Year : 2022

Two design frameworks for optimizing microbial community functions

Shibasaki Shota

Shibasaki Shota, 2022, Two design frameworks for optimizing microbial community functions

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive http://serval.unil.ch
Document URN : urn:nbn:ch:serval-BIB_BBAE9A5D70E92

UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Département de Microbiologie Fondamentale

Two design frameworks for
optimizing microbial community functions

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Shota Shibasaki

Master of Multidisciplinary science, the University of Tokyo

Jury

Prof. Déla Golshayan, Présidente
Prof. Sara Mitri, Directrice de thèse
Prof. Laurent Lehmann, Co-directeur de thèse
Prof. Yolanda Schaerli, Experte
Dre. Xiang-Yi Li Richter, Experte

Lausanne
(2022)

UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Département de Microbiologie Fondamentale

Two design frameworks for
optimizing microbial community functions

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Shota Shibasaki

Master of Multidisciplinary science, the University of Tokyo

Jury

Prof. Déla Golshayan, Présidente
Prof. Sara Mitri, Directrice de thèse
Prof. Laurent Lehmann, Co-directeur de thèse
Prof. Yolanda Schaerli, Experte
Dre. Xiang-Yi Li Richter, Experte

Lausanne
(2022)

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Madame	Prof.	Déla	**Golshayan**
Directeur·trice de thèse	Madame	Prof.	Sara	**Mitri**
Co-directeur·trice	Monsieur	Prof.	Laurent	**Lehmann**
Expert·e·s	Madame	Prof.	Yolanda	**Schaerli**
	Madame	Dre	Xiang-Yi	**Li Richter**

le Conseil de Faculté autorise l'impression de la thèse de

Shota Shibasaki

Master of Arts and Sciences, The University of Tokyo, Japon
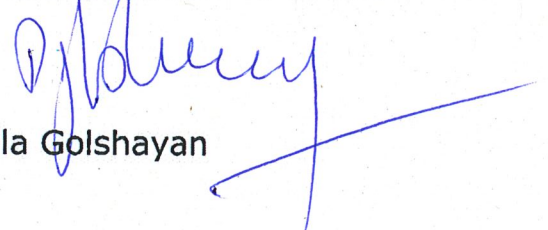
intitulée

Two design frameworks for optimizing microbial community functions

Lausanne, le 26 juillet 2022

pour le Doyen
de la Faculté de biologie et de médecine

Prof. Déla Golshayan

Contents

Acknowledgement

世路多岐 人海遼闊 揚帆待發清曉
There are many roads in life, human societies are as large as the ocean,
I am ready for sailing this morning.
誨我諄諄 南針在抱 仰瞻師道山高
Thanks to the earnest coaching, I have the compass in my mind,
my supervisor's lessons pile up like a mountain
『靑靑校樹』郭輝

Abstract

In the contexts of food or beverage production, biotechnology, and human health, microbial communities provide human societies with many types of benefits called microbial community functions. Although researchers have tried to optimize these microbial community functions, ecological and/or evolutionary dynamics can drive the communities away from the states where community functions are maximized. In this thesis, I theoretically investigate two design frameworks to optimize microbial community functions. One approach is to control the microbial dynamics by fluctuating environmental conditions (Chapter 2), and the other approach is to introduce hierarchical spatial networks (Chapter 4) so that we can restrict species interactions. I show algorithms to reveal the optimal control of the environmental conditions and the optimal allocations of microbes into a given hierarchical spatial structure, respectively. In addition, the environmental fluctuations and the hierarchical spatial structures can affect the fundamental aspects of ecology. When the environmental conditions become harsh, the intensity of demographic noise increases, which affects species diversity (Chapter 3). In a hierarchical spatial structure, stability of a downstream community depends on upstream species because the upstream communities can change the downstream environments by secreting or absorbing chemical compounds that flow downstream. In this case, positive interactions from upstream species to downstream species increases the stability of downstream communities (Chapter 5). These studies show the importance of environmental fluctuations and spatial structures in applied and fundamental microbial ecology and evolution.

Résumé

Dans les contextes de la production d'aliments ou de boissons, de la biotechnologie et de la santé humaine, les communautés microbiennes offrent aux sociétés humaines de nombreux types d'avantages appelés fonctions des communautés microbiennes. Bien que les chercheurs aient tenté d'optimiser ces fonctions, les dynamiques écologiques et/ou évolutives peuvent éloigner les communautés des états où les fonctions des communautés microbiennes sont maximisées. Dans ma thèse, j'étudie théoriquement deux cadres de conception pour optimiser les fonctions des communautés microbiennes. Une approche consiste à contrôler les dynamiques microbiennes en faisant fluctuer les conditions environnementales (Chapitre 2), et l'autre approche consiste à introduire des réseaux spatiaux hiérarchiques (Chapitre 4) afin de restreindre les interactions entre espèces. Je montre des algorithmes pour révéler le contrôle optimal des conditions environnementales et les allocations optimales des microbes dans une structure spatiale hiérarchique, respectivement. De plus, les fluctuations environnementales et les structures spatiales hiérarchiques peuvent affecter les aspects fondamentaux de l'écologie Quand les conditions ambiantes deviennent sévère, l'intensité du bruit démographique augmente, ce qui affecte la diversité des espèces (Chapitre 3). Dans une structure spatiale hiérarchique, la stabilité des communautés aval dépend des communautés en amont car les communautés en amont peuvent modifier les environnements en aval en sécrétant ou en absorbant des composés chimiques qui s'écoulent en aval. Dans ce cas, les interactions positives des espèces en amont aux espèces en aval augmente la stabilité des communautés en aval (Chapitre 5). Ces études montrent l'importance des fluctuations environnementales et des structures spatiales au sein de l'écologie et l'évolution microbiennes appliquées et fondamentales.

Chapter 1

General introduction

1.1 Historical overview

1.1.1 Microbes in human societies

Historically, microorganisms have provided human societies with a variety of benefits. Many types of alcoholic beverages, for example, have been produced with the help of some microorganisms across the world. Wine has been drunk since ca. 6000 BC in Georgia (McGovern et al., 2017), beer has been brewed since 11,000 BC in Israel (Liu et al., 2018), or since ca. 3000 BC in Mesopotamia (Michel et al., 1992) and in China (Wang et al., 2016), and *sake* is considered to have been produced for at least 1300 years in Japan (Kitagaki and Kitamoto, 2013). In addition to these beverages, many types of food have been produced in fermentation: cheese has been made since ca. 5-6,000 BC in Europe (Salque et al., 2013; McClure et al., 2018), and fermented soybean products such as *sofu* or *furu* in China and *natto* in Japan are mentioned in written records in ca. 200 and 1400 AD, respectively (Han et al., 2001; Shurtleff and Aoyagi, 2012). The compositions of microbial species and their dynamic in these types of beverages and food have been investigated in recent studies (Cocolin, 2000; Abriouel et al., 2008; Bokulich et al., 2014; Portillo and Mas, 2016; Yunita and Dodd, 2018; Cason et al., 2020; Bossaert et al., 2021; Tyakht et al., 2021; Somerville et al., 2022).

Microorganisms are also used in some modern biotechnology: to produce biofuels such as bioethanol (Antoni et al., 2007) or to remove toxic compounds such as heavy metals from the environments: i.e., biodegradation and bioremediation (Vidali, 2001; Dua et al., 2002; O'Brien and Buckling, 2015). Typically, genetically engineered *Escherichia coli* (Shin et al., 2010), cyanobacteria (Nozzi et al., 2013; Lea-Smith and Howe, 2017), or algal species (Georgianna and Mayfield, 2012; Brodie et al., 2017) are used in biofuel production. In the context of bioremediation of heavy metals, previous studies investigated *E. coli* (Lakshmanan et al., 2008), *Pseudomonas aeruginosa* (O'Brien et al., 2014), *Rhizobium* (Alfadaly et al., 2021), *Saccharomyces cerevisiae* (Soares and Soares, 2012), and the communities of *Viridibacillus arenosi*, *Sporosarcina soli*, and two strains of *Enterobacter cloacae* (Kang et al., 2016). Some of these species are naturally isolated while others are engineered species. For removing pollutants other than heavy metals

using microorganisms, see, for example, Chen et al. (2009), Dell'Anno et al. (2012), and Kalantary et al. (2014).

Another type of benefit from microorganisms can be found inside human bodies. For microbial species, each person or host can be seen as an "island" in meta-community theory (Costello et al., 2012), and the microorganisms interact with their hosts as other larger organisms interact with their environments. The benefits from the microbial species to the hosts are providing nutrients, training the immune system, and protecting against pathogens, while the hosts' behaviours affect the assembly of the microbial communities (Karkman et al., 2017). Recent studies show that the human gut microbiota are associated with some diseases, although mechanisms are rarely known (Xu et al., 2013; Marchesi et al., 2016; Sánchez et al., 2017; Selber-Hnatiw et al., 2017). For example, relative abundances of Bacteroidetes are associated with obesity (Ley et al., 2006), and transfer of intestinal microbiota from lean donors increases the insulin sensitivity of recipients with metabolic syndrome (Vrieze et al., 2012). In addition, recent studies suggest a relationship between gut microbiota composition and depression (Evrensel and Ceylan, 2015; Cheung et al., 2019). For these reasons, many researchers have investigated how to design gut microbiota to improve host health (Pham et al., 2017; Swann et al., 2020; Clark et al., 2021).

Many of the benefits mentioned above are achieved by microbial communities rather than single-species populations. For this reason, I call these benefits community functions in this thesis. One of the main motivations of my thesis is to investigate how to design microbial communities to improve their community functions. To address this question, we need to know the principles of ecology and evolution that microorganisms follow. In the next subsection, I review the concepts in ecology and evolution that are important for designing microbial communities to maximize their community functions.

1.1.2   Ecological and evolutionary background

To design microbial communities, I introduce the basic ideas in ecology and evolution here. First of all, we need to consider how species interact with one another in a community. A simple classification of species interactions is whether they are positive, negative, or neutral by comparing the growth of the focal species (recipient) with and without a partner species (donor). In addition, we can also classify species interactions by combining the signs of effects between two species (Godsoe et al., 2017; Kong et al., 2018): cooperation or mutualism ($+/+$), competition ($-/-$), exploitation or predation ($+/-$), commensalism ($+/0$), amensalism ($-/0$), and neutralism ($0/0$). However, these classifications are phenomenological. Species interactions can be classified according to their mechanisms (Estrela et al., 2019; Dos Santos et al., 2022). For example, a positive interaction can occur when a donor species produces resources that a recipient species consumes (i.e., cross-feeding) (Seth and Taga, 2014; Smith et al., 2019) or degrading/deactivating toxins that the recipient species is sensitive to (cross-protection) (Yurtsev et al., 2016). On the other hand, a negative interaction can occur by competing for limiting resources (exploitative competition) or by secreting toxins (interference competition) (Birch, 1957; Case and Gilpin, 1974; Cornforth and Foster, 2013; Ghoul and Mitri, 2016).

These interactions between multiple species affect their coexistence. Especially, conditions under which competitive species can coexist have been investigated by many researchers in ecology. In 1934, Gause showed that two competitive microbial species, *Paramecium caudatum* and *P. aurelia*, cannot coexist (Gause, 1934). This experimental result conceptualized the competitive exclusion principle: if two species occupy precisely the same ecological niche, either of the two species drives the other species to extinction (Hardin, 1960). This principle suggests that the difference in the niche is necessary for species' coexistence. Indeed, Tilman (1980) graphically showed that two species competing for two resources can coexist if the growth of the two species is limited by the different resources and if the two species consume more of one resource that limits their growth. This idea is known as the resource-ratio theory (Miller et al., 2005). From this theory, one may consider that the number of coexisting species cannot exceed the number of resources. However, this is not accurate. There are limiting factors other than resources: predators and external environmental factors such as temperatures can be limiting factors and the number of coexisting species cannot exceed the number of limiting factors (Levin, 1970; Krebs, 2013).

One of the leading frameworks for understanding species coexistence these days was developed by Chesson (Chesson, 2000a) and his colleagues. This theory is called the modern coexistence theory (Saavedra et al., 2017). The modern coexistence theory is based on the invasion growth rate of each species: the necessary condition for species coexistence is that each species has a positive long-term mean growth rate at low density (Chesson, 2000a) under temporally (Chesson, 1994) and/or spatially (Chesson, 2000b) variable environments. The modern coexistence theory clarifies that both niche difference between species and their fitness difference matter in the coexistence of species, see Chesson (2000a), Letten et al. (2017), or Godwin et al. (2020) for the details in the two-species scenario, and Saavedra et al. (2017) for the expansion to the three- or more-species scenario. In addition, the spatial or temporal fluctuations partition the invasion growth rate into the invasion growth rate in the absence of frequency-dependent effects, the fluctuation-independent effects, the relative non-linearity (difference in response to the variance of limiting factors), the storage effect (covariance between environments and density-dependent interaction parameters), and the growth-density covariance (in case of spatial variation), see Barabás et al. (2018) for more details. The spatial and/or temporal variations enable species to coexist via these factors. The modern coexistence theory has been applied in some empirical studies with microorganisms (Narwani et al., 2013; Tan et al., 2017; Grainger et al., 2019; Li et al., 2019).

In addition to species coexistence, we also need to consider species diversity. This is because some studies reveal the positive relationships between species diversity and community functions (Dell'Anno et al., 2012; Demeter, 2015). Before the 1970s, researchers considered that complex communities, which have many species (higher diversity) and/or many interspecific interactions (higher connectance), were more stable than simple ones (MacArthur, 1955). This suggested that the coexistence of many species would be trivial. However, Gardner and Ashby (1970) and May (1972) showed this is not the case, assuming random interactions between pairs of species.

Instead, these studies revealed that complex communities are likely to be unstable. These studies have led to the stability-complexity (or stability-diversity) debate (McCann, 2000; Jacquet et al., 2016): why complex ecosystems are ubiquitous in nature although the theoretical studies suggest that such systems should be unstable. I emphasize that the models of Gardner and Ashby (1970) and May (1972) should be regarded as null models: for example, these models do not include any spatial and/or temporal fluctuations in environments, which allow species to coexist. In addition, species interaction networks are not always random (Landi et al., 2018). For example, Mougi and Kondoh (2012) show that complexity increases the probability that the community is stable when species interactions are composed of mutualistic and exploitative interactions in a 3 : 1 ratio. Coyte et al. (2015) show that introducing competitive species in mutualistic communities increases the stability although increasing the number of species has a destabilizing effect. In addition, Bairey et al. (2016) show that diversity can increase the stability in randomly interacting communities by introducing high-order interactions, which are synergistic effects that cannot be captured in pairwise interactions.

So far, I summarize studies related to species interactions, coexistence, and diversity to design microbial communities with high community functions. It is also worth explaining what stability means in ecology and evolution because the designed microbial communities should be stable over time. In ecology, many definitions of stability have been used (Pimm, 1984; Ives and Carpenter, 2007; Donohue et al., 2013; Landi et al., 2018). One of the most widely used definitions in theoretical ecology is the linear (or local) stability: whether the dynamics converge to the initial equilibrium after a small perturbation in states of the systems. This stability is binary (i.e., stable or not) and is investigated by the Jacobian matrix: see Murray (2002) for the mathematical details. When a focal equilibrium is linearly stable, one can also analyze another stability, resilience: the time until the dynamics converge to the initial equilibrium. Higher resilience indicates that the dynamics return to the initial equilibrium faster, and theoretically this stability is measured by the absolute value of the dominant eigenvalue of the Jacobian matrix. Another stability that has recently been used in ecology is structural stability. This stability measures the volume of parameters under which species composition does not change, i.e. the perturbation is applied to the *parameters* of the model. Although structural stability typically considers the volume of parameters where all species coexist (Rohr et al., 2014; Butler and O'Dwyer, 2018; Saavedra et al., 2017; Song and Saavedra, 2018; Cenci and Saavedra, 2018; O'Sullivan et al., 2019), we can consider this stability for any species compositions (Portillo et al., 2021). While these definitions of stability focus on certain equilibrium, we can also consider stability in the whole system. For example, the number of linearly stable states (also known as alternative stable states) can be seen as the degree of instability of the system (Ives and Carpenter, 2007). Robustness refers to the resistance to additional extinctions: i.e., perturbation drives one species to extinction (Dunne et al., 2002). In the contexts of community assembly and invasion biology, we can also consider stability as the resistance to invasion: stable communities are less likely to be colonized by new species (Post and Pimm, 1983). This stability increases with the number of species (species richness) in the community (Case, 1990; Stachowicz et al., 1999; Bonanomi et al., 2014; Hromada

et al., 2021).

In designing microbial communities, we sometimes need to consider stability over evolutionary time scales. Previous studies show that microbial community functions can be lost due to mutants that do not contribute to the community functions (Ellis et al., 2007; O'Brien et al., 2014; Rugbjerg et al., 2018b,a). There are two frameworks that deal with stability at the evolutionary time scales: evolutionary game theory (Maynard Smith and Price, 1973; Nowak, 2006) and adaptive dynamics (Geritz et al., 1998; Diekmann, 2004; Waxman and Gavrilets, 2005). In both frameworks, an evolutionarily stable strategy represents a trait that is not invaded by any mutants when the majority of individuals in the population (i.e, the resident population) have this stable trait (Broom and Rychtar, 2013). Note that mutants are assumed to have trait values close to that of the residents in adaptive dynamics. Importantly, evolutionarily stable strategies do not necessarily spread in a population. In adaptive dynamics, convergence stability indicates whether a focal trait spreads in a population or not: a trait value gradually approaches a convergence stable strategy when the initial trait value of the resident population is enough close to the convergence stable strategy (Brännström et al., 2013; Brown, 2016).

1.1.3 Mathematical background

Next, I introduce theoretical or mathematical frameworks that are useful to design microbial communities. First, we need to consider how to represent the ecological and/or evolutionary dynamics of microbial communities. Here, I introduce two frameworks based on ordinary differential equations. One is a generalized Lotka-Volterra model, which is an extension of a prey-predator Lotka-Volterra model (Lotka, 1910; Volterra, 1926) to include other types of species interactions. Typically, the models in this framework include linear terms representing the net effects of pairwise species interactions, although some recent studies introduced high-order interactions (Bairey et al., 2016; Mayfield and Stouffer, 2017). An important feature of generalized Lotka-Volterra models is that species interactions are phenomenological: if two interaction terms of a pair of species are $(-/-)$, we understand that these two species are competing, but the model does not tell us the mechanisms behind competition. One advantage of this feature is that it is easy to fit this type of model to empirical data: we do need time-series data of species abundances but not time-series data of resource and toxin concentrations. The disadvantages of this framework are, in contrast, (i) that some community functions cannot be directly evaluated, and (ii) that species interactions can be too simplified. The first disadvantage comes from the fact that the models in this framework usually include only the dynamics of species. If a target community function is producing and/or degrading certain compounds, we need either to introduce the dynamics of the target compounds (i.e., dealing with the target compounds as "species") or to estimate concentrations of the target compounds from species abundances (e.g., the productivity of a given compound is estimated by abundances of producer species). The second disadvantage comes from the fact that many models in this framework use linear terms to capture species interactions (Momeni et al., 2017), which cannot capture the environmental dependencies of signs and strength of species interactions (Chamberlain et al., 2014; Piccardi et al., 2019). One solution

to the second problem may be introducing non-linear terms in species interactions, but we need to choose the best formulations for species interactions in this case.

The second framework for representing microbial community dynamics is known as a consumer-resource model (MacArthur, 1970; Tilman, 1986; Chesson, 1990), where the dynamics of species and environmental factors are considered and species indirectly interact with other species. This type of model tends to investigate the dynamics of (a)biotic resources and consumer species (i.e., the models include resource competition and prey-predator interactions), but we can include the dynamics of toxin and production/secretion of resources and/or toxins, implementing many species interactions in microbial communities. Therefore, the advantages of consumer-recourse models are (i) to explicitly evaluate the production and degradation of target compounds and (ii) to represent environmentally mediated species interactions. On the other hand, this framework requires the dynamics of the environmental factors to fit the models, but such data are not always available.

In the context of evolutionary game theory, we can similarly consider two frameworks. In one framework, the payoffs of the strategies are given by a payoff matrix (in two-player games) (Nowak, 2006) or a function of the densities/frequencies of the strategies in the population (in n-player games) (Archetti and Scheuring, 2011; Sanchez and Gore, 2013). In this case, the results of the games are phenomenological. On the other hand, we can also consider a game with environmental feedback to include mechanistic interactions: the dynamics of the strategies in the population are determined by the environment, while environmental dynamics are given by the densities/frequencies of the strategies in the population (Weitz et al., 2016; Gong et al., 2018; Tilman et al., 2020).

After mathematically representing the dynamics of microbial communities, we need to formulate an objective function, a community function to be maximized or minimized. Because we have seen community functions as the production or degradation of chemical compounds, the community functions in this thesis depend on the chemical compounds in the system. Mathematically, we can write a community function at time t as follows:

$$\phi\left(\vec{C}(t)\right) \propto \vec{w}\vec{C}^{\mathrm{T}}(t), \tag{1.1}$$

where $\vec{C} = (C_1, \ldots, C_M)$ represents the concentrations of chemical compounds $1, \ldots, M$ in the system, $\vec{w}$ is the weight vector whose absolute values represent the importance of production or the degradation of the chemical compound while the signs are determined by whether we want to produce or degrade the compounds, and T represents the transpose. We can also consider an objective function when we grow a microbial community from time $t = 0$ to time $t = T$:

$$\Phi(T) \equiv \int_0^T dt\phi\left(\vec{C}(t)\right). \tag{1.2}$$

If microbial community dynamics converge to a linearly stable equilibrium quickly relative to T, $\Phi(T)$ is approximately proportional to $\phi\left(\vec{C}(t)\right)$. This indicates that evaluating community

functions in the form of $\phi\left(\vec{C}(t)\right)$ at a stable equilibrium is reasonable when the dynamics quickly converge to the linearly stable equilibrium. In contrast, the objective functions in the form of $\Phi(T)$ are useful when the convergence to the linearly stable equilibrium is slow (i.e., low resilience) or when the dynamics converge to other types of attractors, such as limit cycles. In both ways of formulating the objective functions, they include the time-dependent variable $\vec{C}(t)$. Even at an equilibrium state, we are rarely able to explicitly write down how $\vec{C}(t)$ changes depending on the designs of microbial communities. For this reason, we cannot reduce the optimization problems in this thesis to typical (non-)linear programming problems (Bhunia et al., 2019).

