Predicting global dynamics of spatial microbial communities from local interaction rules

Simon van Vliet[a,1], Christoph Hauert[b], Martin Ackermann[c,d], and Alma Dal Co[c,d,e,1]

[a]Department of Zoology; University of British Columbia, Vancouver, BC, Canada
[b]Department of Mathematics; University of British Columbia, Vancouver, BC, Canada
[c]Department of Environmental Microbiology; Eawag, Duebendorf, Switzerland
[d]Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland
[e]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
[1]To whom correspondence should be addressed. E-mail: vanvliet@zoology.ubc.ca; alma.dalco@gmail.com

Abstract

Interactions between cells drive biological processes across all of life, from microbes in the environment to cells in multicellular organisms. Interactions often arise in spatially structured settings, where cells mostly interact with their neighbors. A central question is how the properties of biological systems emerge from local interactions. This question is very relevant in the context of microbial communities, such as biofilms, where cells live close by in space and are connected via a dense network of biochemical interactions. To understand and control the functioning of these communities, it is essential to uncover how community-level properties, such as the community composition, spatial arrangement, and growth rate, arise from these interactions. Here, we develop a mathematical framework that can predict community-level properties from the molecular mechanisms underlying the cell-cell interactions for systems consisting of two cell types. Our predictions match quantitative measurements from an experimental cross-feeding community. For these cross-feeding communities, the community growth rate is reduced when cells interact only with few neighbors; as a result, some communities can co-exist in a well-mixed system, but not in a spatial one. In general, our framework shows that key molecular parameters underlying the cell-cell interactions (e.g. the uptake and leakage rates of molecules) determine community level properties. Our framework can be extended to a variety of systems of two interacting cell types, within and beyond the microbial world, and contributes to our understanding of how biological functions arise from interactions between single cells.

Introduction

Biological interactions pervade all of life. Interactions at lower levels of organization can confer new functionality at higher levels. For example, interactions between different cell types determine the functioning of organs and tissues in multicellular organisms and interactions between different species determine the processes an ecosystem performs [1,2]. In natural systems, interactions often arise in spatially structured settings, where individual entities interact mostly with others that are close by in space [3]. When interactions are local, the spatial organization of the different entities defines their network of interaction. A central question is how the properties of biological systems emerge from this network of interactions. This has primarily been studied

in the context of multicellular organisms, however it is also particularly relevant in the context of microbial communities [4, 5].

Microbial communities perform essential processes on our planet and often consist of different interacting cell types [6, 7]. Many of these communities form spatially structured biofilms, where cells are embedded in an extracellular polymeric matrix that limits their movement [8]. In these communities, cells modify their local environment by secreting and taking up biochemical compounds. In this way, neighboring cells can influence each others' growth, survival, and metabolic activity [9–16]. Most interactions are mediated by diffusible molecules and their strength decays with the distance between two cells [17–19]. As a result, cells only interact within a finite distance, and the spatial organization of the microbial system determines which interactions are realized.

To predict and control the functioning of microbial systems, we need to uncover how cells interact in space and how these interactions determine community-level properties, such as the community composition, spatial arrangement, and growth rate. Recent studies have made progress in this direction by characterizing the spatial arrangement of cells [10, 20, 21], the range over which cells interact [10, 11, 22], and how the spatial arrangement of cells affects community properties [5, 19, 20, 23–25]. Despite these recent advances, we do not understand well how local interaction rules scale up to determine community-level properties.

Connecting processes at different scales presents one of the biggest challenges in biology. Here, we develop a mathematical framework that connects community-level properties to the local interaction rules between cells. This work conceptually complements our previous work were we measured these local interaction rules in a microbial community and derived them from the underlying molecular mechanisms [11].

Our framework takes as input the local interaction rules, set by key molecular parameters, and predicts community-level properties, such as the equilibrium frequency of cell types, their degree of spatial clustering, and the community growth rate. We validate our model's predictions on experimental observations of a synthetic microbial community. Our results suggests that, at least for simple systems, it is possible to scale up from molecular-level properties, to individual-level properties, to community-level properties. Properties at each level can be predicted from a few key quantities of the level below (Fig. 1). These findings further our understanding of how scales are connected in biological systems.

Results

We present a mathematical framework that predicts community-level properties from local interaction rules for a variety of systems consisting of two interacting cell types. First we apply it to microbial cross-feeding communities and compare its predictions to experimental observations. Subsequently, we derive general predictions for other systems.

To address our first goal, we build on previous work were we measured the local interaction rules for a cross-feeding community [11]. This community consists of two bacterial cell types that exchange essential amino-acids: one cell type (ΔP) is unable to produce proline, while the other (ΔT) is unable to produce tryptophan. Amino acids are passively released into the extracellular environment and cells in this community can thus grow by exchanging proline and tryptophan [17, 26–29]. We found that cells receive amino acids from partner cells residing within a finite interaction range; the growth rate of a cell increases linearly with the frequency of the partner type within this interaction range [11]. The local interaction rules can be fully described by the interaction range and the maximum growth rate that a cells achieves when all its neighbors are of the partner type (Fig. 1B).

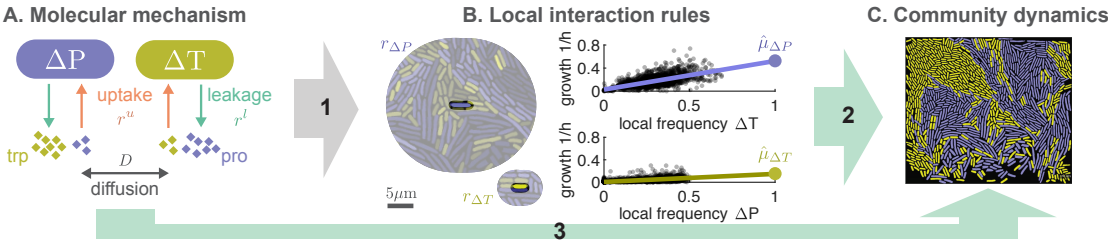

Fig. 1: A mathematical framework to scale up from molecular-level properties, to individual-level properties, to community-level properties. We previously measured the local-interaction rules for a cross-feeding community (B) and showed that these can be derived (arrow 1) from the molecular mechanisms of the interaction (A). Here we developed a mathematical model that can predict community-level properties (C) either from measured local rules (arrow 2) or directly from the underlying molecular mechanisms (arrow 3). (A) The community consists of two types of *Escherichia coli*: ΔP is unable to produce the amino acid proline and ΔT is unable to produce the amino acid tryptophan. Cells exchange amino acids with the environment through active uptake (with rate r^u) and passive leakage (with rate r^l). Amino acids are exchanged between cells through diffusion in the environment (with rate D). All rates differ between the two amino acids. (B) Local interaction rules can be fully described by two fundamental quantities: the size of the interaction neighborhood (r, left); and the growth function of a cell (charecterized by $\hat{\mu}$, right). Each dot corresponds to the measured growth rate of a single cell, $n = 2610$ for ΔP and $n = 2162$ for ΔT, the line shows the result of a linear regression, data reproduced from [11]. (C) Our framework makes predictions for the community-level properties, such as the equilibrium frequency of the two types, their spatial arrangement, and growth rate.

