933.x)
6. Hamilton, W. D. 1964 The genetical evolution of social behaviour. I. *J. Theor. Biol.* **7**, 1–16.
7. Grafen, A. 2006 Optimization of inclusive fitness. *J. Theor. Biol.* **238**, 541–563.
8. Taylor, P. D. & Frank, S. A. 1996 How to make a kin selection model. *J. Theor. Biol.* **180**, 27–37.
9. Rousset, F. & Billiard, S. 2000 A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. *J. Evol. Biol.* **13**, 814–825. (doi:10.1046/j.1420-9101.2000.00219.x)
10. Lion, S. 2009 Relatedness in spatially structured populations with empty sites: An approach based on spatial moment equations. *J. Theor. Biol.* **260**, 121–131. (doi:10.1016/j.jtbi.2009.05.035)
11. Tarnita, C. E., Ohtsuki, H., Antal, T., Fu, F. & Nowak, M. A. 2009 Strategy selection in structured populations. *J. Theor. Biol.* **259**, 570–81. (doi:10.1016/j.jtbi.2009.03.035)