Instead, we can apply optimal control theory to maximize microbial community functions. $\Phi(T)$ in Eq (1.2) can be seen as an objective function in this framework (Sontag, 1998; Lenhart and Workman, 2007). We can control the community dynamics by changing environmental conditions. For example, fermentation processes can be improved by feedback control of inflow media, see literature in Simutis and Lübbert (2015). In the contexts other than fermentation, Angulo et al. (2019) show the schemes to control the microbial communities by combining prebiotics and bactericides. In other studies, researchers use frameworks other than control theory while introducing environmental changes: Treloar et al. (2020) use reinforcement learning and control the medium flowing into a chemostat to maximize the community functions in *silico*, and Wang et al. (2020) introduce the idea of driving communities to a desired steady state by modifying species interactions that are mediated by environments. Typically, the control approaches via environmental fluctuations use feedback systems based on the species' abundances and/or compound concentrations in the system, but measuring these community states or environmental conditions is not always easy; for example, one cannot measure species' abundances by optical density unless the growth medium is transparent.

In such cases, we can consider two alternative frameworks to maximize community functions. One way is to consider the best species compositions (presence or absence of each species) of the communities without environmental fluctuations and spatial structures (Stein et al., 2018; Xu et al., 2019). I call such problems optimal composition problems in this thesis. In this case, we evaluate the community functions with different species compositions. Notably, the best communities are not necessarily composed only of species that directly contribute to the community functions (Xu et al., 2019). For example, species can indirectly improve community functions by facilitating the growth of other species that directly contribute to those functions.

The second idea is to introduce spatial structure and consider where we grow which species (Ben Said and Or, 2017; Ben Said et al., 2020). Hereafter, such an optimization problems is called an optimal allocation problem. The idea behind this approach is that spatial structures enable us to restrict which species interacts with which one, leading to species coexistence that cannot appear in a well-mixed scenario (Kim et al., 2008). For example, we can introduce spatial structures by using multi-stage chemostats, where multiple chemostats are connected by tubes to share the media (Bayrock and Michael Ingledew, 2001; Lin et al., 2002; Raninger and Steiner, 2003; Cinquin et al., 2006; Payne et al., 2012a). In the acyclic (or hierarchical) networks of multi-stage chemostats, only species in upstream chemostats can affect the dynamics of species
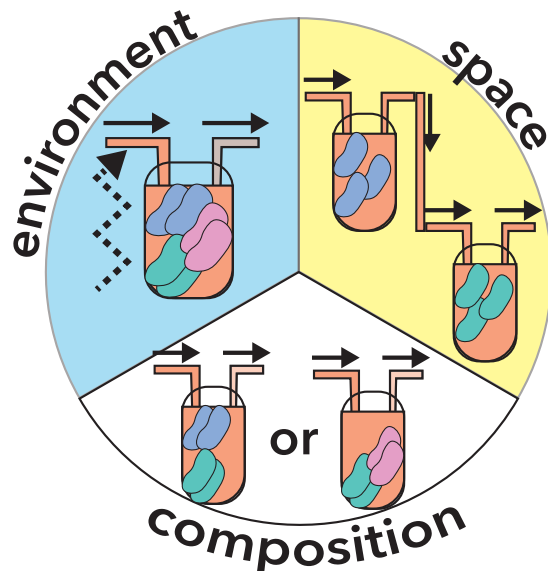
Figure 1.1 – Three optimization approaches

In optimization of microbial community functions, we can consider three types of approaches depending on what we modify: environmental conditions (i.e., controlling microbial dynamics by environmental fluctuations), spatial structures (restricting species interactions by allocating species into a spatial structure), and species composition (analyzing species combinations where species coexist). This figure shows an example when we have three species that have different colors in chemostats.

downstream while the downstream species cannot interact with upstream species. If we have two species competing for the same resources (and thus they cannot coexist without spatial structures), these species can coexist by growing weaker species, which goes extinct without spatial structure, in the upstream chemostat and the other species in the downstream. In short, hierarchical spatial structures enable us to remove some interspecific interactions from the communities. The three approaches to optimize microbial community functions are schematically represented in Fig. 1.1.

However, we need some algorithms to find the best compositions or allocations efficiently because analyzing all combinations or allocations (i.e., the brute force search) is not feasible either in experiments or in simulations (see Chapter 4). Imagine that we have N species and we want to maximize the production of certain beneficial chemical compounds. In the optimal composition problem, the brute force search analyzes $2^N - 1$ combinations: we need to evaluate the effect of the presence and absence of each species, but we can ignore the cases where no species exist. Recently, George and Korolev (2021) introduced a heuristic algorithm to find the best composition of species and they evaluated when this algorithm worked, although they did

not mention the computational costs of the algorithm. In the optimal allocation problem with M-stages of chemostats, the brute force search analyzes N^M allocations when we assume that we allocate one species to each chemostat while allowing the allocation of the same species to multiple chemostats. If we allocate multiple species to each chemostat, the computational cost of the brute force search is $(2^N - 1)^M$. Therefore, we need some efficient search algorithms for both optimal composition and allocation problems.

1.2 Formulating research questions

In the previous section, I reviewed examples of microbial community functions and the background in ecology, evolution, and the relevant theoretical frameworks. In this section, I formulate the two types of research questions I address in the thesis. The first type of research question is "how does one design stable microbial communities that maximize their community functions?" and the second type is "how do these design approaches affect the fundamental aspects of ecology?" In this thesis, I consider community functions as the production or degradation of chemical compounds by microbial species as we have seen in subsection 1.1.1. I evaluate community functions in the form of Eq (1.1) at a linearly stable equilibrium because it is easy to calculate, although I also use the community function similar to Eq (1.2) in Box 1 of Chapter 2. Subsection 1.1.3 shows three types of designing (or optimizing) approaches (Fig. 1.1): controlling the dynamics by environmental fluctuations, allocating species to a spatial structure, and analyzing the best species composition. I use the first two frameworks because these two approaches may affect fundamental aspects in ecology such as species interactions, diversity, and stability. For example, the modern coexistence theory emphasizes the importance of spatial and temporal environmental variations in species coexistence, and these two types of variations relate to the two designing approaches I analyze in this thesis. Therefore, the two types of the research questions are connected.

In the controlling approach, I change the resource and/or toxin concentrations flowing into a chemostat. As many previous studies use feedback control (see the references in subsection 1.1.3), I consider control without a feedback loop. In other words, the environment in the chemostat randomly fluctuates in my controlling approach, leading to the changes in the growth rates of microorganisms. Then, we could ask how frequently we should change the environment to maximize community functions in the long term. In addition to growth rates, environmental fluctuations also affect the abundance of microorganisms. For example, if a harsh environmental condition (e.g., low resource concentration or high toxin concentration) continues for a long time, the abundance of each species decreases and we cannot ignore demographic noise, which could affect species extinction and species compositions. This suggests that the rate of environmental changes can change the intensity of demographic noise and that both environmental fluctuations and demographic noise may together affect species diversity.

In the allocation approach, I introduce spatial structure by considering multi-stage chemostats. As the motivation for introducing spatial structure is restricting species interactions, I assume

acyclic networks of multi-stage chemostats: we have hierarchical spatial structures where upstream species can affect the growth of downstream species, but downstream species cannot interact with the upstream species. To maximize community function, we can ask first "what is the best network structure of a multi-stage chemostat?" Once the network structure of the multi-stage chemostat is determined, we need to find the best allocations of species that maximize the community functions. Using the brute force search is, however, time-consuming and we need to ask "how can we find the best allocation efficiently?" In the context of fundamental ecology, the acyclic networks given by the multi-stage chemostats can be seen as the simplification of a gut (Cinquin et al., 2006; Payne et al., 2012b) or a river (Brown and Swan, 2010; Carraro et al., 2020) and we can consider the meta-community dynamics in such networks. Although there have been many studies of meta-communities with various network structures (Hanski and Gyllenberg, 1993; Leibold et al., 2004; Economo and Keitt, 2007; Chisholm et al., 2011), these studies assume that species interact within each patch and the networks affect the migration rates of species. In contrast, our spatial structures allow species to affect those in the same patch (chemostat) and those in downstream patches. Then, we can ask "how does the upstream community affect the stability of the downstream community?" Stability in this context has two aspects: The first is resistance to invasion and the other is resistance to environmental changes caused by the upstream species, which is similar but not identical to structural stability (see Chapter 5).

In summary, the research questions I address in this thesis are as follows:

1. How can one maximize community function by fluctuating the environment without feedback?

2. How do environmental fluctuations affect species composition and diversity while affecting the intensity of demographic noise?

3. How can one maximize community function by introducing hierarchical spatial structures?

    (a) What is the best spatial structure to maximize community function?

    (b) What are efficient methods to find the best allocations of species to a given network structure?

4. How does an upstream community affect the stability of a downstream community in a hierarchical network?

See also Fig. 1.2 for a schematic illustration. Research questions 1 and 2 correspond to the controlling approaches (introducing environmental fluctuations) while research questions 3 and 4 are based on the allocating approach (introducing spatial structures). In addition, research questions 1 and 3 are toward applied ecology and evolution while questions 2 and 4 investigate more fundamental aspects of ecology.

1.3   Thesis structure

In the following four chapters, I address the above research questions. In Chapter 2, I consider the long-term optimization of detoxification of a single type of toxic compound using evolutionary
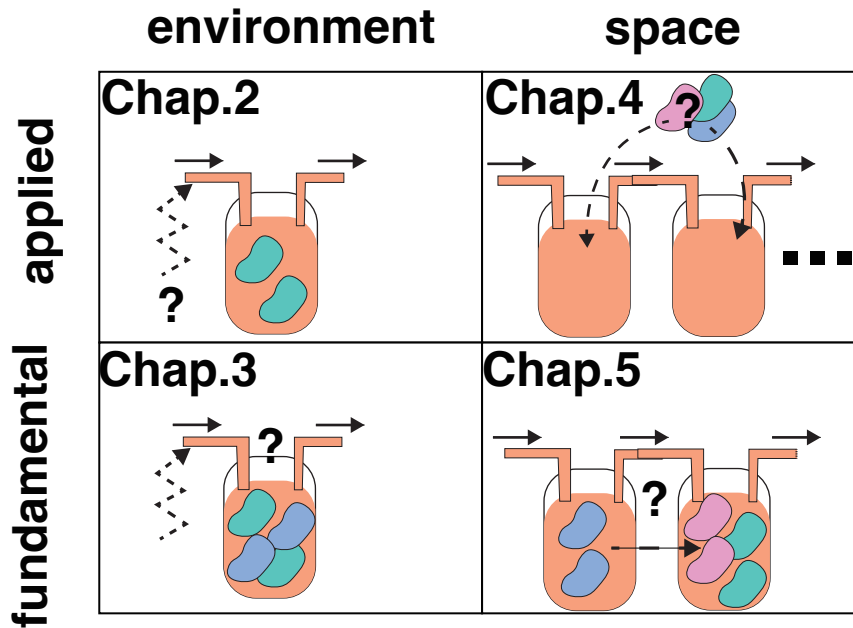
Figure 1.2 – Schematic representation of research questions and corresponding chapters

The left column corresponds to the introduction of the environment changes by fluctuating the media flowing into the chemostat. On the other hand, the right column shows the the introduction of a hierarchical (acyclic) spatial structure using multi-stage chemostats where microbes are allocated. In the top row, I address research questions in applied ecology and evolution. In the bottom panel, I ask questions related to fundamental ecology. In each panel, different species are represented by different colors, and the solid arrows indicate the flow of media in the (multi-stage) chemostats. The research questions are as follows. Chapter 2: how does one maximize community function by fluctuating the environment (the zigzag dashed arrow)? Chapter 3: how do environmental fluctuations and demographic noise affect species composition and diversity? Chapter 4: How can one maximize community functions by allocating species (dashed curve arrows) to hierarchical spatial structures? Chapter 5: how do upstream species affect the stability of the downstream community (the dashed straight arrow)?

game theory. This model shows that changing the in-flowing toxin concentration enables us to introduce strains that can degrade the toxin even when non-detoxifying strains spread in the population. Then, we can see how often we should re-introduce the detoxifiers while changing the inflow toxin concentration to maximize detoxification efficiency. Chapter 3 addresses research question 2 by implementing environmental fluctuations and demographic noise using the Gillespie algorithm (Gillespie, 1977). First, this chapter analyzes the competition between two species and shows that the probability that the species with the slower growth rate outcompetes the other species (i.e, exclusion of the fittest) shows various patterns over the rate of environmental fluctuations depending on the parameters related to the intensity of demographic noise. In multi-species communities, heterogeneity of communities (beta diversity) in simulations also shows various patterns, which can be predictable from pairs of species. Chapter 4 analyzes the optimal allocation problem in a multi-stage chemostat. In the main text of this chapter, I address research question 3b in the chain networks because this network structure maximizes community function in many cases. I also show examples of the optimal allocations using experimental data

from the lab in Chapter 4. For the analysis of the network structure of the multi-stage chemostats (research question 3a), see Appendix C.3. In Chapter 5, I ask research question 4 in the simplest chain network, where we have only two patches: upstream and downstream. This can be seen as a simplified gut system and I ask how we can stabilize the downstream community. Statistical analyses show three ideas to stabilize the downstream gut microbiota: imitating communities of healthy hosts that can be assembled over time, modifying species migration from outside of the system, and increasing the strength of positive interactions from the upstream to the downstream community. Chapter 6 synthesizes these results and discusses future research directions.

Chapter 2

Optimizing bioremediation in evolutionary time scales

2.1 Brief summary

In this chapter, I analyze the simplest problem of optimizing microbial community functions: degrading a single type of toxins using single species. Some microorganisms can degrade compounds that are toxic both to human beings and microorganisms themselves. For example, *Pseudomonas aeruginosa* can degrade heavy metals. Previous studies, however, show that mutants that do not degrade such toxins can spread in the populations because the mutants can receive the benefits of detoxification by the resident individuals (detoxifiers). In this chapter, I build a mathematical model of bioremediation in a chemostat based on evolutionary game theory. This model analyzes the evolution of the following two traits: degradation of the toxin (i.e., public resistance), and private resistance to the toxin, which does not change the toxin concentration in the system. If two strains have identical levels of private resistance to the toxin, the non-detoxifiers outcompete the detoxifiers. In contrast, the difference in private resistance allows the detoxifiers to outcompete or to coexist with the non-detoxifiers under a given toxin concentration. This result indicates that bioremediation can recover by manipulating the concentration of toxins flowing into the system. In addition, this study shows how often we should change the toxin concentration flowing into the chemostat to maximize long-term detoxification efficiency.

2.2 Author contributions

This project was conceived by Shota Shibasaki (SS) and Sara Mitri (SM). The analyses of the mathematical model were performed by SS. The manuscript of Shibasaki and Mitri (2020) was written by SS and SM. This article appears in the next pages.

ORIGINAL ARTICLE

Controlling evolutionary dynamics to optimize microbial bioremediation

Shota Shibasaki (iD) | **Sara Mitri** (iD)

Department of Fundamental Microbiology, University of Lausanne, Lausanne, Switzerland

Correspondence
Sara Mitri, Department of Fundamental Microbiology, University of Lausanne, Lausanne, Switzerland.
Email: sara.mitri@unil.ch

Abstract

Some microbes have a fascinating ability to degrade compounds that are toxic for humans in a process called bioremediation. Although these traits help microbes survive the toxins, carrying them can be costly if the benefit of detoxification is shared by all surrounding microbes, whether they detoxify or not. Detoxification can thereby be seen as a public goods game, where nondegrading mutants can sweep through the population and collapse bioremediation. Here, we constructed an evolutionary game theoretical model to optimize bioremediation in a chemostat initially containing "cooperating" (detoxifying) microbes. We consider two types of mutants: "cheaters" that do not detoxify, and mutants that become resistant to the toxin through private mechanisms that do not benefit others. By manipulating the concentration and flow rate of a toxin into the chemostat, we identified conditions where cooperators can exclude cheaters that differ in their private resistance. However, eventually, cheaters are bound to invade. To overcome this inevitable outcome and maximize detoxification efficiency, cooperators can be periodically reinoculated into the population. Our study investigates the outcome of an evolutionary game combining both public and private goods and demonstrates how environmental parameters can be used to control evolutionary dynamics in practical applications.

KEYWORDS
chemostat, detoxification, eco-evolutionary feedback, evolutionary game theory, public goods game

1 | INTRODUCTION

Microbes can be detrimental or vital to our health, our environment, and our economy. Much of applied microbiology strives to control these species, by promoting the growth of beneficial species, and suppressing that of harmful ones. We have achieved huge breakthroughs over centuries in eliminating pathogens and preventing dangerous diseases in humans, animals, and plants. At the same time, microbes have played important roles in enhancing agriculture and food production (Wolfe & Dutton, 2015), and more recently in the production of biofuels or other chemicals (Antoni, Zverlov, & Schwarz, 2007; Giri, Shitut, & Kost, 2020; Quin & Schmidt-Dannert, 2014; Ryan Georgianna & Mayfield, 2012), and in the degradation or "bioremediation" of toxic compounds, such as heavy metals or waste water (Atashgahi et al., 2018; Bertrand et al., 2015; Dixit et al., 2015; Kang, Kwon, & So, 2016; Zaccaria et al., 2020).

Despite these exciting advances, these approaches remain only partially successful. In medical microbiology, the emergence of resistant pathogens has led to a life-threatening global health crisis (Tacconelli et al., 2018). On the engineering side, we lack sufficient understanding to maximize the benefits gained from microbes, in terms of production rates or the efficiency of degradation of toxic compounds (Giri et al., 2020).

In all the examples above, there are two obstacles for controlling microbes. First, microbes' response to changes in their environment and to the behavior of neighboring cells may cause the extinction of species that contribute to community function or that protect against pathogens (ecological instability). Second, their large population size and short generation times mean that microbes quickly acquire mutations that can lead to evolutionary instability. Species can thereby gain resistance to antibiotics or lose their ability to perform a desired community function (Akita & Kamo, 2015; Bull & Barrick, 2017; Kumar, Maschke, Friehs, & Schügerl, 1991; Rugbjerg, Myling-Petersen, Porse, Sarup-Lytzen, & Sommer, 2018; Rugbjerg, Sarup-Lytzen, Nagy, & Sommer, 2018). In the bioremediation of heavy metals, for example, mutants that do not degrade these harmful compounds may invade the original population (Ellis, Lilley, Lacey, Murrell, & Godfray, 2007; O'Brien, Hodgson, & Buckling, 2014) and exclude it. To optimize the functional efficiency of a microbial system, therefore, we need to consider both ecological and evolutionary dynamics of microbial populations (Schuster et al., 2010).

In this study, we focus in on a bioremediation problem and develop a mathematical model to investigate the simple case of a single species degrading a toxic compound in a chemostat. We first explore its ecological and evolutionary stability and, second, use this knowledge to optimize the efficiency of toxin degradation over time. In our model, we assume that toxins are harmful to the microbes (e.g., heavy metals), who degrade them by secreting a costly product (e.g., extracellular enzymes). The benefit of degrading toxins is shared by the whole population as, for simplicity, our model does not include spatial structure. Toxin degradation can therefore be regarded as a public goods game (Broom, Pattni, & Rychtář, 2018; Samuelson, 1954), as defined in microbiology (Hummert et al., 2014; Smith & Schuster, 2019; West, Diggle, Buckling, Gardner, & Griffin, 2007). It is important to note that not all bioremediation systems correspond to public goods games (e.g., Röling et al., 2002; Smith, Graham, & Cleland, 1998), but here we focus on a subset of these systems where the compound to be degraded is toxic to the microbes and its degradation is costly. In such a system, the evolutionary instability of bioremediation is expected (Ellis et al., 2007; O'Brien et al., 2014).

In our model, microbes can adopt one of four strategies: They can be product secretors that pay a cost to contribute to the public good (cooperators) or nonsecretors that do not (cheaters) (see O'Brien et al. (2014) for an empirical example). In addition, microbes can be sensitive to the toxins or can acquire resistance, for example, by activating efflux pumps to expel toxins from within the cell (Blair, Webber, Baylay, Ogbolu, & Piddock, 2015; Bottery, Wood, & Brockhurst, 2016; Rojo-Molinero, Macià, & Oliver, 2019), or thickening the cell wall. In essence, public good secretion can also be seen as a form of

extracellular resistance to the toxin. In other words, here we consider toxin resistance through private or public means, whereby a cell benefits only itself or also the remaining population, respectively.

The population and evolutionary dynamics are then analyzed using evolutionary game theory, where a strategy is considered to be evolutionarily stable if it is not invaded by mutants with another strategy (Maynard Smith & Price, 1973). Evolutionary game theory typically considers the frequencies of strategies (i.e., frequencies sum to one), for example, in the replicator dynamics (Cressman & Tao, 2014). Here, however, since toxin concentration decreases with the absolute number or density of degrader microbes (cooperators), our model describes the dynamics of the densities of strategies (as in Hauert, Holmes, and Doebeli (2006); Hauert, Wakano, and Doebeli (2008); Gokhale & Hauert, 2016). And since the microbes' death rate depends on toxin concentration in the environment and their resistance level, our model also includes environmental feedback (as in Gong, Gao, and Cao (2018); Tilman, Plotkin, and Akçay (2020); Weitz, Eksin, Paarporn, Brown, and Ratcliff (2016)), where each strategy affects the environment differently, and the changing environment affects the fitness of each strategy differently.

We use this model to derive a protocol for optimal toxin degradation. We first show that cooperators that secrete toxin-degrading enzymes are excluded by cheaters that do not, if they have the same level of resistance to the toxin. This recapitulates a well-known result that can be explained by the tragedy of the commons (Hardin, 1968). However, we then show that cooperators can invade a population of cheaters if their level of toxin resistance is different. Since we assume that cheaters are unlikely to quickly acquire double mutations leading to cooperators with a different resistance level, maintaining degradation is only possible if we periodically inoculate these cooperators back into the chemostat. The success of this approach relies on the ability of cooperators to invade cheaters of different resistance. We then calculate the values of the experimentally controllable parameters (inoculation probabilities of cooperators, chemostat dilution rates, and inflowing toxin concentrations) that maximize the cumulative efficiency of detoxification.

In sum, our model combines population dynamics, evolutionary dynamics, and environmental feedback to optimize a population-level function. Integrating ecology and evolution into microbial public goods games is increasingly appreciated in microbial applications (Moreno-Fenoll, Cavaliere, Martínez-García, & Poyatos, 2017; Sanchez & Gore, 2013). And while optimizing bioremediation is the case study we are considering here, our approach of controlling evolutionary dynamics by changing environmental parameters can be applied to many other microbial functions.