We also showed that the local interaction rules can be derived from key biophysical parameters of the underlying amino acid exchange (arrow 1 in Fig. 1). The interaction range primarily depends on the cell density and on the ratio between the uptake and diffusion rate of the molecules (Methods, Eq. 4). The maximum growth rate primarily depends on the leakage rate of the molecules (Eq. 6). Uptake, leakage, and diffusion rates are molecule specific; as a result the interaction range and maximum growth rate will generally differ between cell types, as we also observed for our specific community (Fig. 1B). Our goal here is to use these molecular parameters, which determine the local interaction rules, to predict community-level properties (arrow 3 in Fig. 1).

A mathematical framework for cross-feeding communities

Our framework builds on previous work using evolutionary graph theory [30–32]. We consider a spatial system of two cell types, A and B, that exchange essential cellular building blocks. We assume that the total number of cells is constant and that each cell interacts with a fixed number of neighbors. In contrast to previous work, we allow the two cell types to interact with a different number of neighbors, as suggested by our experimental observations [11]. We assume that type A interacts with r_A neighbors and that type B interacts with r_B neighbors. We call this quantity the *neighborhood size* and set $r_B \geq r_A$ (Fig. 2B). To allow for different neighborhood sizes, we need to use directed graphs, instead of the undirected graphs used in previous work. As a result, several simplifications no longer hold and we discuss in the SI the required extensions to previous work.

Each cell (of type A or B) interacts with all (r_A or r_B) cells in its interaction neighborhood and these interactions determine its growth rate. We assume that the growth rate of a cell increases linearly with the frequency of the partner type within its interaction neighborhood

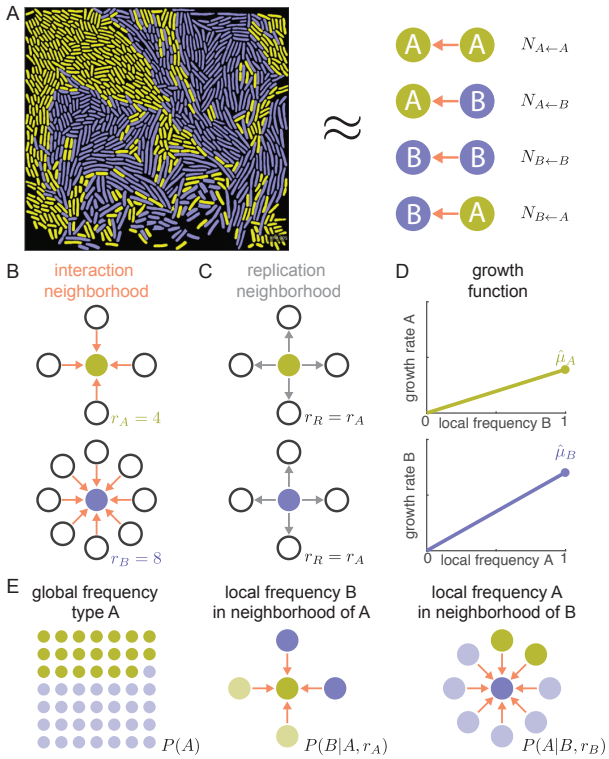

Fig. 2: Pair approximation allows for a quantitative description of spatial systems. (A) Pair approximation assumes that a spatial system (left side) can be fully described by tracking the number of all pairwise links (right side). $N_{X\leftarrow Y}$ indicates that focal cell X interacts with its neighbor Y. (B-D) A system is fully described by the local rules, as specified by: (B) the interaction neighborhoods of both types, characterized by the neighborhood sizes r_A and r_B; (C) the replication neighborhood, assumed to be identical to the smallest interaction neighborhood; and (D) the growth functions of both types, characterized by the maximum growth rates $\hat{\mu}_A$ and $\hat{\mu}_B$. (E) Pair approximation has three dynamical variables that describe the global composition, $P(A)$, and local composition $P(B|A, r_A)$ and $P(A|B, r_B)$ of the system.

(Fig. 2D), as observed in the experimental communities [11]. Each type achieves its *maximum growth rate*, $\hat{\mu}$, when it is completely surrounded by the partner type, and its value is different for the two types (Fig. 2D).

When a cell divides, it replaces a random neighbor within the replication neighborhood. We assume that the replication neighborhood is identical to the smaller interaction neighborhood ($r_R = r_A$, Fig. 2C). This assumption has its limitations: in reality neighboring cells are not replaced, but rather pushed aside. Moreover, the replication neighborhood could be different from the interaction neighborhood. However, this assumption does capture one of the most essential features of real biological systems: that cells place their offspring close by in space and thus become surrounded by their own type (see SI for a more detailed discussion).

We implemented our model in two complementary ways: we used a cellular automaton to simulate the dynamics numerically and we used pair approximation to derive analytical predictions. Pair approximation assumes that the dynamics of a spatial system can be described by only specifying how often the different cell types are found in each others neighborhoods (Fig.

2A) [30, 33, 34]. In our case, this means that the outcome of an interaction between a focal individual and one of its neighbors depends only on the identity of these two cells and not on the wider context in which these two cells are found. This approximation allows us to describe the dynamics of a spatial community (Fig 2A left) by only tracking how often cells of the different types are found within each others interaction neighborhood, i.e. by tracking the number of A←A, A←B, B←A, and B←B links (Fig 2A right). In our convention, X←Y means the focal cell X interacts with Y, i.e. that the focal cell X receives building blocks from Y. These links are directional: because the two types have different interaction ranges, cell X might receive building block from cell Y, even when the inverse is not true. This is the primary difference with previous work, where interactions were assumed to be symmetric (and links undirected) [30].

With these assumptions, we obtain the dynamical equations for the number of pairwise links in time (see SI). Because the number of cells is constant, we only have three independent variables (e.g. the number of A←A, A←B, B←A links) and we can express the fourth one (the number of B←B links) as a function of the other three. Instead of tracking the number of links directly, it is convenient to change variables and track the following dynamical variables: $P(A)$, the global frequency of A, $P(B|A, r_A)$, the local frequency of type B within the interaction neighborhood of type A, and $P(A|B, r_B)$, the local frequency of type A within the interaction neighborhood of type B (Fig. 2E). The first quantity describes the global composition of the community, while the last two quantities describe the average local composition of a cells' neighborhood. The average growth rate of a cell in a spatial system can be directly calculated when we know the average local composition, i.e. the average frequency of the partner type within its interaction neighborhood.