2 | MODEL

In our scenario (Figure 1a), cells can take on one of four strategies depending on whether they produce the enzymes that degrade the toxin (cooperate) or not (cheat), and whether the cells are resistant to the toxin (resistant) or not (sensitive): sensitive cooperator (sCo),
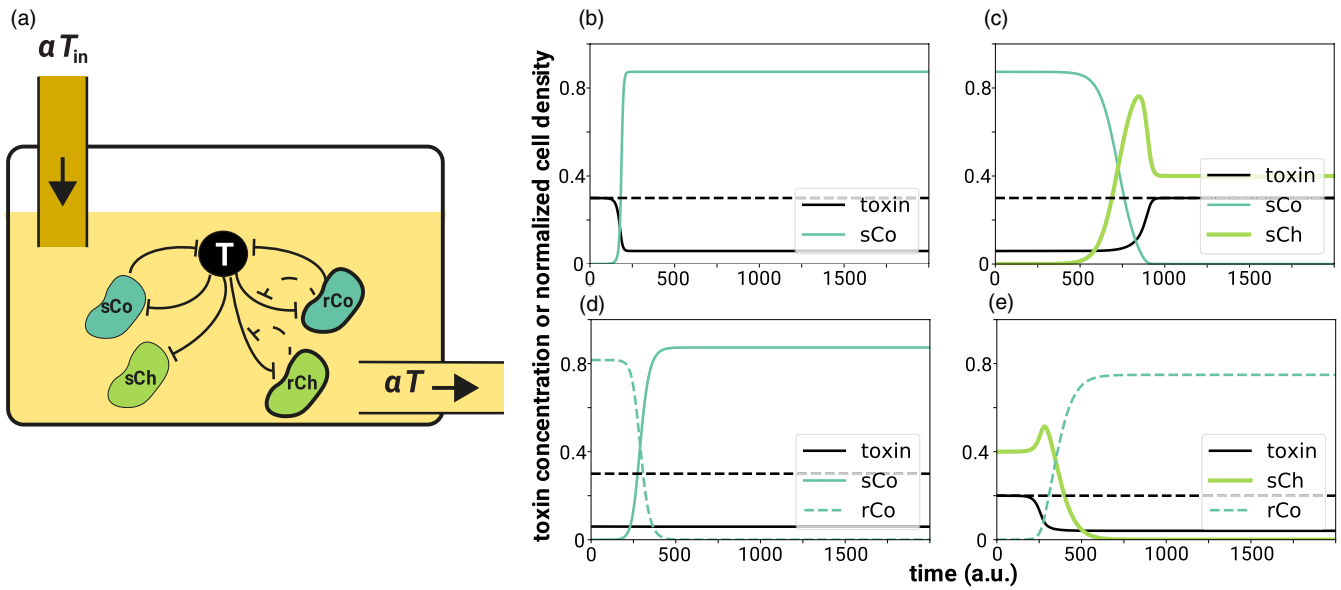
FIGURE 1 Schematic illustration of the model and examples of the dynamics. (a) In our scenario, a fluid with toxin concentration T_{in} flows into the chemostat, while the same amount of fluid with toxin concentration T flows out. The dilution rate of the chemostat is α. Each cell can exhibit one of four strategies: sensitive cooperator (sCo), resistant cooperator (rCo), sensitive cheater (sCh), or resistant cheater (rCh). Cooperators produce enzymes that degrade the toxin, while cheaters do not. The toxin kills cells depending on its concentration, but resistant cells have a lower death rate compared to sensitive cells. Whether the cells are cooperators or cheaters is independent of their resistance level and vice versa. (b–e) Examples of the dynamics in the absence of mutation are shown (a.u. = "arbitrary units"). In each panel, the black solid line represents the toxin concentration T and the dashed black line the toxin concentration flowing into the chemostat T_{in}. Detoxification efficiency at each time-point is proportional to the vertical distance between dashed and solid black lines. Other colored lines represent the cell densities of one of the four strategies (solid dark green: sCo, dashed dark green: rCo, thick lime green: sCh). (b) sCo grows and degrades the toxins. (c) sCo is invaded and excluded by sCh. (d) rCo is invaded and excluded by sCo. (e) sCh is invaded and excluded by rCo. Note that the initial conditions in (c–e) are the stable equilibria of the mono-culture of the resident strategies, while in (b) we begin with a low density of sCo and $T(0) = T_{in}$. Parameter values are $\alpha = 0.1$, $T_{in} = 0.2$ (e) or 0.3 (otherwise), $f_{max} = 0.5$, $K_d = 0.2$, $r = 1$, $c_d = 0.15$, $c_r = 0.2$ (e) or 0.3 (otherwise), $d_{max} = 1$, $K_s = 0.2$ (e) or 0.3 (otherwise), $K_r = 0.6$, and $n = 1$ (e) or 3 (otherwise)

sensitive cheater (sCh), resistant cooperator (rCo), and resistant cheater (rCh).

We begin by defining the bacterial population dynamics in our system. The dynamics of the density x of each strategy $i \in \{sCo, sCh, rCo, rCh\}$ in a chemostat are defined by growth, death, and dilution out of the chemostat:

[Correction added on 13 August 2020, after first online publication: The equation 1 has been corrected.]

$$\frac{dx_i}{dt} = x_i \left[r_i \left(1 - \sum_j x_j \right) - \delta_i(T) - \alpha \right], \quad (1)$$

where r_i is the intrinsic growth rate of strategy i due to nutrients that are not explicitly defined in the model, $\delta_i(T)$ is the death rate of strategy i given toxin concentration T, and α is the dilution rate. The total densities of the four strategies $\sum_i x_i$ should be lower or equal to one in Equation (1); that is, the carrying capacity is equal to one in the absence of death or dilution. In this formulation, a useful proxy for fitness is the ratio between intrinsic growth and death at a given toxin concentration, whether death is by toxin or by dilution:

$$W_i(T) \equiv \frac{r_i}{\delta_i(T) + \alpha}. \quad (2)$$

At an equilibrium, $W_i(T) = 1 / \left(1 - \sum_j x_j \right)$ should be satisfied for any strategy i that exists in the chemostat ($x_i > 0$). In addition, when $W_i(T) > W_j(T)$, strategy i increases faster or decreases slower than strategy j. For simplicity, this basic model assumes that strategies cannot mutate into each other. We extend it to include mutations in Appendix S5.

First, the intrinsic growth rates in this model differ depending on the costs each strategy pays. Cooperators pay a cost, c_d, for producing degrading enzymes, which are regarded as a public good since they reduce environmental toxicity and the death rate of all cells independently of their strategy. In addition, toxin resistance carries a cost, c_r. Such fitness costs, where resistant cells have lower fitness than sensitive ones in the absence of toxins, have been observed in many species (Andersson & Hughes, 2010; Andersson & Levin, 1999; San Millan & MacLean, 2019). In contrast to the production of degrading enzymes, however, where all cells benefit from decreased toxicity, the evolution of resistance can be regarded as an investment into a private good, where only the resistant cells themselves benefit. Assuming that the costs are additive, the intrinsic growth rate r_i of each strategy is defined as follows:

$$r_{sCo} = r(1 - c_d) \quad (3a)$$

$$r_{sCh} = r \tag{3b}$$

$$r_{rCo} = r\left(1 - c_d - c_r\right) \tag{3c}$$

$$r_{rCh} = r\left(1 - c_r\right), \tag{3d}$$

where r is the maximum intrinsic growth rate.

Cellular death rate $\delta_i(T)$ increases with toxin concentration T, and is represented by a Hill equation as is common in models of death by drugs (Chou, 2006):

$$\delta_i\left(T;K\right) = d_{max}\frac{T^n}{T^n + K_i^n}, \tag{4}$$

where d_{max} is the maximum death rate, K_i is the half maximal toxin concentration of strategy i, and n is the Hill coefficient, which determines the steepness of the function. Resistance can be modeled either by increasing K_i or decreasing n (Sampah, Shen, Jilek, & Siliciano, 2011). Here, we assume that resistant cells have a larger K than sensitive cells $K_r > K_s$, such that they reach d_{max} at a higher toxin concentration than the sensitive cells. Note that the toxin concentration T changes over time, as described below.

Due to the dilution in the chemostat, a proportion of cells of each strategy i flows out of the chemostat. The dilution rate into and out of the system is denoted by α

To describe the population dynamics of each strategy, however, it is necessary to also formulate the dynamics of the toxin concentration because it affects the microbes' death rate, and because the toxin concentration changes over time as cooperators detoxify it. The dynamics of the toxin concentration T in the chemostat are defined by the concentration flowing into and out of the chemostat, and detoxification by cooperators:

$$\frac{dT}{dt} = \alpha T_{in} - f\left(x_{Co}\right)T - \alpha T, \tag{5}$$

where T_{in} is the toxin concentration flowing into the chemostat, and $f(x_{Co})$ is the degradation rate, which is assumed to follow a Michaelis–Menten function:

$$f\left(x_{Co}\right) = f_{max}\frac{x_{Co}}{x_{Co} + K_d}, \tag{6}$$

where $x_{Co} = x_{sCo} + x_{rCo}$, that is, the sum of sensitive and resistant cooperators. Whether cooperators are resistant or sensitive has no impact on toxin degradation. K_d represents the density of cooperators x_{Co} that gives half the maximum of $f(x_{Co})$. All parameters of the model are listed in Table 1

TABLE 1 List of parameters

Notation	Range	Description
α	(0, 1]	dilution rate of the chemostat
r	(0, 1]	maximum intrinsic growth rate of the microbe
c_d	(0, 1]	cost of cooperation (production of the degrading enzyme)
c_r	(0, 1]	cost of resistance to the toxin
d_{max}	(0, 1]	maximum death rate by toxin
K_s	[0, K_r)	half maximal effective toxin concentration of the sensitive cells
K_r	(K_s, 1]	half maximal effective toxin concentration of the resistant cells
n	(0, ∞]	Hill coefficient
T_{in}	(0, 1]	toxin concentration flowing into the system
f_{max}	(0, 1]	maximum degradation rate of the toxin
K_d	(0, 1]	half maximal effective cooperator density of the degradation rate
μ_1	[0, 1]	mutation probability in the function of detoxification
μ_2	[0, 1]	mutation probability in the resistance level
m_1	[0, 1]	inoculation probability of sCo
m_2	[0, 1]	inoculation probability of rCo

As the goal of this study is to maximize the efficiency of detoxification, we define detoxification efficiency ϕ as the difference between the toxin concentration flowing into and out of the chemostat multiplied by the dilution rate:

$$\phi\left(\alpha, T_{\text{in}}, T\right) = \alpha\left(T_{\text{in}} - T\right). \tag{7}$$

With this definition, ϕ is proportional to the amount of detoxified liquid and is composed of the degree of detoxification and the amount of liquid flowing out of the chemostat. Although this equation gives the efficiency at any time t, we mainly focus on the efficiency at an equilibrium.

3 | RESULTS

3.1 | All strategies can persist in mono-cultures with no mutation

We first analyze whether cooperators and cheaters can persist in mono-culture. Remember that only cooperators produce public goods that degrade the toxin and thereby increase the survival of all cells in the chemostat. When a few cooperators (either sensitive or resistant, $0 < x_i(0) \ll 1$, $i \in \{sCo, rCo\}$) are introduced into the system, they increase and converge to an equilibrium of positive density (Figure 1b) if and only if

$$r_i > \delta_i\left(T_{\text{in}}\right) + \alpha \Longleftrightarrow W_i\left(T_{\text{in}}\right) > 1, \tag{8}$$
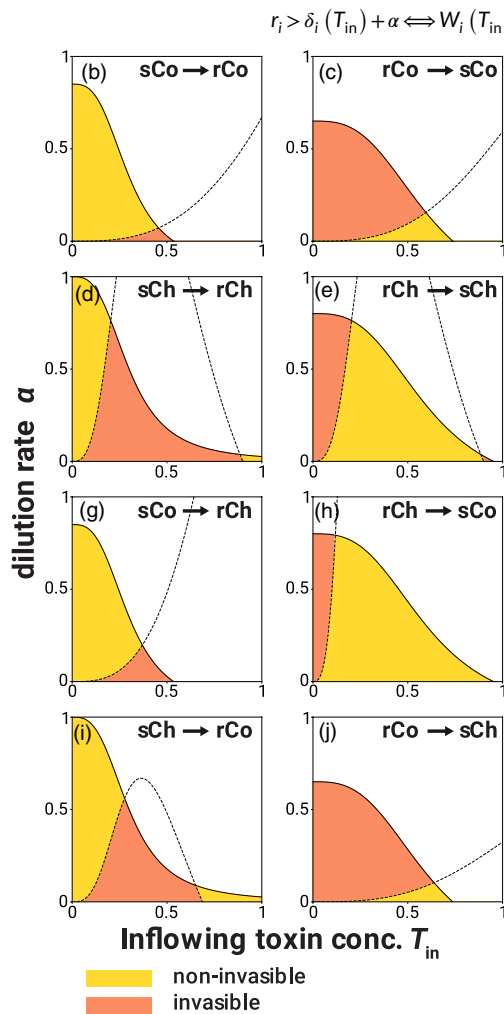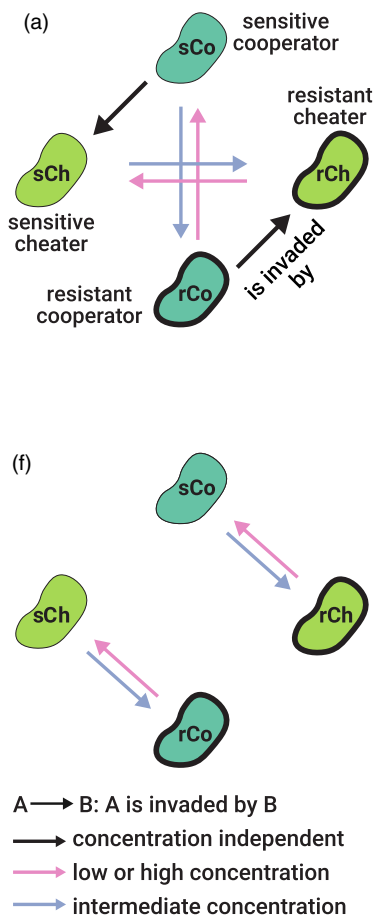
FIGURE 2 Short-term evolutionary dynamics of pairwise invasions. (a, f) Diagrams of pairwise invasion analysis when a single mutation (a) or double mutations (f) occur. A → B represents that A is invaded by B, and the color of each arrow shows the condition for successful invasion. Black arrows represent successful invasion regardless of the toxin concentration, while pink and blue arrows represent toxin concentration-dependent invasion (that invasion succeeds when the toxin concentration is low or high, or when the toxin concentrations are intermediate, respectively). (b–e, g–j) Examples of an invasion state space for each pair of strategies. A → B represents that A is a resident strategy and B is an invader. Pairs of sCo and sCh, and rCo and rCh are omitted since cheaters are fitter than cooperators in these pairs, regardless of the parameter values. In the yellow areas, the resident strategies are not invaded by the invaders, while the invasion succeeds in the orange areas. Each solid line is a boundary under which the resident strategy persists in mono-culture. The dashed line in each panel represents where the fitness proxy W of resident and invader strategies are equal at an equilibrium reached in a mono-culture of the resident strategy. Note that residents can coexist with invaders under certain conditions (see Appendix S3). Parameter values in (b–e, g–j) are $f_{\text{max}} = 0.5$, $K_d = 0.2$, $r = 1$, $c_d = 0.15$, $c_r = 0.2$, $d_{\text{max}} = 1$, $K_s = 0.3$, $K_r = 0.6$, and $n = 3$. Note that panels (b–e, g–j) are examples of the state space given the parameter values; different parameter values will show different invasion landscapes

where i is the focal strategy (see Appendix S1 for derivation). By solving $dT/dt = 0, dx_i/dt = 0$, one can find a trivial equilibrium ($x_i = 0$) and one or more nontrivial equilibria $(T^*, x_i^* > 0)$ that should satisfy:

$$T^* = \frac{\alpha T_{in}}{\alpha + f(x_i^*)}, \tag{9}$$

$$x_i^* = 1 - \frac{\delta_i(T^*) + \alpha}{r_i}, \tag{9b}$$

which we can calculate numerically using Newton's method.

In the absence of cooperators, we assume that the toxin concentration is equal to the incoming toxin because cooperators are the only degrader cells. In the case of a mono-culture of cheaters then, $T = T_{in}$ regardless of cell density, and their density at a stable equilibrium x_i^*, $i \in \{sCh, rCh\}$ is positive and given by Equation (9b) if inequality (8) holds (see Appendix S1 for details). From now on, we focus on conditions where cell density converges to a positive value in mono-culture (i.e., where inequality (8) holds).

3.2 | Cheaters can invade a population of cooperators

Next, we ask what happens when a cheater mutant invades a population of cooperators at its nontrivial equilibrium or vice versa.

Cheater mutants can invade and exclude a population of cooperators at any toxin concentration and independently of whether they are both sensitive or both resistant, as long as the resident and mutant have the same resistance levels (e.g., sCo and sCh) (Figures 1c, and 2a). In contrast, cooperators are unable to invade a population of cheaters of the same resistance level at any toxin concentration. These findings recapitulate the classical result that cheaters will always dominate in a well-mixed environment (West, Griffin, & Gardner, 2007) because cooperators pay a cost for producing degrading enzymes, but are as sensitive to the toxin as cheaters. In other words, the tragedy of commons (Hardin, 1968) occurs in this case.

3.3 | Toxin concentration determines invasion of sensitive and resistant cells

For mutants that differ in their private resistance level (e.g., sCo and rCo), the toxin concentration determines whether invasion succeeds or not (Figure 2b–e). Intuitively, this is because the benefit of being resistant to the toxin is quite low when its concentration is very low. Similarly, when toxin concentration is very high, the death rate of the resistant strain is close to that of the sensitive strain and too high to compensate for the cost of private resistance. Under these conditions, sensitive cells can invade a population of resistant ones (Figure 1D). Instead, resistant cells can invade a population

of sensitive cells when the toxin concentration in the chemostat is intermediate. Two strategies that differ only in their resistance level never stably coexist (see Appendix S3 for derivation).

3.4 | Cooperators can invade and coexist with cheaters of different resistance levels

Thus far, we have considered whether mutants can invade a resident population that differs in only one trait, their private or their public resistance (i.e., cooperative toxin degradation). While we assume that the time to reach an equilibrium following a single mutant invasion is shorter than the time for a second mutation to occur, we nevertheless explore invasions by such double mutants here (Figure 2f). Depending on the concentration of toxins, rCo and sCh can invade each other's populations, as can sCo and rCh (Figure 2g–j). This dependency on toxicity follows the same logic as for the invasion of a resistant mutant into a sensitive population described above: When the toxin concentration is intermediate, rCo has a much lower death rate than sCh, and the benefit of resistance exceeds the sum of the cost of cooperation c_d and resistance c_r (Figure 1e). If, on the other hand, the toxin concentration is either too low or too high, resistance to the toxin does not provide enough of an advantage to overcome its cost, leading instead to the invasion of sCh into a population of rCo. The same logic, albeit with different thresholds, can explain the invasion of sCo into rCh and vice versa (Figure 2g–j). In sum, invasions of double mutants into resident populations that differ in both public and private resistance depend on toxin concentrations (see Appendix S2).

Once a mutant has invaded, whether it will coexist with the resident population is unclear because, as cooperators increase, toxin concentration decreases, which changes the fitness landscape. In other words, increasing cooperator density can decrease the fitness difference between cooperators and cheaters. In Appendix S3, we show that cooperators and cheaters of different resistance levels (e.g., rCo and sCh) can indeed stably coexist at certain parameter ranges. Nevertheless, these two coexisting strains can then be invaded by cheaters with different resistance (e.g., rCh), which excludes the other two strategies (see Appendix S4). In other words, cooperators are never evolutionarily stable because they can be invaded and excluded by cheaters of the same resistance level, as we show next.

3.5 | In the long-term, cooperators are unlikely to be maintained

Having analyzed the outcomes of the invasion of all mutants in the short-term (Figure 2a,f), we can now predict how the population in the chemostat will change in the long-term. Regardless of which type we start with, as cells mutate, the population will transition between different genotypic "states," which can be represented by the state transition diagram in Figure 3. The probability of cooperators
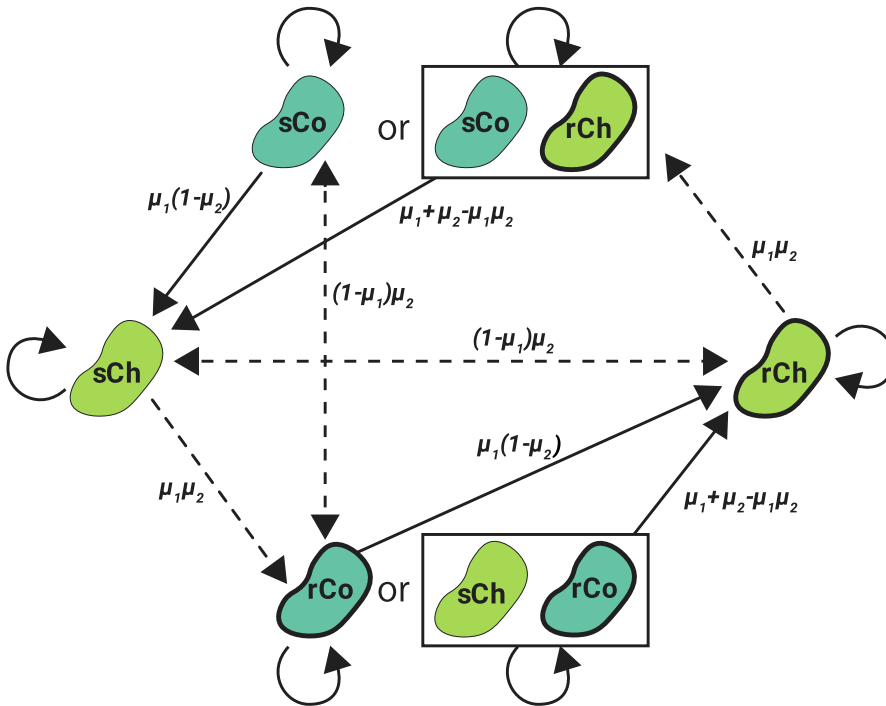
FIGURE 3 Schematic illustration of the state transitions for long-term evolutionary dynamics. Arrows represent state transitions resulting from natural selection. Solid arrows show transitions that are independent of toxin concentration, and dashed arrows transitions that depend on toxin concentration, here depicted for a given T_{in} and α. The transition from sCo to rCo and vice versa does not occur if cooperators coexist with cheaters. Values along the arrows represent state transition probabilities

mutating into cheaters or vice versa is given by μ_1, while μ_2 is the probability to change the level of resistance. For simplicity, we assume that these mutation probabilities are very small. Once a mutation occurs, we use Equation (1) as outlined above to take us to the following equilibrium state. We assume that no further mutations will occur before the equilibrium is reached, but relaxing this assumption does not alter the overall dynamics (Appendix S7). For example, sCh will appear in the population of sCo with probability μ_1 $(1–\mu_2)$ and exclude it. Then, rCh can appear in the population of sCh with probability $(1–\mu_1)$ μ_2, but may invade or not, depending on the toxin concentration (Figure 3).