Predicted community properties at steady state

Pair approximation allows us to predict global properties of the community from the local interaction rules between cells. We input the neighborhood size of each type (r_A and r_B) and the growth functions characterized by the maximum growth rate obtained when a cell is completely surrounded by the other type ($\hat{\mu}_A$ and $\hat{\mu}_B$, Fig. 2 B and D). As output we obtain a system of dynamical equations describing the global, $P(A)$, and local, $P(B|A, r_A)$ and $P(A|B, r_B)$, composition of the system (Fig. 2E). By solving the dynamical equation for steady state, we can find analytical expressions for the global and local compositions at equilibrium (see SI).

Pair approximation predicts that the global composition of the community reaches a steady state in which the frequency of type A is given by:

$$P(A) = \frac{\hat{\mu}_A \cdot \frac{r_A - 2}{r_A} + \left(\frac{\hat{\mu}_A}{r_A} - \frac{\hat{\mu}_B}{r_B}\right)}{\hat{\mu}_A \cdot \frac{r_A - 2}{r_A} + \hat{\mu}_B \cdot \frac{r_B - 2}{r_B}} \tag{1}$$

When the neighborhood size is large ($r_A, r_B \gg 1$), this equation simplifies to the predicted frequency in a well-mixed system (see SI):

$$P(A)_{WM} = \frac{\hat{\mu}_A}{\hat{\mu}_A + \hat{\mu}_B} \tag{2}$$

The same global community composition can correspond to many different spatial arrangements of the two cell types. For example, the two cell types could be highly mixed in space or completely segregated. The spatial arrangement matters because cells interact only locally: the growth of a cell depends on the local frequency of the partner type, $P(B|A, r_A)$ and $P(A|B, r_B)$, which can be different from the global frequency, $P(B)$ and $P(A)$, when the cell types are clustered in space. A relevant quantity to describe a spatial system is thus the ratio between the

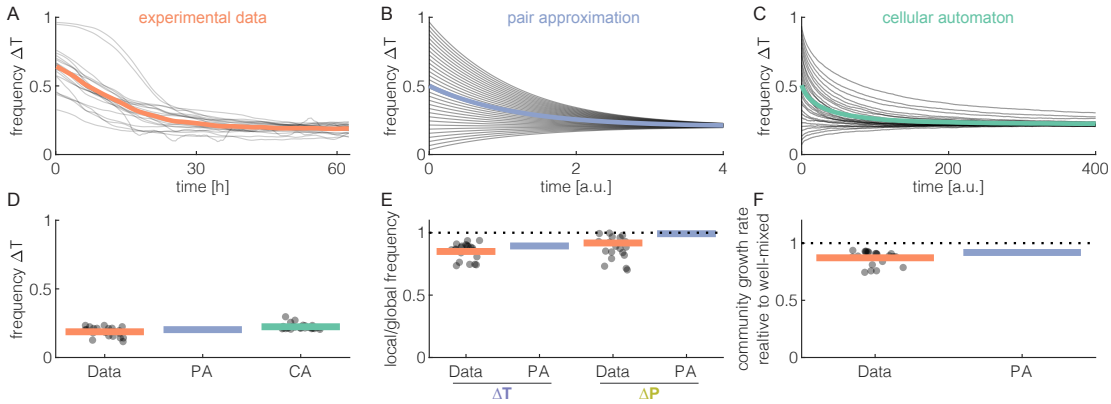

Fig. 3: Our mathematical framework can explain experimentally observed community properties. The model parameters for the experimental community were calculated from the biophysical model (Table S2). (A) Experimental communities approach a stable frequency of ΔT. Individual communities (thin lines, n=22) and their average value (thick line) are shown. (B) Pair approximation predicts a unique stable equilibrium. The dynamical equations (SI Eq. 13-15) were solved numerically starting from different initial frequencies. Time is in arbitrary units (a.u.). (C) Cellular automaton simulations also reach a unique stable equilibrium. (D) The observed equilibrium frequency of ΔT is consistent with the model predictions. Data: $P(A)$=0.19 (95% confidence interval (CI): 0.17–0.20), Pair-approximation: $P(A)$=0.20, and cellular automaton: $P(A)$=0.22 (CI: 0.22–0.23). The difference between the model prediction and data is less than 8%. (E) The frequency of the partner cell type within the interaction neighborhood (local frequency) is lower that the global frequency because cells are surrounded by their own offspring. Dots show measurements for 21 replicate communities, bar shows mean value. Pair approximation predicts (Eq. 3) a decrease in frequency of 0.99 for ΔP and 0.89 for ΔT, the experimental values are 0.87 (CI: 0.85–0.90) and 0.85 (CI: 0.82–0.88), respectively. The difference between the model prediction and data is less than 13%. (F) The average growth rate of the community is reduced due to cell clustering. An *in-silico* (see Methods) shows that the growth rate in clustered communities, with experimentally observed spatial arrangements, is reduced by a factor of 0.87 (CI: 0.85–0.90) compared to randomized communities, where cell clusters were disrupted (Data reproduced from [11]). Pair approximation predicts a decrease by a factor of 0.94 (SI Eq. 22-23), the difference between the model prediction and data is less than 6%.

local and global frequency of the partner type, and pair approximation predicts that at steady state:

$$\frac{P(B|A, r_A)}{P(B)} = \frac{r_A - 2}{r_A - 1}, \quad \frac{P(A|B, r_B)}{P(A)} = \frac{r_B - 2}{r_B - 1} \tag{3}$$

where $P(B) = 1 - P(A)$. This equation shows that the frequency of the partner type in the interaction neighborhood is lower than one would expect from the global composition of the community. This happens because cells place offspring close to themselves when they divide. When the neighborhood size is small (e.g., $r = 3$) this reduction in the local frequency is as much as 50%, but when it is large ($r \gg 1$), the local frequency approaches the global frequency.

Predicting global properties of an experimental microbial community

We validate our mathematical framework by comparing its predictions to experimental observations of our cross-feeding bacterial community. We parameterize our model directly from the biophysical parameters describing the amino acid exchange (arrow 3 in Fig. 1, Table S2). The

model can be fully parameterized by specifying the size of the interaction neighborhood (r), which mostly depends on the uptake and diffusion rates (Eq. 4), and the maximum growth rate of a cell ($\hat{\mu}$), which mostly depends on the leakage rate (Eq. 6, see Methods). Alternatively, we can parameterize the model directly from the measured interaction range and measured maximum growth rate (arrow 2 in Fig. 1), the predictions obtained from both parameterization match closely (Table S3).