Although in principle cooperators can invade a population of cheaters that differ in the level of resistance (e.g., rCo can invade sCh), (a) invasion success depends on α and T_{in}, (b) double mutations are expected to be rare ($\mu_1\mu_2$ is close to 0), and (c) if a cheater mutant of the same resistance as the cooperator invades (e.g., rCh), it will dominate the population and replace the cooperators. Accordingly, it is very difficult to maintain cooperators in the chemostat due to natural selection. This brings us to one of the main findings of the study: Even though cooperators and cheaters can coexist under some conditions, to maintain costly microbial detoxification, it is necessary to inoculate cooperators manually and to change α and T_{in} to favor their survival. Crucially, though, because cooperators are able to invade cheaters of opposite sensitivity, these inoculations can maintain cooperators—and thereby detoxification—in the short-term. In the following sections, we show how to control the values of α and T_{in} and inoculation probabilities to maximize the efficiency of detoxification.

3.6 | Culture conditions can be controlled to optimize detoxification efficiency

Ultimately, our goal is to maximize the efficiency of detoxification ϕ, which depends on the absolute abundance of the two types of cooperators in the chemostat. In turn, these abundances can be controlled by changing the culture conditions through two parameters: the dilution rate α and the toxin concentration flowing into the chemostat T_{in}.

To maximize the objective function in Equation (7), we consider three stable equilibrium states with different toxin concentrations flowing out of the chemostat: (a) $T = T_{in}$ when only cheaters are present, regardless of the values of α and T_{in}; (b) $T = T_{ij}^{\dagger}$ when cooperators i coexist with cheaters j, which have different resistant levels; and (c) $T = T_i^*$ when only one type of cooperators i is present. In the latter two cases, one can calculate the equilibria (analytically or numerically) and their corresponding detoxification efficiency ϕ for each culture condition (values of α and T_{in}). We can then find the optimal culture conditions that maximize this efficiency (Figure 4), although the equilibrium can be ecologically unstable for some parameter values.

Intuitively, the maximum efficiency is larger in a mono-culture of cooperators than in a co-culture of cooperators and cheaters of different levels of resistance (see Appendix S6). If cheaters can be excluded from the population by changing α and T_{in}, the optimal strategy for cultivation is (a) to exclude the cheaters by adjusting the culture conditions and then (b) to change the culture conditions to maximize the productivity of a mono-culture of cooperators.

FIGURE 4 Optimal detoxification at each equilibrium. The efficiency of detoxification at an equilibrium state $\phi(\alpha, T_{in}, T)$ given α and T_{in} is represented by color in each panel: (a) only sCo, (b) coexistence of sCo with rCh, (c) only rCo, and (d) coexistence of rCo with sCh. The red stars represent the maximum efficiency of detoxification in each. In the areas above the dashed gray lines in panels (a) and (c), the cooperator cannot persist in mono-culture (inequality 8). Parameter values are $f_{max} = 0.5$, $K_d = 0.2$, $r = 1$, $c_d = 0.15$, $c_r = 0.3$, $d_{max} = 1$, $K_s = 0.3$, $K_r = 0.5$, and $n = 1$ (a, b) or $n = 3$ (c, d). We used different values of n in the top and bottom rows to allow cooperators to coexist with cheaters with a different level of resistance
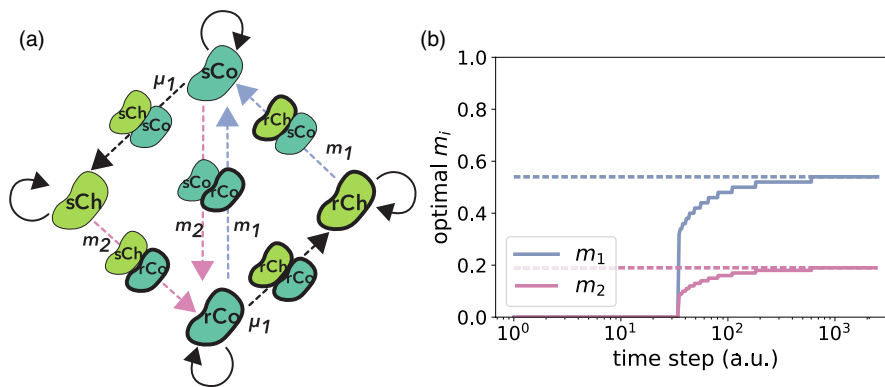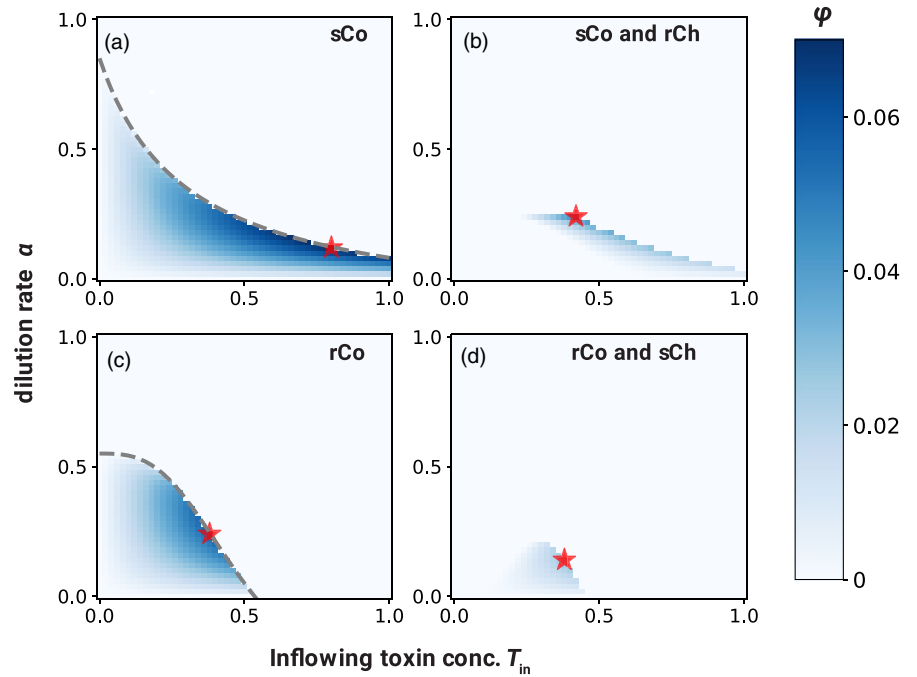
FIGURE 5 Optimization problem assuming that resistance mutations do not occur. (a) A simplified schematic illustration of the state transition diagram when $\mu_2 = 0$ and cooperators can exclude cheaters of different resistance levels. A →B represents a transition from state A to state B. Colored arrows represent transitions that occur through the introduction of sensitive and resistant cooperators in blue and pink, respectively. Dashed arrows indicate transient states where two strategies coexist. The full diagram containing 14 states is shown in Figure A.9. (b) The optimal values of m_1 and m_2 (the blue and pink solid lines, respectively) which maximize the cumulative efficiency defined by Equation (S.67) calculated using Dynamic Programming. The two dashed lines represent the values of m_1 and m_2 which maximize Equation (11). The initial state of the population is a mono-culture of sCo. Around 1,000 time steps, the optimal values of m_1 and m_2 for Equation (S.67) converge to the values which maximize Equation (11). In practice, the detoxification efficiency at each of the 14 states of the model as well the mutation probabilities would be experimentally measured and plugged into the model to calculate the values of m_1 and m_2 that would maximize cumulative efficiency. The plot above was generated using the following fictitious values, as an illustration: $\mu_1 = 0.01$, $\mu_2 = 0$, $\phi = \{0.4, 0.2, 0, 0.15, 0.3, 0.15, 0, 0.2, 0.35, 0.35, 0.3, 0, 0.2, 0\}$. See Appendix S7 for more detail

3.7 | Inoculating cooperators to optimize detoxification efficiency

Above, we showed that even though they are unlikely to appear by double mutation, cooperators can invade a population of cheaters if their level of resistance is different (Figure 3). Instead of waiting for these mutants to arise naturally, it would be more efficient to manually inoculate cooperators into the population, and to change α and T_{in} to allow them to invade successfully and to exclude the cheaters.

Assuming that we cannot observe the prevalence of each strategy at will, the problem is how often to inoculate sensitive or resistant cooperators to maximize detoxification efficiency over time. If cooperator inoculation probabilities are too small, cheaters will dominate the population, leading to a detoxification efficiency of zero. If they are too large, we may inoculate cooperators unnecessarily (e.g., sCo into a mono-culture of sCo) or when they cannot invade (e.g., sCo into a mono-culture of sCh). Such unfavorable inoculations can be costly because they can require a higher inflowing toxin

concentration T_{in}, and result in reduced detoxification efficiency for some time.

To optimize cooperator inoculation probabilities m_1 and m_2 for the sensitive or resistant cooperators, respectively, we consider population state transitions as a Markov chain with discrete time steps s and find the values of m_1 and m_2 that maximize the total amount of toxin degradation (Figure 5a and S9). Transitions in this Markov chain model can occur either due to mutations at probabilities $\boldsymbol{\mu} = (\mu_1, \mu_2)$ or inoculations at probabilities $\boldsymbol{m} = (m_1, m_2)$, resulting in the transition matrix $P(\boldsymbol{m};\boldsymbol{\mu})$. The state distribution vector $\pi(s) = (\pi_i(s))$ where $\pi_i(s)$ is the probability that the population is in state i at time step s ($\sum_i \pi_i(s) = 1$ for $s = 0, 1, \cdots, \infty$). Because the Markov chain is ergodic when cooperators can exclude cheaters that differ in their resistance level (Figure 5a, but see Appendix S7 for a case where cooperators cannot exclude cheaters), the probability distribution of the population states converges to a unique stationary distribution $\boldsymbol{\pi}^*$ in the limit of $s \to \infty$, regardless of the initial distribution $\boldsymbol{\pi}(0)$:

$$\boldsymbol{\pi}^*(\boldsymbol{m};\boldsymbol{\mu}) = \boldsymbol{\pi}^*(\boldsymbol{m};\boldsymbol{\mu}) P(\boldsymbol{m};\boldsymbol{\mu}). \tag{10}$$

Even though the transitions are probabilistic, we assume that the establishment of strategies following mutation or inoculation (i.e., short-term dynamics) is deterministic. Relaxing this assumption by introducing an establishment probability ε into the transition matrix $P(m;\mu)$ does not change the ergodicity of the Markov chain, and we

arrive at the stationary distribution in the same manner. In Appendix S7, we show how to calculate the expected cumulative efficiency of detoxification Φ from the beginning of the cultivation to time step s (defined in Equation (A.67)). This calculation is somewhat cumbersome, but for large s, Φ is approximately proportional (see Appendix S7) to the expected efficiency of detoxification at the stationary distribution $\boldsymbol{\pi}^*$:

$$\widehat{\Phi}(\mathbf{m};\boldsymbol{\mu}) \equiv \sum_i \phi_i \pi_i^*(\mathbf{m};\boldsymbol{\mu}). \tag{11}$$

To maximize the expected cumulative detoxification efficiency Φ, therefore, we can calculate $\boldsymbol{m}$ that maximizes Equation (11). In Box 1, we show an example of how to use this approach in practice.

4 | DISCUSSION

In this study, we have shown how to control the ecological and evolutionary dynamics of a microbial population growing in a chemostat in order to optimize the bioremediation of a toxic liquid. Public goods games where cooperators increase the growth rates of surrounding cells at a cost to themselves have been extensively studied, both empirically and theoretically (Allen, McNally, Popat, & Brown, 2016; Gokhale & Hauert, 2016; Griffin, West, & Buckling, 2004; Hauert

---

Box 1 An example of optimizing detoxification efficiency.

Imagine that we have set up the experimental system described, and would like to compute the optimal inoculation probabilities. We first need to define the Markov chain to make predictions, and second, we need to experimentally measure the parameters of our bacterial strains, in particular their degradation efficiency ϕ_i at each of the different states i of the Markov chain.

To establish the Markov chain, we begin with a few simplifying assumptions: (i) that $\mu_2 = 0$, such that cooperators can only invade a population of cheaters that differ in the level of resistance by inoculation, and (ii) that mutations and the manual inoculation of a cooperator strategy can occur only in a mono-culture (i.e., at most two strategies can exist simultaneously in the population). We further assume that the parameters are in a range where at certain (α, T_{in}), (iii) sCo and rCo can mutually exclude each other, and (iv) sCo and rCo can exclude rCh and sCh, respectively. Under these assumptions (we relax (i) below and (ii)–(iv) in Appendix S7), the Markov chain consists of at least 14 states: four mono-culture situations, six transient situations where two strategies coexist, and four situations where the introduction of cooperators is unfavorable (see Fig. S.9 for the diagram). A simplified schematic of this model is shown in Figure 5a.

By experimentally measuring mutation probabilities $\boldsymbol{\mu}$ and detoxification efficiencies $\boldsymbol{\phi} = (C\phi_i)$ for each state i, where C is a positive constant to change the time scale of ϕ_i into a discrete time step, we can calculate the probability distribution $\boldsymbol{\pi}(s, \boldsymbol{m}) = (\pi_i(s, \boldsymbol{m}))$ as a function of time step s and inoculation probabilities $\boldsymbol{m} = (m_1, m_2)$. Then, Dynamic Programming (DP) provides the optimal $\boldsymbol{m}$ that will maximize the cumulative expected detoxification efficiency Φ when the experiment finishes at time step s. Starting from a monoculture of sCo, the two solid lines in Figure 5b represent m_1 and m_2 provided by DP given some fictitious yet reasonable values of $\boldsymbol{\mu}$ and $\boldsymbol{\phi}$. At first, the optimal values of m_1 and m_2 are zero, because the state of the population is most likely to be a mono-culture of sCo, in which case inoculating cooperators would be pointless. However, as time passes, mutations will arise, and the population is likely to transition to a state of sCh mono-culture; then, the optimal values of m_1 and m_2 increase. At about 1,000 time steps, the optimal values of m_1 and m_2 converge to the values which maximize detoxification efficiency at the stationary distribution $\widehat{\Phi}$ described by Equation (11). When $\mu_2 > 0$ (relaxing assumption (i)), the number of states increases and the state transition diagram becomes more complex. As long as the Markov chain is ergodic, however, it is possible to find the stationary distribution $\boldsymbol{\pi}^*$ and the optimal values of m_1 and m_2 that maximize Equation (11). We show how to find the optima for nonergodic Markov chains in Appendix S7.

et al., 2006, 2008; Sanchez & Gore, 2013). Rather than increasing the growth rate of others, cooperators in our model degrade toxic compounds, which decreases the death rate of surrounding cells (O'Brien et al., 2014). This scenario enables us to introduce the evolution of resistance to the toxin, for example, through efflux pumps, as a private good, which we base on studies of drug–dose effect and resistance to it (Chou, 2006; Sampah et al., 2011). Unsurprisingly, cheaters always exclude cooperators with the same private resistance level because detoxification is costly (West, Griffin, et al., 2007). We show, however, that because the benefit of private resistance depends on the toxin concentration in the chemostat, cooperators can invade a population of cheaters that differ in their private resistance. The co-occurrence of two strains that differ both in their degradation ability as well as their resistance level is unlikely to suddenly arise by mutation, especially if we assume that mutations are rare. To maintain the degradation of toxins, therefore, it is necessary to periodically inoculate cooperators into the chemostat while changing the dilution rate and inflowing toxin concentration to guarantee invasion success.

Optimal values for these parameters (cooperator inoculation probabilities, dilution rate, and inflowing toxin concentration) that maximize the detoxification efficiency of the system can be calculated using our model. As input, the model requires experimental measurements of growth and death rates of the chosen microbe and its mutants (i.e., intrinsic growth rate of each strategy r_i, maximum death rate d_{max}, Hill coefficient n, median-effect toxin concentrations K_s, K_r, and degradation efficiencies ϕ of cooperators).

Our model and its results can also apply to problems other than bioremediation that involve survival in toxic environments. In essence, we are studying the evolutionary dynamics of public resistance (which is cooperative) and private resistance (which is not). Consider, analogously, two types of antibiotic resistance mechanisms: Public mechanisms are costly and benefit the producing cell as well as its neighbors such as extracellular secretion of antibiotic-degrading enzymes (e.g., β-lactamases (Yurtsev, Chao, Datta, Artemova, & Gore, 2013)), and private resistance mechanisms only benefit the producing cells, such as efflux pumps. The evolutionary dynamics in a scenario whereby cells can switch between these different resistance mechanisms and being sensitive to the antibiotic correspond to Figure 2a,f. In this case, however, an objective function would aim to minimize rather than maximize the densities of the most resistant strains. Another interesting aspect is that the benefits of resistance depend on toxin concentration in the chemostat, which is affected by the density of cooperators and the toxin concentration flowing into the chemostat. In other words, the public goods game affects the benefit of the private goods. This is why cooperators can invade a population of cheaters when they differ in their resistance level (Figure 2f).

Of course, our model relies on a number of assumptions and focuses only on a subset of possible bioremediation systems. First, we assume that only cooperators can detoxify. In reality, sensitive cells (cooperators or cheaters) may decrease toxin concentrations by passively absorbing them (Bottery et al., 2016). Including toxin absorption would not change our findings because (a) cooperators decrease toxin

concentration more effectively than cheaters with the same resistance levels, (b) the invasion analysis is still valid, and (c) we can still calculate the optimal inoculation probabilities as shown in Section 3.7. If toxin absorption by sensitive cells is significant, sCh could be better detoxifiers than rCo. In this case, it would be unnecessary to inoculate rCo to optimize the detoxification efficiency. Second, we only consider extracellular toxin degradation (e.g., by enzyme secretion), while toxins can also be degraded inside cells (O'Brien & Buckling, 2015). For intracellular degradation, a different functional form of detoxification $f(x_{Co})$ would be necessary, but we expect similar results as long as this function increases monotonically with the density of cooperators. Indeed, the invasion analysis is independent of the form of $f(x_{Co})$. Similarly, we assume that toxins kill the microbes and that their degradation does not contribute to growth. In reality, many compounds that are undesirable for humans are instead used as substrates by microbes (Atashgahi et al., 2018). This latter case is simpler than the one we consider here, since detoxification is no longer cooperative and there is no risk of cheaters arising and collapsing the system. Finally, detoxification may carry a negligible cost, for example, if it the toxic compound is neutralized by a change in pH, which occurs naturally due to a microbe's metabolism.

Another issue is how to define detoxification efficiency ϕ. Rather than Equation (7), one could, for example, define ϕ as the time needed for the toxins to decrease to a negligible concentration. This would change the optimal culture conditions α and T_{in}, but not the procedure to find the optimal introduction probabilities of cooperators m, which are independent of the formulation of ϕ. Our model also fixes some parameters, such as the Hill coefficient n, which can evolve in reality (Sampah et al., 2011). Similarly, the cost of resistance c_r can decrease over time due to compensatory evolution (Andersson & Hughes, 2010; San Millan et al., 2014). Allowing these parameters to evolve would make it more difficult for sCo to invade rCh because the relative fitness of rCh will increase.

We also assume that our system is well mixed and that there are no spatial gradients within the chemostat. Spatial structure, for example, whereby detoxifying enzymes diffuse slowly through the chemostat and have a patchy distribution can favor the coexistence of cooperators and cheaters (Allison, 2005). Indeed, previous empirical bioremediation studies have reported coexistence of cooperators with cheaters (Ellis et al., 2007; O'Brien et al., 2014). Theoretically, this may be due to a difference of resistance levels between cooperators and cheaters as we show here, but a simpler explanation would be the presence of spatial gradients. Relaxing the assumption of a perfectly well-mixed chemostat would make the persistence of cooperators easier. It may also increase the public benefit of toxin resistance, which we have considered to be private here (Rojo-Molinero et al., 2019).

Finally, there may be other ways of periodically introducing cooperators. Experimentally, our constant inoculation probabilities represent a situation where stock strains of cooperators would be manually added into the chemostat. If instead, multiple chemostats are running in parallel, another way of introducing cooperators would be to exchange certain amounts of fluids between chemostats. Ecologically, this would correspond to migration among patches,

and the optimal migration probabilities would depend on the probability distribution of the different strategies in each chemostat. Comparing the optimal introduction probabilities and the cumulative efficiency of detoxification between the model presented here and a multi-chemostat system is left for future work.

In summary, we have combined an ecological model with evolutionary game theory to develop a protocol for the control and optimization of a bioremediation system by microbes, and guard it against collapse through the emergence of cheaters. More broadly speaking, our scenario motivates the integration of important elements from ecological models, such as population densities and environmental feedback, into evolutionary game theory. In essence, our model can be adapted to any practical applications involving costly microbial traits, where manipulating environmental conditions can be used to control evolutionary dynamics. Achieving this will allow us to better anticipate evolutionary change in microbial systems that we strive to control, whether this involves increasing toxin degradation as we have shown here, the production of public goods such as useful chemicals, or eliminating antibiotic resistant pathogens.

CONFLICT OF INTEREST
There is no conflict of interest.

AUTHOR CONTRIBUTIONS
SS and SM conceived of the project, SS built and analyzed the model, and SS and SM wrote the manuscript.

DATA AVAILABILITY STATEMENT
Python code and csv files used for simulations and analytic calculations are available at https://github.com/ShotaSHIBASAKI/ControllingEvolutionadyDynamicsGitHub.