Our model recapitulates the experimentally observed community-level dynamics. In our experiments, the frequency of the ΔT type converges over time to a stable frequency across 22 communities (Fig. 3A). Our model predicts similar dynamics, as we demonstrate by numerically solving the dynamical equations given by pair approximation and by simulating the system using a cellular automaton. Both pair approximation (Fig. 3B) and simulations (Fig. 3C) show that the community reaches a stable equilibrium. From Eq. 1 we find a predicted equilibrium frequency of ΔT of 0.20, which matches well with both the experimentally observed value of 0.19 and cellular automaton simulations value of 0.22 (Fig. 3D).

Cells in microbial communities tend to be surrounded by cells of their own type because when a cell divides the two new cells remain close in space. Therefore, the local frequency of the partner type found in the interaction neighborhood of a focal cell is typically lower than the global frequency in the whole community. Our model predicts (Eq. 3) that the local frequency of the partner type is reduced by a factor of 0.99 for ΔP and 0.89 for ΔT, compared to the global frequency. We measured this reduction in the frequency for the experimental communities and found values of 0.87 for ΔP and 0.85 for ΔT. Qualitatively the experimental observations thus match our model predictions: cells tend to be surrounded by their own type, thus reducing the local partner frequency.

For cross-feeding communities, cells in a spatially structured community grow slower than cells in an equivalent well-mixed community, because cells become surrounded by their own type. Our model predicts that spatial structure decreases the average growth rate in our experimental community by a factor of 0.92 (SI Eq. 22-23). Experimentally we cannot directly compare the growth in spatially structured communities with those in well-mixed communities. However, we previously used an individual-based biophysical model to perform an equivalent *in-silico* experiment: we compared the predicted growth rate for *clustered communities*, in which the spatial arrangement of cells was based on experimental measurements, with *randomized communities*, in which the spatial arrangement of cells was randomized [11]. We found that the growth rate in the clustered communities was reduced by a factor of 0.87 compared to the randomized communities, a value that is comparable with the value of 0.94 predicted by pair approximation.

Global predictions from local interactions

We now investigate what general predictions we can make for cross-feeding communities if we know the local interaction rules. Our model predicts that the equilibrium composition of the community is mostly set by the maximum growth rates of the two types. In general, the type with highest maximum growth rate constitutes the majority (Fig. 4A). When the neighborhood size is large ($r \gg 1$), the equilibrium frequency approaches that of a well-mixed system, while if the neighborhood size is small ($r = 3$), the equilibrium frequency is shifted more to the type that grows faster (Fig. 4A).

When both cell types have the same maximum growth rate, the ratio of neighborhood sizes ($\frac{r_B}{r_A}$) does not affect the equilibrium state of the community (Fig. 4B, black line). When the the two cell types have different growth rates, increasing the neighborhood size of even a single type moves the equilibrium frequency closer to the expected frequency in a well-mixed system (Fig. 4B, colored lines).

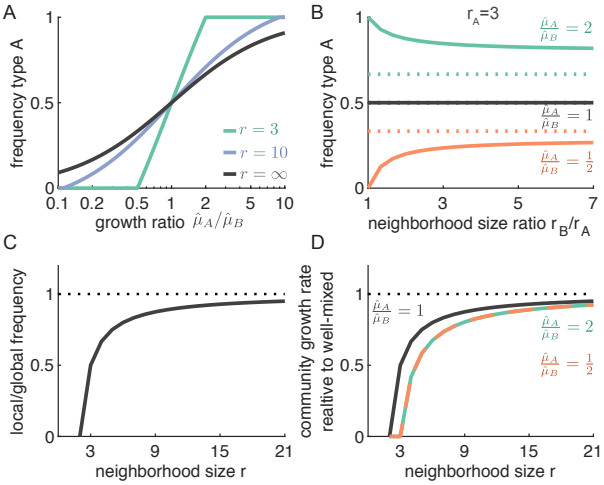

Fig. 4: Community level properties depend on local interaction rules. (A) The global composition of the community (global frequency of ΔT, given by Eq. 1) primarily depends on the ratio of maximum growth rates of the two types. When the neighborhood size r is large (purple line), the equilibrium frequency approaches that of a well-mixed system (black line, given by Eq. 2). When the neighborhood size r is small (green line) the advantage of the type with the higher growth rate reaches is increased. (B) The neighborhood size has only a minor effect on the global composition of the the community. r_A is held constant at 3, while r_B is increased. The equilibrium frequency in the spatial system (solid lines) moves closer to the equilibrium frequency in a well-mixed system (dotted lines) when r_B becomes much larger than r_A. (C) Neighborhood size strongly affects the local composition of the community. The local frequency of the partner cells becomes much lower than the global frequency when the neighborhood size is small. (D) Neighborhood size strongly affects community growth rate. Because of the reduction in the local frequency of the partner type, the average growth rate of cells goes down when the neighborhood size becomes small.

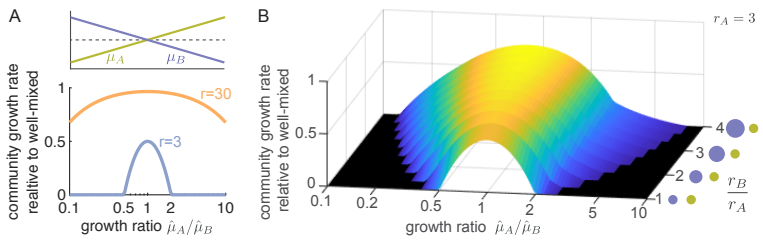

Fig. 5: Small interaction neighborhoods can lead to community collapse. (A) When the neighborhood size is small ($r = 3$), asymmetric communities, where the cell types have different maximum growth rates, grow poorly. When the asymmetry is too large communities cannot grow in a spatially structured environment even though they could grow in a well-mixed one. When the neighborhood size is large ($r = 30$), the growth rate of the spatial community is close to that of the well-mixed one even when there is an asymmetry. (B) Increasing the neighborhood size of even one of the types can increase the growth rate of spatial communities and prevent collapse. r_A is held constant at 3, while r_B is increased.

The neighborhood size strongly alters the local composition of a cells' neighborhood; as the neighborhood size becomes smaller, the partner fraction in each cells' interaction neighborhood decreases compared to the global fraction(Fig. 4C). As a result the average growth rate of the

community decreases relative to an equivalent well-mixed community (Fig. 4D).