ORCID
Shota Shibasaki https://orcid.org/0000-0002-8196-0745
Sara Mitri https://orcid.org/0000-0003-3930-5357

REFERENCES
Akita, T., & Kamo, M. (2015). Theoretical lessons for increasing algal biofuel: Evolution of oil accumulation to avert carbon starvation in microalgae. *Journal of Theoretical Biology*, *380*, 183–191. https://doi.org/10.1016/j.jtbi.2015.05.027

Allen, R. C., McNally, L., Popat, R., & Brown, S. P. (2016). Quorum sensing protects bacterial co-operation from exploitation by cheats. *The ISME Journal*, *10*(JANUARY), 1706–1716. https://doi.org/10.1038/ismej.2015.232

Allison, S. D. (2005). Cheaters, diffusion and nutrients constrain decomposition by microbial enzymes in spatially structured environments. *Ecology Letters*, *8*(6), 626–635. https://doi.org/10.1111/j.1461-0248.2005.00756.x

Andersson, D. I., & Hughes, D. (2010). Antibiotic resistance and its cost: Is it possible to reverse resistance? *Nature Reviews Microbiology*, *8*(4), 260–271. https://doi.org/10.1038/nrmicro2319

Andersson, D. I., & Levin, B. R. (1999). The biological cost of antibiotic resistance. *Current Opinion in Microbiology*, *2*(5), 489–493. https://doi.org/10.1016/S1369-5274(99)00005-3

Antoni, D., Zverlov, V. V., & Schwarz, W. H. (2007). Biofuels from microbes. *Applied Microbiology and Biotechnology*, *77*(1), 23–35. https://doi.org/10.1007/s00253-007-1163-x

Atashgahi, S., Sánchez-Andrea, I., Heipieper, H. J., van der Meer, J. R., Stams, A. J. M., & Smidt, H. (2018). Prospects for harnessing biocide resistance for bioremediation and detoxification. *Science*, *360*(6390), 743–746. https://doi.org/10.1126/science.aar3778

Bertrand, J.-C., Doumenq, P., Guyoneaud, R., Marrot, B., Martin-Laurent, F., Matheron, R., ... Soulas, G. (2015). Applied microbial ecology and bioremediation. In J.-C. Bertrand, P. Caumette, P. Lebaron, R. Matheron, P. Normand, & T. Sime-Ngando (Eds.), *Environmental microbiology: Fundamentals and applications* (Chapter 16, 1st ed., pp. 659–753). Dordrecht, Netherlands: Springer, Netherlands. https://doi.org/10.1007/978-94-017-9118-2.

Blair, J. M., Webber, M. A., Baylay, A. J., Ogbolu, D. O., & Piddock, L. J. (2015). Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology*, *13*(1), 42–51. https://doi.org/10.1038/nrmicro3380.

Bottery, M. J., Wood, A. J., & Brockhurst, M. A. Selective conditions for a multidrug resistance plasmid depend on the sociality of antibiotic resistance. *Antimicrobial Agents and Chemotherapy*, *60*(4), 2524–2527. https://doi.org/10.1128/AAC.02441-15

Broom, M., Pattni, K., & Rychtář, J. (2018). Generalized social dilemmas: The evolution of cooperation in populations with variable group size. *Bulletin of Mathematical Biology*, *81*(11), 4643–4674. https://doi.org/10.1007/s11538-018-00545-1

Bull, J. J., & Barrick, J. E. (2017). Arresting evolution. *Trends in Genetics*, *33*(12), 910–920. https://doi.org/10.1016/j.tig.2017.09.008

Chou, T.-C. (2006). Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacological Reviews*, *58*(3), 621–681. https://doi.org/10.1124/pr.58.3.10

Cressman, R., & Tao, Y. (2014). The replicator equation and other game dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(Suppl 3), 10810–10817. https://doi.org/10.1073/pnas.1400823111

Dixit, R., Wasiullah, Malaviya, D. Pandiyan, K., Singh, U., Sahu, A., ... Paul, D. (2015). Bioremediation of heavy metals from soil and aquatic environment: An overview of principles and criteria of fundamental processes. *Sustainability*, *7*(2), 2189–2212. https://doi.org/10.3390/su7022189

Ellis, R. J., Lilley, A. K., Lacey, S. J., Murrell, D., & Godfray, H. C. J. (2007). Frequency-dependent advantages of plasmid carriage by *Pseudomonas* in homogeneous and spatially structured environments. *ISME Journal*, *1*(1), 92–95. https://doi.org/10.1038/ismej.2007.11

Giri, S., Shitut, S., & Kost, C. (2020). Harnessing ecological and evolutionary principles to guide the design of microbial production consortia. *Current Opinion in Biotechnology*, *62*, 228–238. https://doi.org/10.1016/j.copbio.2019.12.012

Gokhale, C. S., & Hauert, C. (2016). Eco-evolutionary dynamics of social dilemmas. *Theoretical Population Biology*, *111*, 28–42. https://doi.org/10.1016/j.tpb.2016.05.005

Gong, L., Gao, J., & Cao, M. (2018). *Evolutionary game dynamics for two interacting populations in a coevolving environment*. In 2018 IEEE Conference on Decision and Control (CDC), pp. 3535–3540, https://doi.org/10.1109/CDC.2018.8619801

Griffin, A. S., West, S. A., & Buckling, A. (2004). Cooperation and competition in pathogenic bacteria. *Nature*, *430*(7003), 1024–1027. https://doi.org/10.1038/nature02744

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*(June), 1243–1248.

Hauert, C., Holmes, M., & Doebeli, M. (2006). Evolutionary games and population dynamics: Maintenance of cooperation in public goods games. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1600), 2565–2571.

Hauert, C., Wakano, J. Y., & Doebeli, M. (2008). Ecological public goods games: Cooperation and bifurcation. *Theoretical Population Biology*, *73*(2), 257–263.

Hummert, S., Bohl, K., Basanta, D., Deutsch, A., Werner, S., Theißen, G., ... Schuster, S. (2014). Evolutionary game theory: Cells as players. *Molecular BioSystems*, *10*(12), 3044–3065.

Kang, C.-H., Kwon, Y.-J., & So, J.-S. (2016). Bioremediation of heavy metals by using bacterial mixtures. *Ecological Engineering*, *89*, 64–69. https://doi.org/10.1016/j.ecoleng.2016.01.023

Kumar, P. K. R., Maschke, H., Friehs, K., & Schügerl, K. (1991). Strategies for improving plasmid stability in bioreactors. *Trends in Biotechnology*, *9*(8), 279–284.

Maynard Smith, J., & Price, G. R. (1973). The logic of animal conict. *Nature*, *246*(5427), 15–18.

Moreno-Fenoll, C., Cavaliere, M., Martínez-García, E., & Poyatos, J. F. (2017). Eco-evolutionary feedbacks can rescue cooperation in microbial populations. *Scientific Reports*, *7*(1), 42561. https://doi.org/10.1038/srep42561

O'Brien, S., & Buckling, A. (2015). The sociality of bioremediation: Hijacking the social lives of microbial populations to clean up heavy metal contamination. *EMBO Reports*, *16*(10), 1241–1245.

O'Brien, S., Hodgson, D. J., & Buckling, A. (2014). Social evolution of toxic metal bioremediation in *Pseudomonas aeruginosa*. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1787), 20140858.

Quin, M. B., & Schmidt-Dannert, C. (2014). Designer microbes for biosynthesis. *Current Opinion in Biotechnology*, *29*(1), 55–61. https://doi.org/10.1016/j.copbio.2014.02.014

Rojo-Molinero, E., Macià, M. D., & Oliver, A. (2019). Social behavior of antibiotic resistant mutants within *Pseudomonas aeruginosa* biofilm communities. *Frontiers in Microbiology*, *10*(Mar): 570. https://doi.org/10.3389/fmicb.2019.00570

Röling, W. F. M., Milner, M. G., Jones, D. M., Lee, K., Daniel, F., Swannell, R. J. P., & Head, I. M. (2002). Robust hydrocarbon degradation and dynamics of bacterial communities during nutrient-enhanced oil spill bioremediation. *Applied and Environmental Microbiology*, *68*(11), 5537–5548. https://doi.org/10.1128/AEM.68.11.5537-5548.2002

Rugbjerg, P., Myling-Petersen, N., Porse, A., Sarup-Lytzen, K., & Sommer, M. O. (2018). Diverse genetic error modes constrain large-scale bio-based production. *Nature Communications*, *9*(1): 787. https://doi.org/10.1038/s41467-018-03232-w

Rugbjerg, P., Sarup-Lytzen, K., Nagy, M., & Sommer, M. O. A. (2018). Synthetic addiction extends the productive life time of engineered *Escherichia coli* populations. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(10), 201718622. https://doi.org/10.1073/pnas.1718622115

Ryan Georgianna, D., & Mayfield, S. P. (2012). Exploiting diversity and synthetic biology for the production of algal biofuels. *Nature*, *488*(7411), 329–335. https://doi.org/10.1038/nature11479

Sampah, M. E. S., Shen, L., Jilek, B. L., & Siliciano, R. F. (2011). Dose-response curve slope is a missing dimension in the analysis of HIV-1 drug resistance. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(18), 7613–7618. https://doi.org/10.1073/pnas.1018360108

Samuelson, P. A. (1954). The pure theory of public expenditure. *The Review of Economics and Statistics*, *36*(4), 387. https://doi.org/10.2307/1925895

San Millan, A., & MacLean, R. (2019). Fitness costs of plasmids: A limit to plasmid transmission. In F. Baquero, E. Bouza, J. Gutiérrez-Fuentes, T. Coque (Eds.), *Microbial Transmission*, Washington, DC: ASM Press. https://doi.org/10.1128/microbiolspec.MTBP-0016-2017

San Millan, A., Peña-Miller, R., Toll-Riera, M., Halbert, Z. V., McLean, A. R., Cooper, B. S., & Maclean, R. C. (2014). Positive selection and compensatory adaptation interact to stabilize non-transmissible plasmids. *Nature Communications*, *5*, 1–11. https://doi.org/10.1038/ncomms6208

Sanchez, A., & Gore, J. (2013). Feedback between population and evolutionary dynamics determines the fate of social microbial populations. *PLoS Biology*, *11*(4), e1001547. https://doi.org/10.1371/journal.pbio.1001547

Schuster, S., Kreft, J.-U., Brenner, N., Wessely, F., Theissen, G., Ruppin, E., & Schroeter, A. (2010). Cooperation and cheating in microbial exoenzyme production – Theoretical analysis for biotechnological applications. *Biotechnology Journal*, *5*(7), 751–758. https://doi.org/10.1002/biot.200900303

Smith, P., & Schuster, M. (2019). Public goods and cheating in microbes. *Current Biology*, *29*(11), R442–R447. https://doi.org/10.1016/j.cub.2019.03.001

Smith, V. H., Graham, D. W., & Cleland, D. D. (1998). Application of resource-ratio theory to hydrocarbon biodegradation. *Environmental Science and Technology*, *32*(21), 3386–3395. https://doi.org/10.1021/es9805019

Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., ... Zorzet, A. (2018). Discovery, research, and development of new antibiotics: The WHO priority list of antibiotic-resistant bacteria and tuberculosis. *The Lancet Infectious Diseases*, *18*(3), 318–327. https://doi.org/10.1016/S1473-3099(17)30753-3

Tilman, A. R., Plotkin, J. B., & Akçay, E. (2020). Evolutionary games with environmental feedbacks. *Nature Communications*, *11*(1), 915. https://doi.org/10.1038/s41467-020-14531-6

Weitz, J. S., Eksin, C., Paarporn, K., Brown, S. P., & Ratcliff, W. C. (2016). An oscillating tragedy of the commons in replicator dynamics with game-environment feedback. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(47), E7518–E7525. https://doi.org/10.1073/pnas.1604096113

West, S. A., Diggle, S. P., Buckling, A., Gardner, A., & Griffin, A. S. (2007). The social lives of microbes. *Annual Review of Ecology, Evolution, and Systematics*, *38*(1), 53–77. https://doi.org/10.1146/annurev.ecolsys.38.091206.095740

West, S. A., Griffin, A. S., & Gardner, A. (2007). Evolutionary explanations for cooperation. *Current Biology*, *17*(16), R661–R672. https://doi.org/10.1016/j.cub.2007.06.004

Wolfe, B. E., & Dutton, R. J. (2015). Fermented foods as experimentally tractable microbial ecosystems. *Cell*, *161*(1), 49–55. https://doi.org/10.1016/j.cell.2015.02.034

Yurtsev, E. A., Chao, H. X., Datta, M. S., Artemova, T., & Gore, J. (2013). Bacterial cheating drives the population dynamics of cooperative antibiotic resistance plasmids. *Molecular Systems Biology*, *9*, 683. https://doi.org/10.1038/msb.2013.39

Zaccaria, M., Dawson, W., Cristiglio, V., Reverberi, M., Ratcliff, L. E., Nakajima, T., ... Momeni, B. (2020). Designing a bioremediator: Mechanistic models guide cellular and molecular specialization. *Current Opinion in Biotechnology*, *62*, 98–105. https://doi.org/10.1016/j.copbio.2019.09.006

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

Chapter 3

Coupling environmental fluctuations with demographic noise

3.1 Brief summary

In the previous chapter, I analyzed how changing media affects the evolutionary dynamics of microorganisms and their detoxification efficiency in the chemostat model. In this chapter, I analyze how such environmental fluctuations affect microbial community dynamics, rather than community functions. When the environment becomes harsh, population sizes of microorganisms decrease and the intensity of demographic noise increases, which should affect species diversity in turn. In this project, I analyze how the coupling effect of environmental fluctuations with demographic noise affects the heterogeneity of the communities across many simulations (i.e., beta diversity). I first investigate competition between two species (a slower-growing and a faster-growing species) over the rate of environmental fluctuations and the species' toxin sensitivities, which affect population sizes and the intensity of demographic noise. Due to demographic noise, the slower-growing species can outcompete the faster-growing species with some probability, but the sensitivities to the toxin affect the patterns of how this probability changes over the rate of environmental fluctuations. In multi-species communities, beta diversity also shows similar patterns over the environmental fluctuation rate and the mean toxin sensitivities across community members. The effect of coupling environmental fluctuations with demographic noise on beta diversity can be predicted by pairwise analyses of community members.

3.2 Author contributions

This project was designed by SS, Mauro Mobilia (MM), and SM. The simulations and the statistical analyses were performed by SS. The first draft of Shibasaki et al. (2021) was written by SS, and revised by SS, MM, and SM. The accepted author manuscript of Shibasaki et al. (2021) appears in the next pages.

Exclusion of the fittest predicts microbial community diversity in fluctuating environments

Shota Shibasaki[1], Mauro Mobilia[2†], and Sara Mitri[1†]

[1]Department of Fundamental Microbiology, University of Lausanne, Switzerland

[2]Department of Applied Mathematics, School of Mathematics, University of Leeds, United Kingdom

[†]these authors contributed equally

E-mail addresses of all authors.

SS: shota.shibasaki@unil.ch

MM: M.Mobilia@leeds.ac.uk

SM: sara.mitri@unil.ch

Short running title: Exclusion of the fittest and diversity

Key words: competitive exclusion, environmental switching, demographic noise, chemostat, beta diversity, intermediate disturbance hypothesis

Type of article: article

Numbers of words: 187 words in abstract, and 7159 words in main text.

Number of references: 78.

Numbers of figures: 6, tables: 1, text box 0.

Corresponding authors:

SS Quartier UNIL-Sorge, Batiment Biophore, Office 2917.3 , CH-1015 Lausanne. Phone : + 41 21 692 56 00. E-mail: shota.shibasaki@unil.ch.

SM Quartier UNIL-Sorge, Batiment Biophore, Office 2414, CH-1015 Lausanne. Phone : +41 21 692 56 12. E-mail: sara.mitri@unil.ch.

Abstract

Microorganisms live in environments that inevitably fluctuate between mild and harsh conditions. As harsh conditions may cause extinctions, the rate at which fluctuations occur can shape microbial communities and their diversity, but we still lack an intuition on how. Here, we build a mathematical model describing two microbial species living in an environment where substrate supplies randomly switch between abundant and scarce. We then vary the rate of switching as well as different properties of the interacting species, and measure the probability of the weaker species driving the stronger one extinct. We find that this probability increases with the strength of demographic noise under harsh conditions and peaks at either low, high, or intermediate switching rates depending on both species' ability to withstand the harsh environment. This complex relationship shows why finding patterns between environmental fluctuations and diversity has historically been difficult. In parameter ranges where the fittest species was most likely to be excluded, however, the beta diversity in larger communities also peaked. In sum, how environmental fluctuations affect interactions between a few species pairs predicts their effect on the beta diversity of the whole community.

1 Introduction

Natural environments are not static: temperature, pH, or availability of resources change over time. Many studies in microbiology, ecology and evolution have focused on responses to fluctuations in resource abundance in the regime of feast and famine periods (Hengge-Aronis, 1993; Vasi et al., 1994; Srinivasan and Kjelleberg, 1998; Xavier et al., 2005; Merritt and Kuehn, 2018; Himeoka and Mitarai, 2019). These models capture the dynamics within many natural ecosystems. For example, the gut microbiome of a host is exposed to fluctuating resources that depend on its host's feeding rhythm, which may affect microbiota diversity (Cignarella et al., 2018; Li et al., 2017; Thaiss et al., 2014). In addition to their magnitude, environmental fluctuations can also differ in their time scales – for the gut microbiota, a host's feeding rhythm may vary from hourly to daily, or even monthly if feeding depends on seasonal changes (Davenport et al., 2014; Smits et al., 2017) – or in the type of substrates taken up, which can be nutritious or harmful for the microbiota.

How environmental fluctuations (EFs) affect species diversity has been a highly contested topic in ecology. Here, EFs refer to changes that are not caused by the organisms themselves, but nevertheless affect their dynamics (e.g., abiotic resource supplies). The intermediate disturbance hypothesis argues that intermediate intensity and frequency of disturbances maximize species diversity (Connell, 1978; Grime, 1973) because highly competitive species dominate at a low level of disturbance, while only species that have adapted to the disturbance dominate at high disturbance (Grime, 1977). This hypothesis is controversial (Fox, 2013) and other relationships between disturbance and species diversity have been reported both empirically and theoretically (Mackey and Currie, 2001; Miller et al., 2011).

Another framework that is used to predict species diversity under EFs is the modern coexistence theory (Chesson, 1994), which explains the maintenance of diversity through species' differing responses to and preferences for environmental conditions, which can vary over spatial and/or temporal scales through fluctuations (Amarasekare, 2019; Chesson, 2000a,b; Letten et al., 2018b; Barabás et al., 2018; Ellner et al., 2019). The modern coexistence theory divides environmental factors into those that are independent of species abundances (e.g., temperature) and those that depend on them (e.g., amounts of resources). The latter environmental factors mediate the sign and/or magnitude of interspecific interactions (Hoek et al., 2016; Piccardi et al., 2019; Zuñiga et al., 2019), and whether species tend to cooperate or compete can, in turn, drive community diversity and stability (Mougi and Kondoh, 2012; Coyte et al., 2015; Marsland III et al., 2019; Butler and O'Dwyer, 2018, 2020). Microbial communities often experience extreme environmental fluctuations that can alter interactions between species and hence affect species diversity (Rodríguez-Verdugo et al., 2019; Nguyen et al., 2020b).

Another important factor potentially influencing the outcome of interactions between species is demographic noise (DN) arising from randomness in birth and death events in finite populations. DN is negligible in large populations but strong in small ones, where it can lead to species extinction or fixation, and affect community composition (Roughgarden, 1979; Ewens, 2004). As EFs affect population sizes, they modulate the strength of demographic noise, leading to a coupling of EFs and DN. This interdependence has been understudied until recently, despite its potentially important consequences on eco-evolutionary dynamics (Wienand et al., 2017,

2018; West and Mobilia, 2020; Taitelbaum et al., 2020).

To understand how the interplay between EFs and DN affect community diversity, we set up a stochastic model of multiple species subject to a varying supply of nutrients and/or toxins. Our model then allows us to ask: How do EFs, coupled to DN, affect species interactions and diversity?

We include toxins in our model as they are important in natural communities, but often missed in similar models. Toxins that typically come to mind are pesticides or antibiotics (Pérez et al., 2005; Xu et al., 2011), but in principle they can be anything that inhibits microbial growth compared to some optimal condition. For example, oxygen is harmful to anaerobic microbes (Guittar et al., 2021) and bile acids are toxic to microbes in the human gastrointestinal tract (Ruiz et al., 2013; Molinero et al., 2019). These toxins can, however, be degraded by microbes. In the example of bile acids, *Lactobacillus* and *Bifidobacterium* strains degrade them by producing bile-salt hydolases and other extracellular enzymes (Ruiz et al., 2013; Molinero et al., 2019). Such degradation of toxic compound by microbes is expected to be quite common, as microbes will be selected to become resistant to toxins by diminishing environmental toxicity. In our model then, toxins contribute to environmental harshness and reduce population sizes, which can modulate the strength of DN. We also expect toxins to affect inter-species interactions, as species that absorb or degrade them can potentially facilitate the growth of others (Hoek et al., 2016; Piccardi et al., 2019; Zuñiga et al., 2019).

The next section describes our theoretical model, explaining how DN and EFs are implemented and how species interactions and diversity are measured (section 2). We first use it to explore how often one single species goes extinct due to environmental switching coupled with DN (section 3.1). We then add a second slower-growing species, still focusing on the behavior of the first, and ask how its extinction probability is affected by the slower-grower, and how different properties of the two species change the effects of the switching rate (section 3.2). We find that sensitivity to toxins increases the strength of DN because it decreases species abundances. At toxin sensitivities that are high enough to increase DN but low enough that either species is likely to survive, the slower-growing species can outcompete the fast-growing one (we call this phenomenon *exclusion of the fittest*, sections 3.3 and 3.4), a result that is a direct consequence of coupling DN and EFs. Finally, we expand our model to larger communities and show that our first analysis of the exclusion of the fittest species predicts beta diversity patterns in larger communities (section 3.5): community composition is the most diverse when strong species are most likely to be excluded. Section 4 discusses the importance of coupling EFs and DN on species interactions and diversity. Finally, we compare our results with the intermediate disturbance hypothesis and the modern coexistence theory. Technical and computational details are given in a series of Appendices.

2   Model & Methods

In order to investigate the interdependence of EFs and DN on species interactions and diversity, we study an idealized chemostat model that combines DN and environmental switching (Fig. 1). The chemostat is meant to represent natural ecosystems subject to in- and outflow, such as a gut or a river. Within the chemostat,

we consider a well-mixed population consisting of N species of bacteria, and $N/2$ types of resources as well as $N/2$ types of toxins (without loss of generality, N is assumed to be even). Our system can exhibit different levels of species richness at equilibrium depending on parameter values: only one species may persist as species interact with all resources and toxins, or all species may coexist as the environment has as many limiting factors (resources and toxins) as the number of species. The special case of a model with a single resource and a single toxin in a mono-culture is considered for completeness in Section 3.1. Although many models ignore the role of toxins, they play two important roles in our model: first, sensitivities to toxins changes the strength of DN because population sizes of sensitive species are low. Second, toxins can enable a pair of species to facilitate each other (Piccardi et al., 2019) because toxin absorbance or degradation decreases the death rate of other species. In this manuscript, we consider toxins that are not released from cells when they die, e.g., antibiotics that inhibit DNA or RNA synthesis (Cangelosi and Meschke, 2014), although the release of toxins from dead cells could be introduced with a small modification of our model (see Huang et al. (2013) for example). The absence of toxins is similar to having a low toxin sensitivity (Fig. A.22).
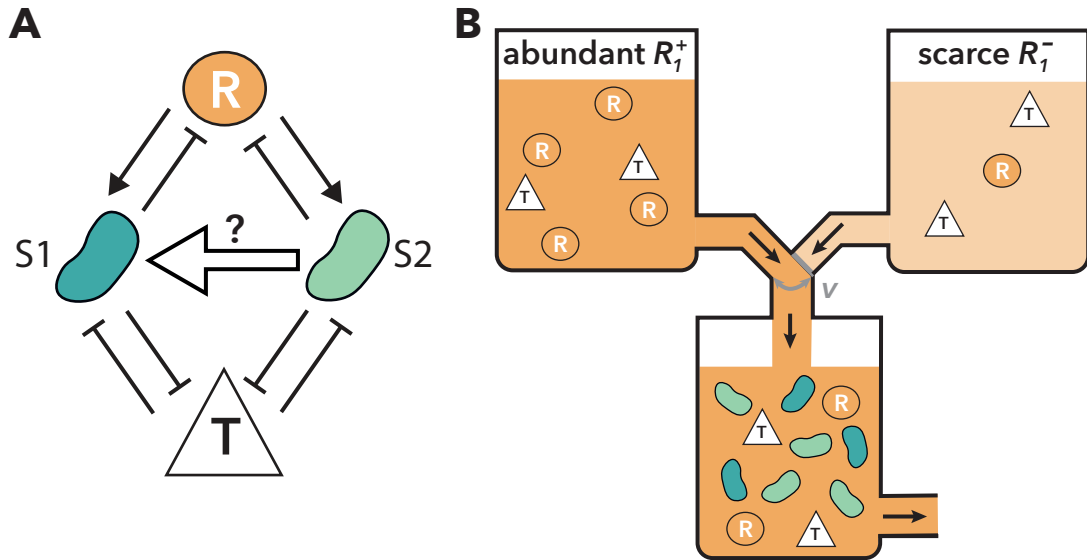
Figure 1: Schematic illustration of the model

A chemostat model with environmental switching. A: Interaction network when $N = 2$. A → B represents that A increases B while A ⊣ B represents that A decreases B. Two species compete for the same resource (R in a circle) but are killed by the same toxic compound (T in a triangle). As a proxy for species interactions, we follow the net effect of the slower-growing species 2 on species 1 (large arrow from species 2 to 1). B: An example of a chemostat model with environmental switching and $N = 2$. Environmental switching is realized by changing the media flowing into a chemostat. In this example, the current environmental condition is abundant resource supply R_1^+.

In our model, communities follow a continuous-time multi-variate birth-and-death process (see, e.g. Novozhilov et al. (2006); Allen (2010)) in a time-fluctuating binary environment (see, *e.g.*, Wienand et al. (2017, 2018); West and Mobilia (2020); Taitelbaum et al. (2020)). More specifically, we consider that at time σ the community consists of an amount $r_i(\sigma)$ of resources of type i ($i = 1, \ldots, N/2$), an amount $t_j(\sigma)$ of toxin of type j ($j = 1, \ldots, N/2$), and an abundance $s_k(\sigma)$ of microbial species k ($k = 1, \ldots, N$). Here, resources are assumed to be nutrients for all species, allowing them to grow at different rates, while the toxins kill all species at different

rates when species have different toxin sensitivities. hence the species extinction probability.

In a static environment (no EFs), this chemostat model evolves according to the birth-and-death process defined by the following "birth" and "death" reactions:

$$r_i \xrightarrow{\tau_{r_i}^+} r_i + 1, \tag{1a}$$

$$r_i \xrightarrow{\tau_{r_i}^-} r_i - 1, \tag{1b}$$

$$t_j \xrightarrow{\tau_{t_j}^+} t_j + 1, \tag{1c}$$

$$t_j \xrightarrow{\tau_{t_j}^-} t_j - 1, \tag{1d}$$

$$s_k \xrightarrow{\tau_{s_k}^+} s_k + 1, \tag{1e}$$

$$s_k \xrightarrow{\tau_{s_k}^-} s_k - 1, \tag{1f}$$

occurring with transition rates:

$$\tau_{r_i}^+ = \alpha R_i, \tag{2a}$$

$$\tau_{r_i}^- = \sum_{k=1}^{N} \frac{\mu_{ik}}{Y_{ik}^r} \frac{r_i}{r_i + K_{ik}^r} s_k + \alpha r_i, \tag{2b}$$

$$\tau_{t_j}^+ = \alpha T_j, \tag{2c}$$

$$\tau_{t_j}^- = \sum_{k=1}^{N} \frac{\delta_{jk}}{Y_{jk}^t} \frac{t_j}{t_j + K_{jk}^t} s_k + \alpha t_j, \tag{2d}$$

$$\tau_{s_k}^+ = \sum_{i=1}^{N/2} \mu_{ik} \frac{r_i}{r_i + K_{ik}^r} s_k, \tag{2e}$$

$$\tau_{s_k}^- = \sum_{j=1}^{N/2} \delta_{jk} \frac{t_j}{t_j + K_{jk}^t} s_k + \alpha s_k, \tag{2f}$$

where α is the dilution rate of the chemostat, $\xi = \pm 1$ (see below) represents changing environmental conditions in terms of in-flowing resource and/or toxin amount. $R_i(\xi)$ $(T_j(\xi))$ is resource i's (toxin j's) supply under the environmental condition ξ, Y_{ik}^r (Y_{jk}^t) is species k's biomass yield for resource i (toxin j), μ_{ik} is the maximum growth rate of species k by resource i, δ_{jk} is the maximum death rate of species k by toxin j (species k's sensitivity to toxin j), and K_{ik}^r (K_{jk}^t) is the amount of resource i (toxin j) that gives the half-maximum growth (death) rate for species k (see also Table A.1). These transition rates hence reflect that (i) the amounts of resources and toxins increase depending on the product of their in-flow concentrations and the dilution rate, (ii) the amounts of resources and toxins decrease with the dilution rate and with the consumption/absorption by species, (iii) the growth and death rates depend in Monod functional forms on the amounts of resources and toxins, respectively, (iv) the dilution rate α sets the time scale at which the environment attains the state of abundance or scarcity (after a time $\sim 1/\alpha$), see below and Appendix 2.2 and (v) all effects are additive. When

Table 1: Different scenarios of environmental switching

Scenario	$R_i\,(\xi=1)$	$R_i\,(\xi=-1)$	$T_j\,(\xi=1)$	$T_j\,(\xi=-1)$
1	R_i^+	R_i^-	$\langle T_j \rangle$	$\langle T_j \rangle$
2	$\langle R_i \rangle$	$\langle R_i \rangle$	T_j^-	T_j^+
3	R_i^+	R_i^-	T_j^-	T_j^+

R_i^+, R_i^-, and $\langle R_i \rangle$ represent abundant, scarce, and mean resource supplies, respectively. $R_i^+ > R_i^-$ and $\langle R_i \rangle = \left(R_i^+ + R_i^-\right)/2$ for $i = 1, \ldots, N/2$. Similar representation and relations hold for toxin supply T_j. In each condition, $\xi = 1$ ($\xi = -1$) means mild (harsh) environment, respectively.

R_i and T_j are constant, the environment is static and the birth-and-death process defined by Eqs (1a) – (2f) naturally accounts for the DN arising in the population, which is the sole source of noise.

We model EFs by considering a time-fluctuating binary environment resulting in R_i and/or T_j to be *dichotomous random variables*, i.e., $R_i = R_i(\xi(\sigma))$ and/or $T_j(\xi(\sigma))$, where $\xi(\sigma) = \pm 1$ represents the binary state of the environment ($\xi = 1$ represents a mild environment while $\xi = -1$ represent a harsh environment). We hence assume that R_i and/or T_j switch between two values at a rate ν, according to a time-continuous colored dichotomous Markov noise (random telegraph noise) (Bena, 2006; Horsthemke and Lefever, 2006)

$$\xi \xrightarrow{\nu} -\xi. \tag{3}$$

We call ν a "switching rate" because we implement a symmetrically switching environment as a simple example of EFs (but see Appendix 2.2 for a further discussion of this choice and Appendix 5 for other forms of EFs). Although we use arbitrary units in the model, the unit of the switching rate applies to all events. For example, if the unit for the dilution rate is per hour, $\nu = 10^0$ means that the environment switches every hour on average. We investigate three environmental switching scenarios, where either or both resource and toxin supplies fluctuate over time, see Table 1. In the main text, we focus on the scenario where only resource supplies switch between abundant and scarce supplies (R_i^+ and R_i^-, respectively, such that $R_i^+ > R_i^- > 0$) while the amounts of toxin supplies remain constant over time ($\langle T_j \rangle \equiv (T_j^+ + T_j^-)/2$); see Appendix 3 for the remaining scenarios. The initial resource supply in the main text is $R_i\,(\xi\,(0)) = R_i^+$ with probability 0.5 and otherwise $R_i\,(\xi\,(0)) = R_i^-$.

We assume that ξ switches symmetrically between the states ± 1 (see Taitelbaum et al. (2020) and Appendix 5 for the cases of asymmetric switching). In all our simulations, ξ is stationary and thus ξ has zero mean, $\langle \xi(\sigma) \rangle = 0$, autocorrelation $\langle \xi(\sigma)\xi(\sigma') \rangle = \exp\left(-2\nu|\sigma - \sigma'|\right)$, where $\langle \cdot \rangle$ denotes the ensemble average over the environmental noise, and finite correlation time $1/(2\nu)$. A great advantage of dichotomous noise is its simplicity: ξ is bounded (R_i and T_j are always well defined) and straightforward to simulate. This choice allows us to model suddenly changing environmental conditions, which reflect situations in microbial life such as exposure to resource or toxin oscillations that can be reproduced in lab-controlled experiments (Sunya et al., 2013). Other forms of EFs are also possible, e.g. ξ could be a Gaussian random variable, but then R_i and T_j would be unbounded and vary continuously and could take unrealistic values. Modeling EFs with Eq (3) is arguably the simplest biologically-motivated choice to couple fluctuations in resource/toxin supplies with demographic

noise, and allows us to investigate questions that are not specific to dichotomous noise, see Appendix 2.2. Interestingly, the analysis of the long-term dynamics of the two-species, one-resource-one-toxin model under symmetric dichotomous noise in Appendix 2.2 reveals that this simple form of environmental variability leads to distributions of the total population size n (i.e., the sum of species, resources and toxins abundances $n \equiv \sum_k s_k + \sum_i r_i + \sum_j t_j$) that varies greatly with the rate of switching relative to the dilution rate: when $\nu/\alpha \gg 1$ (fast switching), the total population size n is unimodal; when $\nu/\alpha \ll 1$ (slow switching), n is bimodal and fluctuates between two very different values; intermediate scenarios interpolating between unimodal and bimodal distribution of n arise when $\nu/\alpha \sim 1$ (intermediate switching), see Wienand et al. (2017, 2018); West and Mobilia (2020); Taitelbaum et al. (2020) and Fig. A.6. This results in an explicit coupling of DN to EFs in multi-species communities, via the modulation of the DN intensity by ν/α. This is a distinctive feature of our model comparing previous studies. For example, some models (Leigh, 1981; Engen and Lande, 1996; Kalyuzhny et al., 2015) do not couple DN and EFs, while other models (Engen and Lande, 1996; Kamenev et al., 2008; Chisholm et al., 2014; Fung et al., 2015) study the coupling of DN and EFs in single species scenarios. See Appendix 2.2 for more detailed discussion and analysis.

The master equation for this model is defined by combining the dynamics of the amounts of resources and toxins, abundances of species, with the environmental switching (Eqs (1a) - (3)):

$$
\begin{aligned}
\frac{d}{d\sigma} P\left(\vec{r}, \vec{t}, \vec{s}, \xi, \sigma\right) = & \sum_{i=1}^{N/2} \left(\mathbb{E}_{r_i}^- - 1\right) \left\{\tau_{r_i}^+ P\left(\vec{r}, \vec{t}, \vec{s}, \xi, \sigma\right)\right\} \\
& + \sum_{i=1}^{N/2} \left(\mathbb{E}_{r_i}^+ - 1\right) \left\{\tau_{r_i}^- P\left(\vec{r}, \vec{t}, \vec{s}, \xi, \sigma\right)\right\} \\
& + \sum_{j=1}^{N/2} \left(\mathbb{E}_{t_j}^- - 1\right) \left\{\tau_{t_j}^+ P\left(\vec{r}, \vec{t}, \vec{s}, \xi, \sigma\right)\right\} \\
& + \sum_{j=1}^{N/2} \left(\mathbb{E}_{t_j}^+ - 1\right) \left\{\tau_{t_j}^- P\left(\vec{r}, \vec{t}, \vec{s}, \xi, \sigma\right)\right\} \\
& + \sum_{k=1}^{N} \left(\mathbb{E}_{s_k}^- - 1\right) \left\{\tau_{s_k}^+ P\left(\vec{r}, \vec{t}, \vec{s}, \xi, \sigma\right)\right\} \\
& + \sum_{k=1}^{N} \left(\mathbb{E}_{s_k}^+ - 1\right) \left\{\tau_{s_k}^- P\left(\vec{r}, \vec{t}, \vec{s}, \xi, \sigma\right)\right\} \\
& + \nu \left\{P\left(\vec{r}, \vec{t}, \vec{s}, -\xi, \sigma\right) - P\left(\vec{r}, \vec{t}, \vec{s}, \xi, \sigma\right)\right\}
\end{aligned}
\tag{4}
$$

where $P\left(\vec{r}, \vec{t}, \vec{s}, \xi, \sigma\right)$ gives the probability to find the population in state $\left(\vec{r}, \vec{t}, \vec{s}, \xi\right)$ at time σ, with $\vec{r} = (r_i)$, $\vec{t} = (t_j)$, $\vec{s} = (s_k)$. Here, $\mathbb{E}_{r_i}^\pm$ is a shift operator such that

$$
\mathbb{E}_{r_i}^\pm P\left(\vec{r}, \vec{t}, \vec{s}, \xi, \sigma\right) = P\left(r_1, \ldots, r_i \pm 1, \ldots, r_{N/2}, \vec{t}, \vec{s}, \xi, \sigma\right),
\tag{5}
$$

and $\mathbb{E}_{t_j}^\pm$ and $\mathbb{E}_{s_k}^\pm$ are the equivalent shift operators for t_j and s_k, respectively. Note that $P\left(\vec{r}, \vec{t}, \vec{s}, \xi, \sigma\right) = 0$ whenever any of $r_i, t_j, s_k < 0$. The first to sixth lines on the right-hand-side of Eq (4) represent the birth-and-

death processes (1), while the last line accounts for environmental switching.

The master equation (4) fully describes the model dynamics and can be simulated exactly with the Gillespie algorithm (Gillespie, 1977). Owing to the stochastic nature of the model, after a time that diverges exponentially with the community size, DN will cause the eventual collapse of the population (Spalding et al., 2017; Assaf and Meerson, 2017). This phenomenon is practically unobservable when resource levels remain sufficiently large, i.e. if $R_i^- \gg 1$, and the population settles in a long-lived quasi-stationary distribution. Here we focus on the quasi-stationary regime that is attained when distributions of species abundances appear to be stationary for a long time (see supplementary video): the distributions of two species' abundances change little from time $\sigma = 130$ onward when the parameter values are as shown in Table A.1 with $\nu = 10^{-1}$ and $\delta = 0.2$, which are typical parameter values used in this study. It would be reasonable then to set $\sigma_{end} = 200$ expecting that species' abundances reach a quasi-stationary state in many of our chosen parameter values.

2.1 Evaluating species interactions

First, we analyze how the net effect of species interactions (i.e., resource competition and facilitation via detoxification) can change under DN and EFs by measuring the extinction probabilities in the presence or absence of another species at time σ_{end} (see Appendix 1.1 for details). We begin with two species ($N = 2$) where the sign and magnitude of species interactions can change because the amounts of resources and toxins can change the net effect of resource competition and facilitation via detoxification (Piccardi et al., 2019). We used parameter values such that species 1, if it persists, always outcompetes species 2 in the absence of DN (i.e., species 1 grows faster than species 2, see Appendix 2.1 for analysis and Table A.1 for parameter values) to identify the effects of DN. Importantly, EFs alone do not change species 1's extinction probabilities in mono- versus co-culture in the absence of DN. Under DN coupled with EFs in our chosen parameter range, either of the two species or both species tend to go extinct. As a proxy for interactions, we focus on the net effect of species 2 on species 1, which is defined by the extinction probability of species 1 in mono-culture minus that in co-culture with species 2 (the so-called *difference in extinction probability*, see also Eq (A.1)): species 2 has a negative (positive) effect on species 1 if species 1 more (less) frequently goes extinct in co-culture with species 2 than in mono-culture. When species 2 has a negative effect on species 1, one can consider two possibilities in co-culture: (i) species 2 outcompetes species 1 (see Fig. 4) or (ii) both species 1 and 2 go extinct. As explained in section 3.3, we focus on the former probability, the so-called *probability of exclusion of the fittest* to understand the net effect of species 2 on species 1. We performed 10^5 simulations for each switching rate and toxin sensitivity in mono- and co-cultures. We calculated 95% of highest posterior density intervals (HPDI) to measure the uncertainty of species 1's extinction probabilities (see Appendix 1.1) but these intervals are too small to be visible on our plots due to the large number of simulations we ran.

2.2 Evaluating species diversity

To explore how species diversity changes with the switching rate, we ran simulations at different numbers of species ranging from $N = 2$ species to $N = 10$ species and different mean toxin sensitivities (see Appendix 1.3 for details). For each condition (one number of species and one mean toxin sensitivity), we sampled 100 sets of parameters from certain probability distributions. These sets of parameters represented 100 communities composed of N species. In this analysis (section 3.5) we consider a more general scenario and all previous simplifying assumptions on the parameter values are relaxed: all species have randomly drawn (see Appendix 1.3), and hence typically different, growth and death rates.

The dynamics of each community were independently simulated 100 times to see whether the species composition was robust against DN and EFs. These 100 replicate runs can be seen as 100 independent "patches" that initially consist of the same species set, but no species migrate from one patch to another. We measured the beta diversity of each community (Jost, 2007; Chao et al., 2012) and species richness (number of surviving species) as functions of the environmental switching rate in the quasi-stationary distribution (at time σ_{end}). Beta diversity accounts for the heterogeneity of each community across 100 replicates For example, if beta diversity is larger than one but species richness is one in all replicates, different species fixate in each replicate. In contrast, beta diversity is one if all communities show identical species compositions; for example, in the two-species scenario with parameter values shown by Table A.1 and in the absence of DN, species 1 always outcompetes species 2 and thus beta diversity is one. This baseline corresponds to a perfectly deterministic scenario. One could instead consider using another baseline corresponding to a perfectly stochastic or neutral scenario where all species' parameter values are identical (e.g., Fig. A.21), in which case species compositions are determined only by demographic noise coupled with environmental fluctuations. We choose to focus on the deterministic baseline, as beta diversity is always $= 1$ regardless of the number of species, their parameter values, and the environmental switching rate. In contrast, beta diversity in the neutral scenario changes with these parameter values, making it more difficult to compare across conditions.

We compared the patterns of beta diversity of two- or ten-species communities with the probability of exclusion of the fittest in species pairs sampled from these communities (see also Appendix 7). The sampled species pairs may stably coexist, in which case the fittest species in a pair is the one that is more abundant in the absence of any noise. If both species go extinct in the absence of noise, either of the two species is randomly labeled as the fittest. This labeling generalizes what is used in the species interaction analysis above where species 1 (the fittest species) always grows faster than species 2, such that their long term coexistence is impossible in the absence of DN.

2.3 Statistical analysis

Statistical analysis was performed with Python 3.7.6 incorporating Scipy 1.4.1. and pymc3 3.10.0. For statistical tests of Pearson's correlation and Spearman's rank-order correlation, scipy.stats.pearsonr and scipy.stats.spearmanr were used, respectively. For calculation of HPDIs, pymc3.stats.hpd was used.

3 Results

3.1 Toxin sensitivity determines single species' response to coupled DN and EFs
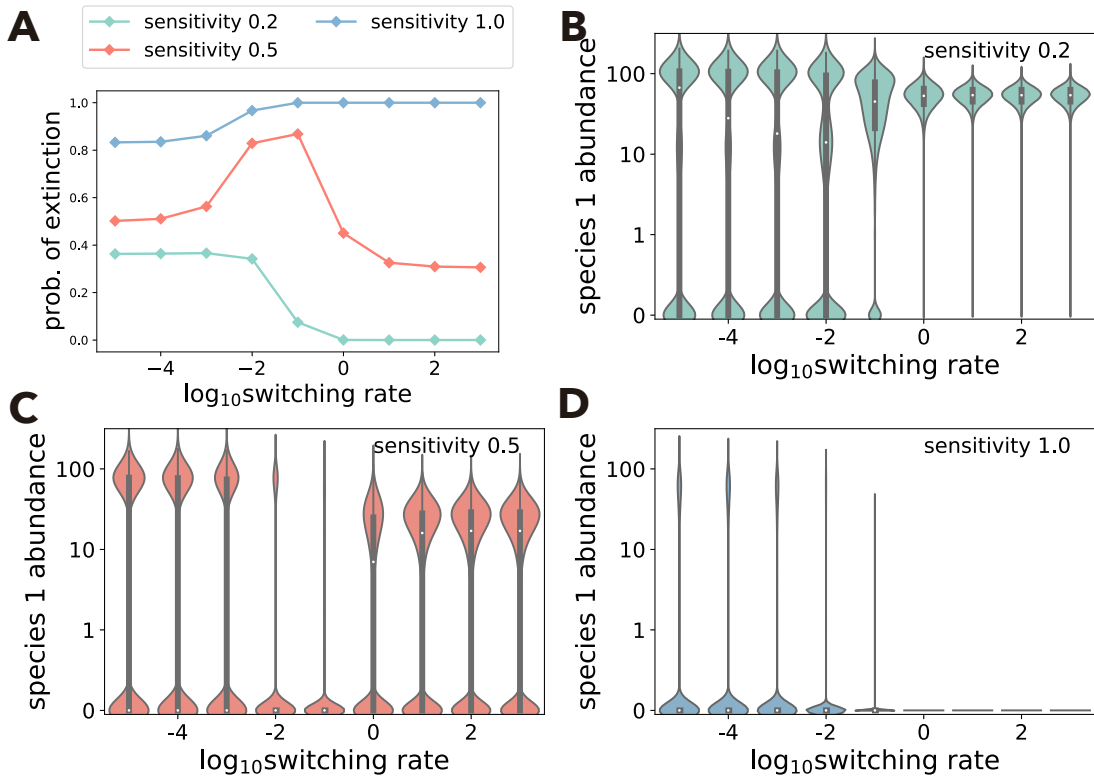
Figure 2: Extinction probability and species abundance in mono-culture

A: Extinction probability of species 1 in mono-culture when the toxin sensitivity is low (green), moderate (orange), or high (blue). 95% HDPIs are too small to see. B–D: violin plots of species 1's abundance at the end of 10^5 simulations with a low (B), moderate (C), or high (D) toxin sensitivity. White dots and black bars represent median values of the abundances and their interquartile ranges, respectively.