Small interaction neighborhoods can lead to community collapse

We have shown that having a small interaction neighborhood reduces the growth rate of mutualistic communities. Here we show that this growth reduction is more pronounced when the two cell types have different maximum growth rates. We call such communities *asymmetric*. Our model predicts that the more asymmetric a mutualistic community is, the more their growth is hindered by small neighborhood sizes, to the point that the community collapses when neighborhood sizes are very small (Fig. 5A). Increasing the neighborhood size of either one (Fig. 5B), or both (Fig. 5A) of the types can prevent the collapse of the community. This finding shows that there is a limit to the stability of mutualistic interactions in spatially structured communities: mutualistic types with different maximum growth rates that can stably coexist in well-mixed environments might not be able to survive in spatially structured environments. For cross-feeding communities, spatial structure can thus have a detrimental effect.

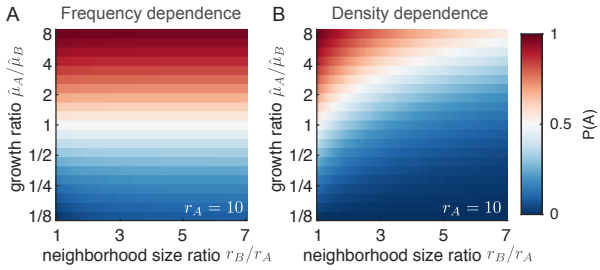

Fig. 6: Neighborhood size strongly affects the equilibrium frequency when the growth rate of a cell depends on the absolute number of partner cells. (A) When the growth rate depends on the *frequency* of the partner type within the interaction neighborhood (frequency dependence), the equilibrium frequency of type A is almost completely determined by the ratio of the maximum growth rates. (B) When the growth rate depends on the *number of cells* of the partner type within the interaction neighborhood (density dependence), the equilibrium frequency of type A depends both on the ratio of the maximum growth rates and on the ratio of the neighborhood sizes. The slow growing type can still dominate the community when it has a much larger interaction range (blue region, top right). In both panels $r_A = 10$.

Spatial effects for other types of communities

So far, we have focussed on cross-feeding systems were the growth rate increases with the frequency of the other type, however for other systems we expect different growth rate functions to be of relevance. As long as the growth rate depends linearly on the neighborhood composition, we can find an analytical solution for the steady state community properties (for non-linear functions the model can be solved numerically, see SI). The global composition of the community at steady state ($P(A)$) depends strongly on the chosen growth functions (SI, Eq. 28). However, the relation between the local and global composition of the community is always the same: i.e. for all growth functions we find that equation 3 holds (see SI).

Here we illustrate results for two different types of system. First, we consider a system where a type's growth rate increases with the number, rather than the frequency, of partner type within the interaction neighborhood (Fig. 6). In this case, the type with the highest product of $\hat{\mu} \cdot r$

reaches the higher global frequency (see SI). Therefore, the global composition of the community depends strongly both on the maximum growth rate and on the neighborhood size (Fig. 6).

Second, we consider systems where two cell types inhibit each other's growth. In such systems, a cell's growth rate decreases (linearly) with the frequency of the other type within the interaction neighborhood. We find that cells that inhibit each other, grow slower when they are well-mixed compared to when they interact within a small neighborhood (see SI). This happens because cells typically are surrounded by their own type, and thus have reduced interactions with the growth inhibiting cells.

Discussion

We developed a mathematical framework that can make quantitative predictions for the dynamics of spatially structured communities consisting of two interacting cell types (Fig. 1). We primarily focussed on microbial cross-feeding systems and used pair approximation to derive analytical predictions for the community level properties from knowledge of the local interaction rules (Fig. 2). These rules are unequivocally defined by two fundamental quantities: the size of the interaction neighborhood and the maximum growth rate that cells achieve when they are completely surrounded by the partner type; both quantities typically differ for the two cell types. We furthermore showed that these local interaction rules can be directly derived from key biophysical parameters of the underlying molecular mechanisms (Fig. 1). We validated this framework using an experimental community of two cross-feeding bacteria and showed that our model reproduces the experimentally observed community level properties (Fig 3). We thus elucidated how the local interaction rules arise from the molecular exchanges, and how the local interaction rules scale up to determine community properties. Generally, we are thus able to scale up from molecular-level properties, to individual-level properties, to community level properties.

For spatial cross-feeding communities we found that the local and global properties can change independently. The global composition of the community (i.e. the equilibrium frequency of the two types) is set by the ratio of the maximum growth rates, which mostly depends on the leakage rates of the exchanged metabolites. In contrast, the composition of the local neighborhood, the one that matters most for the cells, is set by the neighborhood size, which mostly depends on the ratio of the uptake and diffusion rates. Different biophysical parameters thus control the global and local properties of the community.

Small neighborhood sizes reduce the frequency of the partner type around each cell. This happens because cells place their offspring close-by in space. In our model we assumed that cells cannot actively move. However, cells could overcome the negative effect of having a small neighborhood size if they could actively move to locations with a higher frequency of the partner type. Without such active movement, cross-feeding cells in a spatial system grow slower than in equivalent well-mixed system (Fig. 4). This reduction in growth rate can even lead to the collapse of the community: cells in asymmetric communities (i.e. communities where the two type have different maximum growth rates) can stably coexist in well-mixed system, yet they might not coexist in a spatial system, if they interact at small ranges (Fig. 5).

Previous work suggested that short-range interactions provide an advantage to cooperative interactions, because they separate cooperative types from non-cooperators [24, 35–37]. In general, when it is good to be surrounded by your own type, we expect short-range interactions to be beneficial, as for example when two cell types inhibit each other. When it is good to be surrounded by the other type, we expect short-range interactions to be detrimental, as for example when two cell types exchange beneficial resources. When a community consists of more than two cell types the situation can be more complex. Previous studies have shown that short-range

interaction can be beneficial for cross-feeding communities that contain non-producing cells: the short-range interactions harm the producing cells, but they harm the non-producers even more, and thus prevent them from taking over [20, 38, 39].

Our mathematical framework can model a variety of systems of two interacting cell types. Here we considered systems where a type's growth rate increases with the number, or the frequency, of the partner type within the interaction neighborhood, and systems where the two cell types inhibit each other's growth rate. Our framework can model systems beyond the bacterial world, as long as they are composed of two interacting types. For example, it could be adapted to model homeostasis in tissues composed of two different cell types that control each other's growth via growth factors [40, 41].

Many biological systems are spatially structured and multi-scale: they consist of individual entities (e.g. cells or species) that interact with each other in space. The sum of these interactions determines the global properties of the system (e.g. multicellular organism or microbial community). To understand such biological systems, it is essential to scale between levels of organization. Our work provides a contribution to this effort by creating a mathematical framework that can scale from molecular mechanisms, to local interaction rules, to global system properties.

Methods

Model assumptions

- Two cell types, A and B, fully occupy a regular directed graph. Type A interacts with r_A neighbors, type B interact with r_B neighbors, where $r_B \geq r_A$.
- A cell's growth rate increases linearly with the frequency of the other type within its interaction neighborhood.
- Individuals reproduce with a probability proportional to their growth rate, and their offspring replaces a random neighbor within the replication neighborhood.
- The replication neighborhood is identical to the smallest interaction neighborhood $r_R = r_A$.

We used pair approximation to derive analytical predictions and a cellular automaton to run simulations of the model.

Pair approximation

We track the number of all pairwise links $N_{X \leftarrow Y}$, where $X, Y \in A, B$. The total number of links changes through time, because the neighborhood sizes are different for the two cell types (see SI). This is in contrast to previous work were the two neighborhood sizes are the same and the total number of links remains constant. Two events can change the number of links: a type A cell reproduces and replaces a B neighbor, with rate T^+, or a type B cell reproduces and replaces an A neighbor, with rate T^-. During a T^+ event the number of A cells thus increases by one, and during a T^- event it decreases by one. The rate T^+ is given by:

$$T^+ = P(A) \cdot P(B|A, r_A) \cdot \frac{1+P(B|A,r_A)(r_A-1)}{r_A} \frac{\hat{\mu}_A}{\langle \mu \rangle}$$

where the first factor gives the probability of choosing a type A focal cell, the second the probability that this A cells has a type B neighbor, and the third the average growth rate of a type A cell that has at least 1 type B neighbor, relative to the average population growth rate $\langle \mu \rangle$, which is given by:

$$\langle\mu\rangle = P(A)\cdot P(B|A,r_A)\cdot\hat{\mu}_A + (1-P(A))\cdot P(A|B,r_B)\cdot\hat{\mu}_B$$

Similarly, for T^- we have:

$$T^- = P(B)\cdot P(A|B,r_B)\cdot\frac{1+P(A|B,r_B)(r_B-1)}{r_B}\frac{\hat{\mu}_B}{\langle\mu\rangle}\cdot\frac{P(A|B,r_A)}{P(A|B,r_B)}$$

The extra factor at the end is the probability that a type A neighbor that is in the (larger) interaction neighborhood is also part of the smaller replication neighborhood. We can express all probabilities as function of the number of links:

$$P(A) = \frac{N_{A\leftarrow A}+N_{A\leftarrow B}}{rA\cdot N}, \quad P(B|A,r_A) = \frac{N_{A\leftarrow B}}{N_{A\leftarrow B}+N_{A\leftarrow A}}$$
$$P(B) = \frac{N_{B\leftarrow A}+N_{B\leftarrow B}}{rB\cdot N}, \quad P(A|B,r_B) = \frac{N_{B\leftarrow A}}{N_{B\leftarrow A}+N_{B\leftarrow B}}$$

where N is the total number of cells in the systems.

When a T^+ or T^- event happens, the number of $X\leftarrow Y$ link changes by Δ^+_{XY} and Δ^-_{XY}, respectively. These changes can be fully expressed as function of $N_{X\leftarrow Y}$, as we show in the SI. We can then write the dynamical equation for $N_{X\leftarrow Y}$ as:

$$\frac{dN_{X\leftarrow Y}}{dt} = T^+\cdot\Delta^+_{XY} + T^-\cdot\Delta^-_{XY}$$

We can solve these equation numerically to obtain the temporal dynamics, and we can solve them analytically for the steady state, see SI for full derivation.

Model parametrization

To parametrize the model, we need to specify the size of the interaction neighborhood r and the maximum growth rate of a cell $\hat{\mu}$. We previously showed that that interaction-range (R) of a cell (the distance over which amino-acids can be exchanged) is given by [11]:

$$R = \beta\sqrt{\frac{2(1-\rho)^2\cdot D}{\rho(2+\rho)\cdot(r^u+r^l)}}\ln\left[\frac{r^l}{\gamma}\left(1+\sqrt{1+\frac{4\gamma}{r^l}}\right)+4\right] \tag{4}$$

where β and γ are constants, ρ is the volume fraction occupied by cells, and D, r^u, and r^l are the rates of diffusion, uptake, and leakage of the exchanged metabolite (see SI for details and parameter values). The number of neighbors can be directly calculated from the interaction range:

$$r = \left(2R(l-w)+\pi\left(2R+\tfrac{w}{2}\right)^2-\pi\left(\tfrac{w}{2}\right)^2\right)\cdot\rho_{2D} \tag{5}$$

where l and w are the average length and width of cells, and ρ_{2D} is the number of cells per area.

We previously showed that that maximum growth rate of a cell $\hat{\mu}$ is given by [11]:

$$\hat{\mu}\approx\mu_{wt}\cdot\frac{r^l}{\gamma}\left(\sqrt{1+\frac{2\gamma}{r^l}}-1\right) \tag{6}$$

Cellular automaton

In the cellular automaton cells are placed on a square grid of size 100 x 100 with periodic boundary conditions. Cells interact within an extended Moore neighborhood with range d_X, and thus have a neighborhood size of $(2d_X + 1)^2 - 1$. For each cell we calculate the growth rate (using linear growth function) and we randomly pick a cell to reproduce with a probability that is proportional to its growth rate. We then randomly replace one of its neighbors. The grid was initialized with a random arrangement with a given frequency of the two types. We parameterized the experimental community by choosing $d_{\Delta T} = 1$ (8 neighbors) and $d_{\Delta P} = 5$ (120 neighbors). Simulations are implemented in C++.

Experimental communities

The experimental methods are described in detail in reference [11], here we summarize the most relevant details. The community consists of two strains of *Escherichia coli*: ΔT: MG1655 trpC::frt, PR-sfGFP and ΔP: MG1655 proC::frt, PR-mCherry. The deletion of *trpC* and *proC* prevents the production of tryptophan and proline, respectively. Cells are labeled with constitutively expressed fluorescent proteins.

Cells were grown in microfluidic chambers of 60x60x0.76μm, the small height forces cells to grow in a monolayer. The chambers open on one side into a feeding channel of 22μm high and 100μm wide. The microfluidic devices were fabricated from Polydimethylsiloxane (PDMS) using SU8 photoresist molds. Overnight cultures of the two strains (started from single colonies) were mixed in a 1:1 volume ratio and loaded into the microfluidic devices by pipette. Cells were grown on M9 medium (47.8mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.6 mM NaCl and 18.7 mM NH_4Cl) supplemented with 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.2% glucose, and 0.1% Tween-20. For the first 10h the medium was supplemented with 434 mM of L-proline and 98 mM of L-tryptophan to allow cells to grow independently and fill the chambers. Subsequently, cells were grown without externally supplied amino acids.