To establish our intuition on how the coupling of EFs and DN affects the dynamics, we first analyse the extinction probabilities of a single species (species 1) in mono-culture with one type of resource and toxin (Fig. 2A). As the switching rate decreases, the duration of the harsh or mild environments become longer. In the presence of DN, this duration determines whether or not the species goes extinct, together with its sensitivity to toxins in the environment, which can be seen to modulate environmental harshness. When the switching rate is very low ($\nu \to 0$), the species is exposed to the static environment with either abundant or scarce resources (with probability 0.5, respectively) depending on the initial environmental condition $\xi(0)$: it mostly goes extinct under scarce resource supplies even when their sensitivity to toxins is low (Fig. 2B). On the other hand, abundant resource supplies maintain species 1 with some probability even if it is very sensitive to the toxin (Fig. 2D). Over many simulations, low fluctuation rates therefore result in a bimodal distribution of the species' abundance (e.g., Fig. 2B). At the other extreme, very high switching rates ($\nu \to \infty$) expose the bacteria to an environment with mean abundance of resources (Wienand et al., 2017, 2018; West and Mobilia, 2020; Taitelbaum et al., 2020). This is enough to rescue the species with a low toxin sensitivity (Fig. 2B) but not

with a high sensitivity (Fig. 2D). At an intermediate toxin sensitivity (Fig. 2C), the worst situation lies in the intermediate fluctuation rate: the duration of the harsh environment is long enough to drive them extinct, but the time with abundant resource supply is not long enough to rescue them fully. In sum, even when only a single species is present, we see non-trivial patterns in its response to EFs coupled with DN, which depends on its sensitivity to toxins.

3.2 Toxin sensitivity changes how switching rate affects two-species competition

Next, we add another species (species 2) that grows slower than species 1 into the environment and ask how it interacts with species 1 in our model. Rather than measuring interactions through the effect of each species on the other's abundances, we focus on species 1 and analyze how its extinction probability is affected by the presence of species 2, compared to mono-culture (Fig. 2). Our reasoning is that (i) species 1 should always out-compete species 2 in the absence of DN, and that (ii) we already know the extinction probability of species 1 under EFs and DN in mono-culture; measuring any deviation from the mono-culture outcome allows us to quantify how likely it is for the fitter species to be lost in a given community. Such species loss events can be seen as ecological drift. We again explore changes in the species' toxin sensitivity, as we learned above that it affects species abundances via DN (Fig. 2), but also because we expect it to affect species interactions (Piccardi et al., 2019). For now, we varied sensitivity to toxins simultaneously for both species, an assumption that we relax later.

When both species were highly sensitive to the toxin, species 2 had a positive effect on species 1, reducing its extinction probability. This occurs because in the simulations, toxic compounds are degraded more quickly in co-culture than in mono-culture due to the larger initial number of individuals ($s_1(0) + s_2(0) > s_1(0)$). This effect can be recapitulated by a mono-culture with larger initial abundance (Fig. A.20). A larger total initial species abundance in co-culture decreases death rates, which outweighs competition for nutrients in toxic environments (Piccardi et al., 2019). However, for most parameter values in Fig. 3A, species 2 has a negative effect on species 1 by increasing its extinction probability. We therefore focus on competitive interactions in the main text and consider positive interactions in Appendix 4.

As in the single species case (Fig. 2), the extinction of species 1 was highly dependent on the toxin sensitivity of the two species, as we varied the fluctuation rate: monotonically increasing, monotonically decreasing, or non-monotonically changing with a minimum or maximum value at an intermediate switching rate (Fig. 3B). Interestingly, this pattern does not match the single-species behavior (compare Fig. 2A and 3A).

3.3 Behavior at extreme switching rates explains non-monotonic changes in exclusion of the fittest

To better understand why the faster-grower goes extinct in the observed parameter ranges (Fig. 3B), we decompose species 2's effect on species 1 into the probability that species 2 persists but species 1 goes extinct (hereafter, called *probability of exclusion of the fittest*) and the probability of any other outcome (e.g., extinction
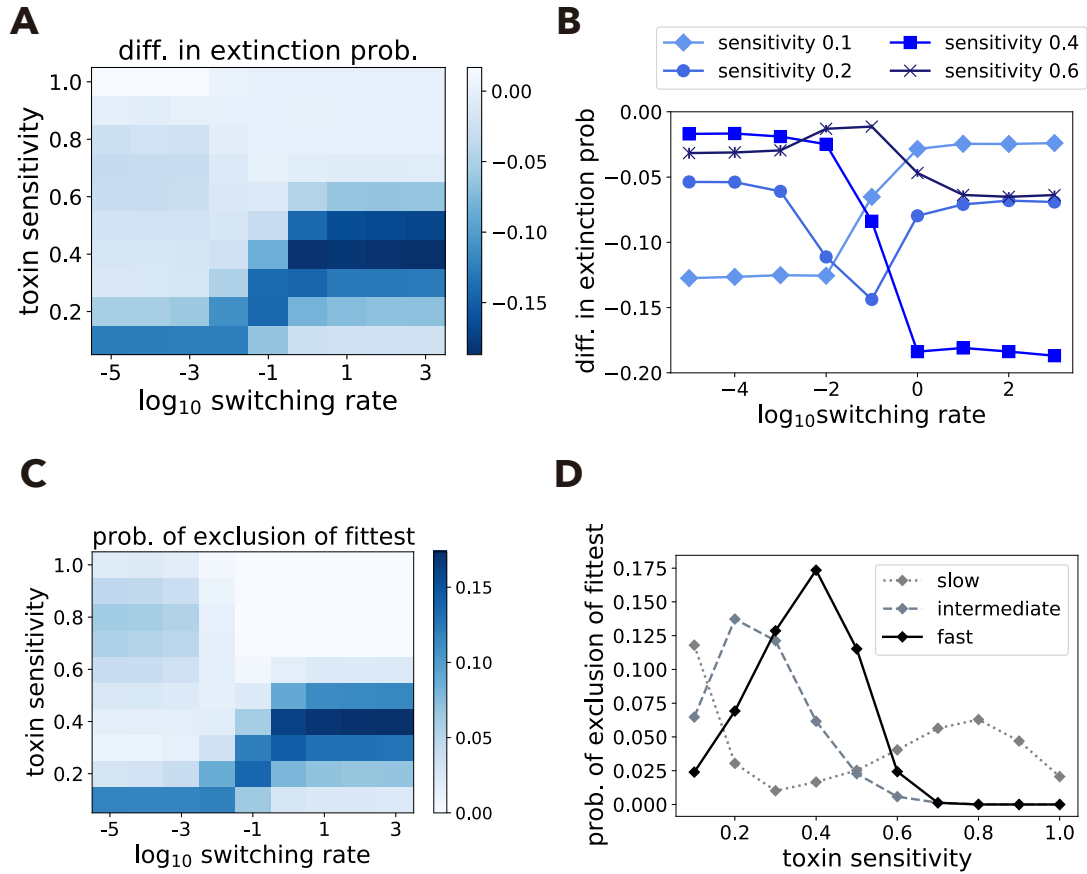
Figure 3: Species interaction strength changes differently over the switching rate

A: Difference in species 1's extinction probability in mono-culture minus co-culture with species 2 (Eq A.1) changes over the switching rate ν and the two species' identical toxin sensitivities. In most of the parameter space, species 2 has a negative effect on species 1 (i.e., species 2 increases the extinction probability of species 1). B: Some illustrative examples from panel A plotted differently to show how species 2's effect on species 1 changes over the switching rate at given toxin sensitivities. The difference in extinction probability can monotonically increase (toxin sensitivity 0.1), monotonically decrease (toxin sensitivity 0.4), or non-monotonically change with a minimum (toxin sensitivity 0.2) or a maximum (toxin sensitivity 0.6) value at an intermediate switching rate. C: Probability that species 2 persists but species 1 goes extinct (i.e., exclusion of the fittest) over the switching rate and the toxin sensitivity. D: probabilities of exclusion of the fittest over the toxin sensitivity, when the environmental switching rate is slow ($\nu = 10^{-5}$), intermediate ($\nu = 10^{-1}$), or fast ($\nu = 10^3$). In panels B and D, 95% HDPIs are too small to see.

of both species, see Eq (A.2)). We focus on the probability of exclusion of the fittest, as it explains the variation in species interaction strength in most cases (compare Figs. 3A and C) and investigate how it changes over the switching rate and depends on toxin sensitivity.

We again let the two extreme switching rates guide our intuition (see also Fig. A.4). At a very slow rate ($\nu \to 0$), the resource supply remains either scarce or abundant, while a very fast environmental switching rate ($\nu \to \infty$) drives resource supply to the mean concentration. We explore system behavior at these three constant resource supplies (scarce, abundant or mean) and different toxin sensitivities, which together represent how harsh the environment is. When the environment is harsh – due to resource scarcity, high toxin supplies, or high toxin sensitivity –, both species are most likely to go extinct rather than to outcompete each other (Figs. 5A – C). As toxin sensitivity goes down and species survival becomes more likely, DN becomes more important and we see a higher probability that even the fitter species (species 1) will be excluded. The more resources are available, the more likely it is that species survive – in particular, that species 1 out-competes species 2, and the peak of competitive exclusion moves to a higher toxin sensitivity (see arrows in Figs. 5A – C). When it is easy for both species to survive (toxin sensitivity is low and/or resources are abundant), DN no longer plays an important role and the faster-growing species 1 is unlikely to be excluded. ntuitively then, the exclusion of the fittest is caused by the coupling of DN with EFs: Harsh environments, where both species' abundances are positive but low (see Appendix 2.2), lead to stronger DN and a higher probability of extinction of the fittest species compared to a static or mild environment (Fig. 4). It is important to stress that the exclusion of the fittest never occurs without DN, regardless of EFs (see Appendix 2).

Hereafter, we refer to the toxin sensitivity that maximizes the probability of exclusion of the fittest in the absence of environmental switching as the "critical toxin sensitivity" (arrows in Figs. 5A – C). We see two critical toxin sensitivities (at 0.1 and 0.8 in Figs. 3C and D, Figs. 5A and C) at $\nu \to 0$ that correspond to the long time spent with either scarce or abundant resources. Instead, at $\nu \to \infty$, where resources remain at mean abundance, there is a single critical toxin sensitivity (at 0.4 in Figs. 3C and D) where exclusion of the fittest is most likely (Fig. 5B). Toxin sensitivities between these critical values can show a maximum or minimum probability of exclusion of the fittest at an intermediate switching rate, resulting in the rugged landscape of Fig. 3C (see Appendix 1.2 for more detail).

3.4 Competition strength changes non-monotonically under different scenarios

We have shown that using a given set of parameters, the rugged landscape shown in Fig. 3C causes the competitive exclusion of a faster-growing species to either increase, decrease or vary non-monotonically across switching rates, depending on toxin sensitivity. We next explore the generality of this finding. In the appendix, we explore scenarios where (i) switching occurs in toxin rather than resource supplies, where (ii) both resource and toxin supplies switch (Table 1, see Appendix 3), or where (iii) we change the amounts of scarce and abundant resource supplies (Appendix 4).

In all these scenarios, the landscapes of species 1's difference in extinction probability and probability of
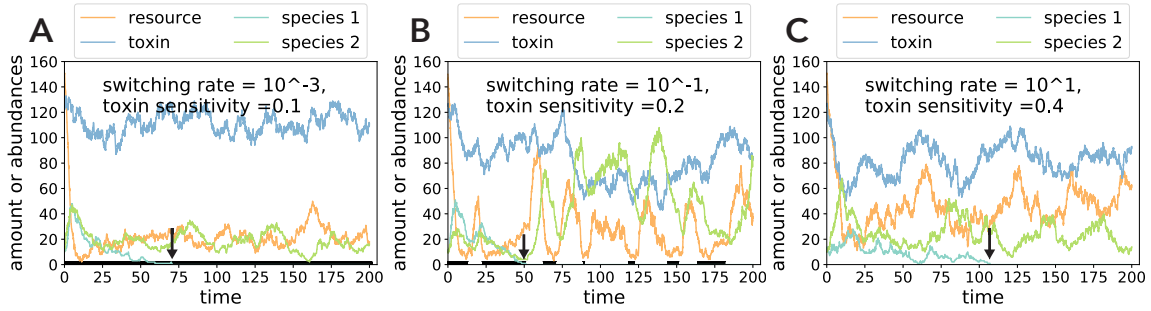
Figure 4: Examples of the dynamics with exclusion of the fittest

In these examples (A: $\nu = 10^{-3}$, and $\delta = 0.1$, A: $\nu = 10^{-1}$, and $\delta = 0.2$, C: $\nu = 10^1$, and $\delta = 0.4$), species 1 goes extinct but species 2 survives at the end of simulation $\sigma_{end} = 200$ due to the DN-EFs coupling. Black lines on the x-axis represents times when the resource supply is scarce ($\xi(t) = -1$) while white lines represent times when the resource supply is abundant ($\xi(t) = 1$). In panels A and B, species 1 decreases its abundance and goes extinct during the scarce resource supply condition (pointed by arrows). In panel C, the environmental conditions are not shown because they are not visible due to the fast switching.
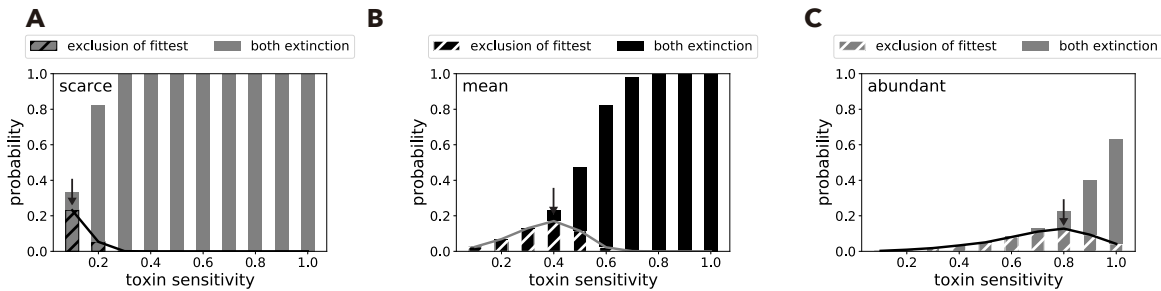
Figure 5: Exclusion of the fittest explains how DN changes with toxin sensitivity and resource supply

Analysis of exclusion of the fittest predicts at which toxin sensitivity DN is strongest. A − C: In the absence of environmental switching, the probability of exclusion of the fittest (i.e., probability that species 2 excludes species 1 and survives, shown by solid lines and hatched bars) is uni-modal over the toxin sensitivity while the probability of both species extinction (bars) monotonically increases. The toxin sensitivities giving the peak values of probabilities of exclusion of the fittest (critical toxin sensitivities, pointed by black arrows) depend on the resource supply: scarce $R_1 = R_1^-$ (A), mean $R_1 = \langle R_1 \rangle$ (B), or abundant $R_1 = R_1^+$ (C). 95% HDPIs are not observable as they are very small.

15

exclusion of the fittest are qualitatively similar (Figs.3 C and D, A.7, and A.9). However, each scenario differs in the three critical toxin sensitivities and likelihood that the difference in extinction probability non-monotonically changes over the switching rate. Accordingly, we asked whether the distances between critical toxin sensitivities might predict the probability of observing non-monotonic behavior. In Table A.2 and Fig. A.14, we show that the distance between critical sensitivities under harsh and mean environments (i.e., very fast environmental switching) correlates positively with the likelihood of observing non-monotonic effects of the switching rate on competition (Fig. A.14, black circles; Spearman's $\rho = 0.77$, P-value: 0.043), but no significant correlation was found with the distance between the critical toxin sensitivities under the mean and mild, or the harsh and mild environments (Fig. A.14, grey diamonds and cross marks; Spearman's $\rho = -0.22$, P-value: 0.64, and $\rho = 0.42$, P-value: 0.35, respectively). Therefore, the non-monotonic change of the difference in extinction probability is likely when there is a large difference in the critical toxin sensitivities under harsh and mean conditions.

3.5 Beta diversity changes similarly to exclusion of the fittest

In the previous sections, we focused on interactions between two species and the conditions under which one may drive the other extinct. Ultimately, however, our interest is to predict how whole communities comprised of tens, hundreds or even thousands of species are affected by fluctuations in the environment.

We set up a model of 100 communities composed of between 2 and 10 species each. Species within each community were defined by parameter values that were randomly sampled from the same distributions with the exception of toxin sensitivity δ, which was sampled from beta distributions with different means, ranging from $\bar{\delta} = 0.1$ to 1. We generated a new set of 100 communities with different numbers of species and different fluctuation rates as above, and ran 100 replicate simulations for each of the 100 communities in each set. We then measured beta diversity across the 100 replicate simulations per community and final species richness (number of surviving species) over all 100 runs for the 100 communities (total: 10'000). In this model design, the 100 replicate runs represent independent "patches" without migration. Their beta diversity then indicates how different the species compositions were across all patches in a given environment (e.g. a given fluctuation rate), and we repeat the exercise 100 times with different species sampled from the same distributions to see generality of the results. A high beta diversity would then indicate that we have different community compositions in each patch, while a beta diversity of 1 would tell us that all patches have the same species composition.

In two-species communities, we obtained qualitatively similar patterns of exclusion of the fittest to those in the species interaction analysis (compare column A in Fig. 6 with Fig. 3C), suggesting that our results in Fig. 3C are unlikely to be specific to the choice of species parameter values in Table A.1. Beta diversity changes over the environmental switching rate similarly to the probability of exclusion of the fittest for four out of the five tested mean toxin sensitivities (columns A and B in Fig. 6, see also Fig. A.15): both monotonically decrease (mean toxin sensitivity $\bar{\delta} = 0.1$ or 1.0), or non-monotonically change with maximum ($\bar{\delta} = 0.2$) or minimum ($\bar{\delta} = 0.6$) values at intermediate switching rates. This similarity can be explained as follows: ignoring extinction of both species, a small probability of exclusion of the fittest indicates that the stronger species 1 fixates in most

simulations, a homogeneous outcome with small beta diversity. In contrast, when the exclusion of the fittest is more likely, the weaker species 2 is more likely to fixate, leading to significant heterogeneity in the simulation results and large beta diversity (species 2 survives alone in a fraction of the runs, and species 1 in the rest). At one mean toxin sensitivity $\bar{\delta} = 0.4$, the patterns of the probability of exclusion of the fittest and beta diversity over the switching rate do not match. At this toxin sensitivity, beta diversity remains high at low switching rates ($\nu = 10^{-5}, 10^{-4}, 10^{-3}$) because both species go extinct in 50% of the runs but they can coexist in about 7%, as illustrated by our measure of final species richness (column C in Fig. 6). Beta diversity ignores the cases of both-species extinction but increases in cases of coexistence, see Eq (A.6). Overall, looking at final species richness (column C in Fig. 6), we see that complete extinction is more likely to occur as sensitivity increases. At the highest toxin sensitivity, the only way a species can survive is if the switching rate is really low and they can benefit from abundant resources for a long time, while at lower toxin sensitivity, complete extinctions only occur at low switching rates, because there is long term exposure to scarce resources.

In communities with ten species (see Figs. A.18 and A.19 for intermediate community sizes), we observe similar patterns between beta diversity (column D in Fig. 6) and the probability of exclusion of the fittest (column A in Fig. 6) Studying interactions between species pairs can therefore predict the behavior of a ten-species community. To explore whether it matters which two species one selects for the interaction analysis, we next repeatedly sub-sampled species pairs from each ten-species community and compared the exclusion of the fittest in the pairs with the beta diversity of the whole community (Fig. A.16). Naturally, the more species pairs one samples, the more accurately we can predict the pattern of beta diversity, but this accuracy appears to saturate at around five species pairs (Appendix 7, Fig. A.17), which is approximately 11% of all possible species pairs. In addition, large beta diversity does not necessarily reflect a large variation in species richness; large beta diversity with small species richness (see mean toxin sensitivity $\bar{\delta} = 0.4$ or 0.6 at switching rate $\nu \geq 10^0$ in columns D and E of Fig. 6) indicate that different species fixate in each run, which also supports the observed relationship between the exclusion of the fittest and beta diversity. In sum, estimating the probability of exclusion of the fittest between a few randomly selected species pairs (section 3.3) is a good predictor for the beta diversity of larger communities under those same environmental conditions (see Appendix 7 for detailed discussion). This similarity is not coincidental: as for the probability of exclusion of the fittest, beta diversity is also maximized when DN is the strongest, such that different species survive in each patch.