The growth of the communities was followed using time-lapse microscopy. Phase contrast and fluorescent images were taken every 10min using fully automated Olympus IX81 inverted microscopes, equipped with a 100x NA1.3 oil objective, a 1.6x auxiliary magnification, and a Hamamatsu ORCA flash 4.0 v2 sCMOS camera. The sample was maintained at 37°C with a microscope incubator.

Image analysis

Microscope images were analyzed using Vanellus (version v1.0 [42]). Images were registered, cropped to the area of the growth chambers, and deconvoluted. Images were segmented on the fluorescent channels using custom build Matlab routines [11] or on a combination of the phase contrast and fluorescent channels using Ilastik (version v1.3.3 [43]). Cell tracking was done using custom build Matlab routines [11] followed by manual correction.

Cell length and width were measured by fitting an ellipse to the cell shape and taking the major and minor axis length, respectively. Cell growth rates were estimated by fitting a linear regression to the log transformed cell lengths over a 40min time window [11].

Data analysis

For each chamber we estimated $P(\Delta T)$ as the number of pixels occupied by ΔT cells divided by the total number of pixels occupied by cells. We estimated $P(\Delta T|\Delta P, r_{\Delta P})$ by first calculating the local ΔT frequency for each ΔP cell, and averaging this over all ΔP cells within the chamber.

The local frequency was calculated as the number of pixels occupied by ΔT cells divided by the total number of pixels occupied by cells, considering only pixels within 12.1μm (the interaction range of ΔP) of the cell perimeter. $P(\Delta P|\Delta T, r_{\Delta T})$ was estimated in similar way, using the interaction range of 3.2μm for ΔT.

The difference between the average growth rates of *clustered* and *randomized communities* was estimated using an *in-silico* experiment. For the clustered communities experimentally observed spatial arrangements of 22 communities were converted to a 40x40 square grid. Each grid site was assigned the cell type that occupied the majority of the corresponding pixels in the real image. Grid sites that remained empty after this procedure were randomly assigned one of the types. This grid was used as input for a previously described individual based model that can predict single cell growth rates [11]. The resulting growth rates were averaged over all cells to obtain the average growth rate for the clustered communities. Subsequently the 40x40 grids were randomized, keeping the frequency of the two cell types constant but permuting their locations. We calculated the average community growth rate for 20 independent randomizations, and we assigned the average over all these permutations as the growth rate of the randomized communities.

Author contributions

SvV: Conceptualization, Methodology, Formal analysis, Writing – original draft. CH: Methodology, Formal analysis, Writing – review & editing. MA: Conceptualization, Writing – review & editing. ADC: Conceptualization, Methodology, Formal analysis, Writing – original draft. The authors declare no conflicts of interest.

Acknowledgements

We thank Michael Doebeli and the members of the Microbial Systems Ecology group, the Doebeli lab, and the Hauert lab for valuable input and discussion. SVV was funded by the SNSF Postdoc.Mobility fellowship nr. 175123, CH was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) grant nr. RGPIN-2015-05795, and MA and ADC were funded by SNSF grant nr. 31003A_169978, Eawag, and ETH Zurich.

References

[1] Stephen R Proulx, Daniel E L Promislow, and Patrick C Phillips. Network thinking in ecology and evolution. *Trends in Ecology & Evolution*, 20(6):345–353, 2005.

[2] Simon A. Levin. Ecosystems and the Biosphere as Complex Adaptive Systems. *Ecosystems*, 1:431–436, 1998.

[3] S T A Pickett and M L Cadenasso. Landscape Ecology: Spatial Heterogeneity in Ecological Systems. *Science*, 269(5222):331 – 334, 1995.

[4] Anurag A Agrawal, David D Ackerly, Fred Adler, A Elizabeth Arnold, Carla Cáceres, Daniel F Doak, Eric Post, Peter J Hudson, John Maron, Kailen A Mooney, Mary Power, Doug Schemske, Jay Stachowicz, Sharon Strauss, Monica G Turner, and Earl Werner. Filling key gaps in population and community ecology. *Frontiers in Ecology and the Environment*, 5(3):145–152, apr 2007.

[5] Otto X Cordero and Manoshi S Datta. Microbial interactions and community assembly at microscales. *Current Biology*, (31):227–234, 2016.

[6] Paul G Falkowski, Tom Fenchel, and Edward F Delong. The microbial engines that drive earth's biogeochemical cycles. *Science*, 320(5879):1034 – 1039, 2008.

[7] Hans-Curt Flemming, Jost Wingender, Ulrich Szewzyk, Peter Steinberg, Scott A Rice, and Staffan Kjelleberg. Biofilms: an emergent form of bacterial life. *Nature Reviews Microbiology*, 14(9):563–575, sep 2016.

[8] Hans-Curt Flemming and Stefan Wuertz. Bacteria and archaea on Earth and their abundance in biofilms. *Nature Reviews Microbiology*, 17(4):247–260, 2019.

[9] Bjarke B. Christensen, Janus A.J. Haagensen, Arne Heydorn, and Søren Molin. Metabolic commensalism and competition in a two-species microbial consortium. *Applied and Environmental Microbiology*, 2002.

[10] M. J. I. Muller, B. I. Neugeboren, D. R. Nelson, and A. W. Murray. Genetic drift opposes mutualism during spatial population expansion. *Proceedings of the National Academy of Sciences*, 111(3):1037–1042, 2014.

[11] Alma Dal Co, Simon van Vliet, Daniel Johannes Kiviet, Susan Schlegel, and Martin Ackermann. Short-range interactions govern the dynamics and functions of microbial communities. *Nature Ecology & Evolution*, pages 1–10, 2020.

[12] Peter Jensen, Mette Kolpen, Kasper N. Kragh, and Michael Kühl. Microenvironmental characteristics and physiology of biofilms in chronic infections of CF patients are strongly affected by the host immune response. *APMIS*, 125(4):276–288, 2017.

[13] Clayton W Hall and Thien-Fah Mah. Molecular mechanisms of biofilm-based antibiotic resistance and tolerance in pathogenic bacteria. *FEMS Microbiology Reviews*, 41(3):276–301, may 2017.

[14] Alma Dal Co, Simon van Vliet, and Martin Ackermann. Emergent microscale gradients give rise to metabolic cross-feeding and antibiotic tolerance in clonal bacterial populations. *bioRxiv*, page 534149, 2019.

[15] Alma Dal Co, Martin Ackermann, and Simon van Vliet. Metabolic activity affects the response of single cells to a nutrient switch in structured populations. *Journal of The Royal Society Interface*, 16(156):20190182, 2019.

[16] Simon van Vliet, Alma Dal Co, Annina R. Winkler, Stefanie Spriewald, Bärbel Stecher, and Martin Ackermann. Spatially Correlated Gene Expression in Bacterial Groups: The Role of Lineage History, Spatial Gradients, and Cell-Cell Interactions. *Cell Systems*, 6(4):496–507.e6, apr 2018.

[17] Bernhard Schink. Synergistic interactions in the microbial world. *Antonie Van Leeuwenhoek Int. J. Gen. Mol. Microbiol*, 81:257–261, sep 2002.

[18] Philip S Stewart and Michael J Franklin. Physiological heterogeneity in biofilms. *Nature Reviews Microbiology*, 6(3):199–210, mar 2008.

[19] Shawn E McGlynn, Grayson L Chadwick, Christopher P Kempes, and Victoria J Orphan. Single cell activity reveals direct electron transfer in methanotrophic consortia. *Nature*, 526:531, sep 2015.

[20] Babak Momeni, Adam James Waite, and Wenying Shou. Spatial self-organization favors heterotypic cooperation over cheating. *eLife*, 2, nov 2013.

[21] Elizabeth G. Wilbanks, Ulrike Jaekel, Verena Salman, Parris T. Humphrey, Jonathan A. Eisen, Marc T. Facciotti, Daniel H. Buckley, Stephen H. Zinder, Gregory K. Druschel, David A. Fike, and Victoria J. Orphan. Microscale sulfur cycling in the phototrophic pink berry consortia of the Sippewissett Salt Marsh. *Environmental Microbiology*, 16(11):3398–3415, 2014.

[22] Sophie E. Darch, Olja Simoska, Mignon Fitzpatrick, Juan P. Barraza, Keith J. Stevenson, Roger T. Bonnecaze, Jason B. Shear, and Marvin Whiteley. Spatial determinants of quorum signaling in a Pseudomonas aeruginosa infection model. *Proceedings of the National Academy of Sciences*, 115(18):201719317, 2018.

[23] Carey D Nadell, Knut Drescher, and Kevin R Foster. Spatial structure, cooperation and competition in biofilms. *Nature Reviews Microbiology*, 14:589, jul 2016.

[24] Knut Drescher, Carey D Nadell, Howard A Stone, Ned S Wingreen, and Bonnie L Bassler. Report Solutions to the Public Goods Dilemma in Bacterial Biofilms. *Current Biology*, 24(1):50–55, 2014.

[25] Jodi L. Connell, Jiyeon Kim, Jason B. Shear, Allen J. Bard, and Marvin Whiteley. Real-time monitoring of quorum sensing in 3D-printed bacterial aggregates using scanning electrochemical microscopy. *Proceedings of the National Academy of Sciences*, 111(51):18255–60, 2014.

[26] Michael T Mee, James J Collins, George M Church, and Harris H Wang. Syntrophic exchange in synthetic microbial communities. *Proceedings of the National Academy of Sciences*, 111(20):E2149–56, 2014.

[27] Glen D'Souza, Silvio Waschina, Samay Pande, Katrin Bohl, Christoph Kaleta, and Christian Kost. Less is more: Selective advantages can explain the prevalent loss of biosynthetic genes in bacteria. *Evolution*, 68(9):2559–2570, 2014.

[28] Glen D'Souza, Shraddha Shitut, Daniel Preussger, Yousif Ghada, Silvio Waschina, and Christian Kost. Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Natural Product Reports*, 35:455–488, 2018.

[29] J Jeffrey Morris. Black Queen evolution : the role of leakiness in structuring microbial communities. *Trends in Genetics*, 31(8):475–482, 2015.

[30] Hisashi Ohtsuki, Christoph Hauert, Erez Lieberman, and Martin A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(25):502–505, 2006.

[31] Feng Fu, Long Wang, Martin A Nowak, and Christoph Hauert. Evolutionary dynamics on graphs: Efficient method for weak selection. *Physical Review E*, 79(4):46707, apr 2009.

[32] Feng Fu, Martin A Nowak, and Christoph Hauert. Invasion and expansion of cooperators in lattice populations: prisoner's dilemma vs. snowdrift games. *Journal of theoretical biology*, 266(3):358–366, oct 2010.

[33] Ronald Dickman. Kinetic phase transitions in a surface-reaction model: Mean-field theory. *Physical Review A*, 34(5):4246–50, 1986.

[34] H. Matsuda, N. Ogita, A. Sasaki, and K. Sato. Statistical Mechanics of Population: The Lattice Lotka-Volterra Model. *Progress of Theoretical Physics*, 88(6):1035–1049, 1992.

[35] Jeffrey A Fletcher and Michael Doebeli. A simple and general explanation for the evolution of altruism. *Proc. Biol. Sci.*, 276(1654):13–19, jan 2009.

[36] Martin A. Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560–1563, dec 2006.

[37] Felix J H Hol, Peter Galajda, Krisztina Nagy, Rutger G Woolthuis, Cees Dekker, and Juan E Keymer. Spatial Structure Facilitates Cooperation in a Social Dilemma : Empirical Evidence from a Bacterial Community. *PLoS ONE*, 8(10):2–11, 2013.

[38] Simon M Stump, Christopher A Klausmeier, and W K Kellogg. Competition and coexistence between a syntrophic consortium and a metabolic generalist, and its effect on productivity. *Journal of Theoretical Biology*, 404:348–360, 2016.

[39] Simon Maccracken Stump, Evan Curtis Johnson, Zepeng Sun, and Christopher A. Klausmeier. How spatial structure and neighbor uncertainty promote mutualists and weaken black queen effects. *Journal of Theoretical Biology*, 446:33–60, jun 2018.

[40] Xu Zhou, Ruth A. Franklin, Miri Adler, Jeremy B. Jacox, Will Bailis, Justin A. Shyer, Richard A. Flavell, Avi Mayo, Uri Alon, and Ruslan Medzhitov. Circuit Design Features of a Stable Two-Cell System. *Cell*, 172(4):744–757.e17, 2018.

[41] Costas A. Lyssiotis and Alec C. Kimmelman. Metabolic Interactions in the Tumor Microenvironment. *Trends in Cell Biology*, 27(11):863–875, 2017.

[42] Daniel J. Kiviet. Vanellus, https://github.com/daankiviet/vanellus/releases/tag/v1.0, 2020.

[43] Christoph Sommer, Christoph Straehle, Ullrich Kothe, and Fred A Hamprecht. Ilastik: Interactive learning and segmentation toolkit. In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 230–233. IEEE, 2011.