4   Discussion

Understanding how species diversity in microbial communities arises and is maintained is a central question in microbial ecology and evolution. While many theoretical and experimental studies have addressed this question in static environments, community diversity is expected to respond to fluctuations between benign and harsh environmental conditions, which can alter the abundance of different species and the interactions between them (see e.g., Rodríguez-Verdugo et al. (2019)). Strong drops in population sizes caused by harsh conditions can increase the strength of DN, which, coupled with EFs may lead to non-trivial outcomes (Wienand
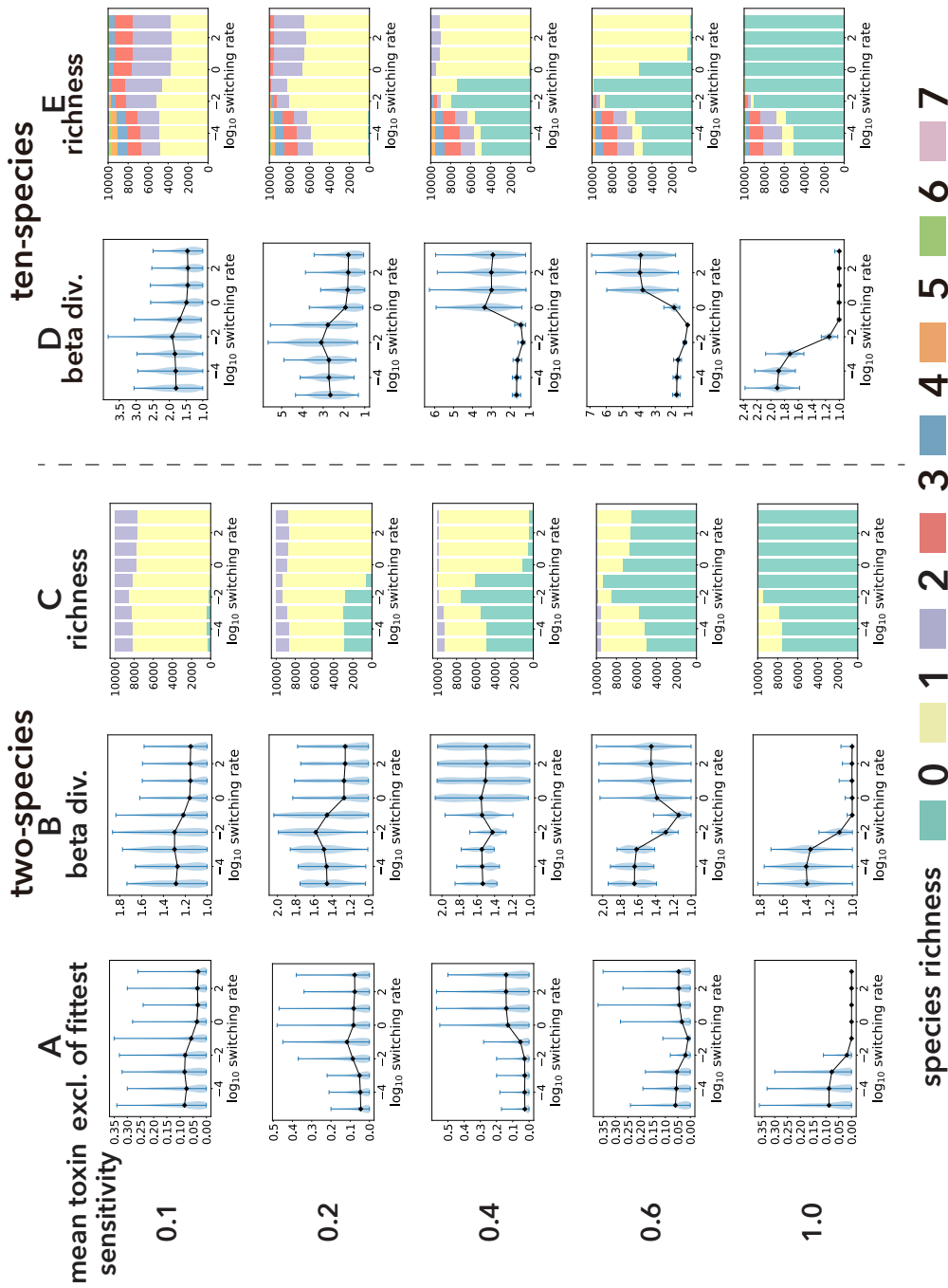
Figure 6: Comparison of exclusion of the fittest and diversity

Probabilities of exclusion of the fittest (column A), beta diversities (column B), and species richness (column C) over the switching rate and the mean toxin sensitivities in two-species communities. Here, competitive exclusion refers the event where the slower-growing species excludes the faster-growing species. Columns D and E show the beta diversity and species richness in ten-species communities, respectively. In the plots of probability of exclusion of the fittest and beta diversity, the black lines show the means and blue areas represent the probability distributions calculated from 10'000 simulations (100 beta diversity and each of them from 100 replicate runs). Each color in the species richness plots represents the proportion of 10'000 runs where at the end of the run there were that number of species surviving. Each row represents the mean toxin sensitivity of communities in those runs.

et al., 2017, 2018; West and Mobilia, 2020; Taitelbaum et al., 2020). Here we have analysed a mathematical model representing a biological scenario such as a gut or soil microbial community that experiences fluctuations between benign and harsh conditions, such as feast and famine, or drought and rain. The model shows how community diversity – mediated by inter-specific interactions and DN – changes with environmental fluctuation rates.

Our study is centred on two main findings. First, we show that the rate at which resource supplies switch changes the ability of a slower-growing species to drive a fast-growing one extinct (Fig. 3). While the fitter species will never be excluded in the absence of DN, in a fluctuating environment with DN, harsh conditions (e.g. scarce resources) strengthen DN to sometimes drive species extinction in spite of their greater fitness. We see this as a form of ecological drift, wherein selection by the environment is not strong enough to maintain the fittest species and thus the fittest species may go extinct due to DN. By changing the length of time spent in the harsh environment, the environmental switching rate affects such competitive exclusion. In addition, the species' ability to withstand environmental harshness can also strengthen DN and lead to stronger ecological drift. We confirmed the generality of these results by exploring various forms of EFs (e.g., asymmetric switching environments and cyclically changing environments in Appendix 5). When we consider many replicate communities with identical initial species compositions, the increased stochasticity resulting from DN means that which species go extinct under these harsh conditions is less dependent on their relative competitiveness, and species composition will be more random in each replicate, leading to greater beta diversity. Recent studies corroborate this finding: smaller communities show larger varieties in species composition due to DN (Gilbert and Levine, 2017), while species composition is robust against EFs when species are insensitive to them (Dedrick et al., 2021).

Although we now understand that beta diversity increases when the probability of exclusion of the fittest is high, precisely when such ecological drift is maximized will be difficult to predict in practice, as it is a function of multiple factors: the form of the EFs (Appendix 3 and Appendix 5), the magnitude of EFs (Appendix 4), the rate of EFs and the sensitivity of species to environmental harshness. Yet, our simulation approach allows us to investigate these different effects efficiently.

Our second main finding is that a good way to predict how beta diversity will respond to EFs is to measure the probability of exclusion of the fittest across fluctuation rates in few pairs of species randomly sampled from a focal community and use it as an indicator for how the beta diversity of the whole community will behave (Fig. 6, Appendix 7). In two species communities, explaining the similarity between the probability of exclusion of the fittest and beta diversity is straightforward: the latter accounts for the probabilities that both or either of the two species persist. We verified that this similarity also holds in neutral scenarios (i.e., two species are identical except for their labels and thus the fittest species is arbitrarily chosen, Fig. A.21). In larger communities, the similarity is not straightforward but our first main finding helps to understand this result. EFs affecting DN strength could exclude species that would persist in a community without noise. If the "fittest" species goes extinct with some probability due to DN, community compositions at time σ_{end} are heterogeneous and beta diversity increases. Therefore, our two findings emphasize the importance of coupling EFs and DN: competition

results can differ from those in static environments and this affects beta diversity.

This brings us to a hypothesis that has been debated at length in ecology: the intermediate disturbance hypothesis (Connell, 1978; Grime, 1973), which states that intermediate intensity and frequency of disturbance maximize species diversity. Fox (2013) argues that the intermediate disturbance hypothesis should be abandoned because many examples disagree with it (Mackey and Currie, 2001; Miller et al., 2011). In our model, fluctuations in resource and toxin can be regarded as disturbances. In agreement with Mackey and Currie (2001) and Miller et al. (2011) then, an intermediate intensity (i.e., toxin sensitivity) or disturbance frequency (environmental switching rate) does not always maximize beta diversity: our analysis shows that intermediate frequencies of disturbance maximize beta diversity only when mean toxin sensitivity is within a certain range. Mean toxin sensitivities at the two thresholds of this range show that beta diversity monotonically decreases or increases over the switching rate. These thresholds depend on scenarios of environmental switching and amounts of resource supplies because these parameters change the probability of competitive exclusion (see Appendix 3 and Appendix 4). High beta diversity at intermediate disturbances is then a consequence of a change in environmental conditions and not expected to apply generally.

The relationship between EFs and species diversity is also an important question in the modern coexistence theory, which predicts that fluctuations will affect species coexistence by changing species growth rates when rare (Chesson, 2000a,b; Barabás et al., 2018; Ellner et al., 2019; Letten et al., 2018a). Compared to the approach taken in this study – where we ask how many and which species persist at the end of a long but fixed time frame (i.e., for a quasi-stationary distribution, Fig. 6) – the modern coexistence theory allows one to analyze whether or for how long a set of species will all coexist (Schreiber et al., 2020). An interesting future direction would be to apply the modern coexistence theory to investigate how environmental fluctuation rates and toxin sensitivities affect the duration of all-species coexistence. This approach would help to propose biological mechanisms behind species coexistence in our setup.

Of course, our model makes some simplifying assumptions and has some limitations. First, we used arbitrary time units, which in practice can be considered to be hours, corresponding to typical bacterial growth rates in relevant experiments (Novick and Szilard, 1950; Lin et al., 2002; Zhao and Lin, 2003). This implies that species interactions and beta diversity will vary when environmental switching ranges from hourly ($\nu = 10^0$) to about once every four days ($\nu = 10^{-2}$) on average, which is shorter than in some experimental studies (Benneir and Lenski, 1999; Rodríguez-Verdugo et al., 2019; Chen and Zhang, 2020) but not impractical. That said, under this assumption, changing environmental switching from a daily to an hourly scale, for example, would show different species compositions or diversity.

Second, our model focuses on competitive exclusion but other types of interactions can also affect diversity (Rodríguez-Verdugo et al., 2019). Positive interactions between pairs of species (e.g., cross-feeding), for example, might increase alpha and gamma diversities, because such interactions enable species to coexist (Sun et al., 2019). This could result in an increase in beta diversity because the extinction of one species increases its partner species' extinction probability (Dunn et al., 2009; Goldberg and Friedman, 2020).

Third, one can consider more complex species-resource (and species-toxin) interaction functions. For ex-

ample, species' growth rates can be limited by the resource with the smallest amount when the resources are complementary (León and Tumpson, 1975). In addition, absorption rates of toxin might correlate with species' growth rate if the toxin targets cell metabolism. These functional forms may enable more species to coexist, collapsing the similarity between beta diversity and the probability of exclusion of the fittest. Our study would also be more general if species could change their growth rates depending on resource concentrations (Nguyen et al., 2020a) or types (Balakrishnan et al., 2021). Introducing such plasticity would affect species' extinction probability and diversity. One could also consider building species-compounds interaction networks. In the current manuscript, we assume that each resource (toxin) has a positive (negative) effect on each species. However, we know that some compounds, (e.g., pesticides) can be resources for some species but toxic to others (Muturi et al., 2017), and other compounds (e.g., that affect pH (Ratzke et al., 2020) and osmolarity (Larsen, 1986; Oren, 2008)) can have either positive or negative effects on growth depending on their concentrations. The way in which species and compounds interact could affect exclusion of the fittest and species diversity.

Although we focused on beta diversity to measure the heterogeneity of communities in this manuscript, other metrics of species diversity could be considered. For example, Fung et al. (2015) and Kalyuzhny et al. (2015) analyze how demographic noise and environmental fluctuations affect species abundance distributions (SAD). Grilli (2020) instead calculate the mean abundance distribution (MAD), which is defined as the distribution of mean species abundances over communities and follows a log-normal distribution. While SAD characterizes species diversity within a single community, it does not explain the heterogeneity of species compositions across communities, which is what we can capture with beta diversity. Combining multiple SADs or probability distributions of SADs is possible, but would not be as easy to analyze. Similarly, MAD ignores the variation of species' abundances across communities and may therefore not capture the heterogeneity of communities caused by demographic noise and environmental fluctuations.

Finally, our community analysis considers up to ten microbial species, which is orders of magnitude below the size of natural microbial communities, according to genomic sampling (Gans, 2005; Roesch et al., 2007). However, it may also be reasonable to assume that species live in structured environments where they cannot possibly interact with more than a handful of other genotypes (Tecon et al., 2019). This suggests that a 10-species community may already be biologically meaningful.

In conclusion, the time scale of environmental fluctuations changes the importance of species fitness for survival and thus community beta diversity. This occurs in our model when EFs affect the strength of DN, leading to the occasional exclusion of strong species. Predicting how the strength of DN changes is not simple because it is affected by both environmental and species' parameters (resource and/or toxin supplies and toxin sensitivities in our model). This may be one explanation as to why the intermediate disturbance hypothesis does not always hold, but rather there are many relations between diversity and disturbance (Mackey and Currie, 2001; Miller et al., 2011). Nevertheless, we found similarities between how competitive exclusion plays out between species pairs and beta diversity at the community level. In the event that we would like to predict how the diversity of a given ecosystem, such as a soil community or a bioremediation ecosystem, responds to environmental fluctuations, it may be sufficient to isolate a few culturable species and analyze their interactions

over different fluctuation rates. This approach promises to greatly facilitate our ability to study large and complex natural communities and their response to harsh conditions.

Author Contributions

SS, MM and SM designed the study, SS performed simulations, analyzed data, and wrote the first draft, and all authors contributed to revisions.

Acknowledgement

Data accessibility

The programming codes and simulation data in csv files for this manuscript are available in Github.

Funding Statement

References

Allen, L. J. S. An Introduction to Stochastic Processes with Applications to Biology. Chapman and Hall/CRC, New York, 2nd edition, 2010. ISBN 9780429184604. doi: 10.1201/b12537. URL https://www.taylorfrancis.com/books/9781439894682.

Amarasekare, P. The evolution of coexistence theory. Theoretical Population Biology, 133:49–51, 2019. ISSN 10960325. doi: 10.1016/j.tpb.2019.09.005. URL https://doi.org/10.1016/j.tpb.2019.09.005.

Assaf, M. and Meerson, B. WKB theory of large deviations in stochastic populations. Journal of Physics A: Mathematical and Theoretical, 50(26):263001, 6 2017. ISSN 1751-8113. doi: 10.1088/1751-8121/aa669a. URL http://dx.doi.org/10.1088/1751-8121/aa669a.

Balakrishnan, R., Hwa, T., and Cremer, J. Suboptimal proteome allocation during changing environments constrains bacterial response and growth recovery. bioRxiv, 2021. doi: https://doi.org/10.1101/2021.04.28.441780.

Barabás, G., D'Andrea, R., and Stump, S. M. Chesson's coexistence theory. Ecological Monographs, 0(0):1–27, 2018. ISSN 00129615. doi: 10.1002/ecm.1302. URL http://doi.wiley.com/10.1002/ecm.1302.

Bena, I. Dichotomous Markov noise: Exact results for out-of-equilibrium systems. A review. International Journal of Modern Physics B, 20(20):2825–2888, 6 2006. ISSN 02179792. doi: 10.1142/S0217979206034881. URL http://arxiv.org/abs/cond-mat/0606116.

Benneir, A. F. and Lenski, R. E. Experimental evolution and its role in evolutionary physiology. American Zoologist, 39(2):346–362, 1999. ISSN 00031569. doi: 10.1093/icb/39.2.346.

Butler, S. and O'Dwyer, J. Cooperation and Stability for Complex Systems in Resource-Limited Environments. Theoretical Ecology, 13:239–250, 2020. doi: 10.1101/514018. URL https://doi.org/10.1007/s12080-019-00447-5.

Butler, S. and O'Dwyer, J. P. Stability criteria for complex microbial communities. Nature Communications, 9(1):2970, 12 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05308-z. URL http://www.nature.com/articles/s41467-018-05308-z.

Cangelosi, G. A. and Meschke, J. S. Dead or Alive: Molecular Assessment of Microbial Viability. Applied and Environmental Microbiology, 80(19):5884–5891, 10 2014. ISSN 0099-2240. doi: 10.1128/AEM.01763-14. URL https://journals.asm.org/doi/10.1128/AEM.01763-14.

Chao, A., Chiu, C. H., Hsieh, T. C., and Inouye, B. D. Proposing a resolution to debates on diversity partitioning. Ecology, 93(9):2037–2051, 9 2012. ISSN 00129658. doi: 10.1890/11-1817.1. URL http://doi.wiley.com/10.1890/11-1817.1.

Chen, P. and Zhang, J. Antagonistic pleiotropy conceals molecular adaptations in changing environments. Nature Ecology & Evolution, 2 2020. ISSN 2397-334X. doi: 10.1038/s41559-020-1107-8. URL http://www.nature.com/articles/s41559-020-1107-8.

Chesson, P. Multispecies Competition in Variable Environments. Theoretical Population Biology, 45:227–276, 1994.

Chesson, P. Mechanisms of Maintenance of Species Diversity. Annual Review of Ecology and Systematics, 31(1):343–366, 11 2000a. ISSN 0066-4162. doi: 10.1146/annurev.ecolsys.31.1.343. URL http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.31.1.343.

Chesson, P. General theory of competitive coexistence in spatially-varying environments. Theoretical Population Biology, 58(3):211–237, 2000b. ISSN 00405809. doi: 10.1006/tpbi.2000.1486.

Chisholm, R. A., Condit, R., Rahman, K. A., Baker, P. J., Bunyavejchewin, S., Chen, Y.-Y., Chuyong, G., Dattaraja, H. S., Davies, S., Ewango, C. E. N., Gunatilleke, C. V. S., Nimal Gunatilleke, I. A. U., Hubbell, S., Kenfack, D., Kiratiprayoon, S., Lin, Y., Makana, J.-R., Pongpattananurak, N., Pulla, S., Punchi-Manage, R., Sukumar, R., Su, S.-H., Sun, I.-F., Suresh, H. S., Tan, S., Thomas, D., and Yap, S. Temporal variability of forest communities: empirical estimates of population change in 4000 tree species. Ecology Letters, 17(7): 855–865, 7 2014. ISSN 1461023X. doi: 10.1111/ele.12296. URL http://doi.wiley.com/10.1111/ele.12296.

Cignarella, F., Cantoni, C., Ghezzi, L., Salter, A., Dorsett, Y., Chen, L., Phillips, D., Weinstock, G. M., Fontana, L., Cross, A. H., Zhou, Y., and Piccio, L. Intermittent Fasting Confers Protection in CNS Autoimmunity by Altering the Gut Microbiota. Cell Metabolism, 27(6):1222–1235, 2018. ISSN 19327420. doi: 10.1016/j.cmet.2018.05.006. URL https://doi.org/10.1016/j.cmet.2018.05.006.

Connell, J. H. Diversity in Tropical Rain Forests and Coral Reefs. Science, 199(4335):1302–1309, 1978.

Coyte, K. Z., Schluter, J., and Foster, K. R. The ecology of the microbiome: Networks, competition, and stability. Science, 350(6261):663–666, 2015. ISSN 0036-8075. doi: 10.1126/science.aad2602. URL http://www.sciencemag.org/cgi/doi/10.1126/science.aad2602.

Davenport, E. R., Mizrahi-Man, O., Michelini, K., Barreiro, L. B., Ober, C., and Gilad, Y. Seasonal variation in human gut microbiome composition. PLoS ONE, 9(3), 2014. ISSN 19326203. doi: 10.1371/journal.pone.0090731.

Dedrick, S., Akbari, M. J., Dyckman, S. K., Zhao, N., Liu, Y.-Y., and Momeni, B. Impact of Temporal pH Fluctuations on the Coexistence of Nasal Bacteria in an in silico Community. Frontiers in Microbiology, 12 (February):1–12, 2021. ISSN 1664302X. doi: 10.3389/fmicb.2021.613109.

Dunn, R. R., Harris, N. C., Colwell, R. K., Koh, L. P., and Sodhi, N. S. The sixth mass coextinction: are most endangered species parasites and mutualists? Proceedings of the Royal Society B: Biological

Sciences, 276(1670):3037–3045, 9 2009. ISSN 0962-8452. doi: 10.1098/rspb.2009.0413. URL https://royalsocietypublishing.org/doi/10.1098/rspb.2009.0413.

Ellner, S. P., Snyder, R. E., Adler, P. B., and Hooker, G. An expanded modern coexistence theory for empirical applications. Ecology Letters, 22(1):3–18, 1 2019. ISSN 1461-023X. doi: 10.1111/ele.13159. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.13159.

Engen, S. and Lande, R. Population Dynamic Models Generating Species Abundance Distributions of the Gamma Type. Journal of Theoretical Biology, 178(3):325–331, 2 1996. ISSN 00225193. doi: 10.1006/jtbi.1996.0028. URL https://linkinghub.elsevier.com/retrieve/pii/S0022519396900284.

Ewens, W. J. Mathematical population genetics. Springer, 2004. doi: 10.1007/978-0-387-21822-9.

Fox, J. W. The intermediate disturbance hypothesis should be abandoned. Trends in Ecology & Evolution, 28 (2):86–92, 2 2013. ISSN 01695347. doi: 10.1016/j.tree.2012.08.014. URL https://linkinghub.elsevier.com/retrieve/pii/S0169534712002091.

Fung, T., Villain, L., and Chisholm, R. A. Analytical formulae for computing dominance from species-abundance distributions. Journal of Theoretical Biology, 386:147–158, 2015. ISSN 10958541. doi: 10.1016/j.jtbi.2015.09.011. URL http://dx.doi.org/10.1016/j.jtbi.2015.09.011.

Gans, J. Computational Improvements Reveal Great Bacterial Diversity and High Metal Toxicity in Soil. Science, 309(5739):1387–1390, 8 2005. ISSN 0036-8075. doi: 10.1126/science.1112665. URL https://www.sciencemag.org/lookup/doi/10.1126/science.1112665.

Gilbert, B. and Levine, J. M. Ecological drift and the distribution of species diversity. Proceedings of the Royal Society B: Biological Sciences, 284(1855):20170507, 5 2017. ISSN 0962-8452. doi: 10.1098/rspb.2017.0507. URL https://royalsocietypublishing.org/doi/10.1098/rspb.2017.0507.

Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. Journal of Physical Chemistry, 81 (25):2340–2361, 1977. ISSN 00223654. doi: 10.1021/j100540a008.

Goldberg, Y. and Friedman, J. Positive interactions within and between populations decrease the likelihood of evolutionary rescue. bioRxiv, 2020. ISSN 26928205. doi: 10.1101/2020.08.06.239608. URL http://dx.doi.org/10.1371/journal.pcbi.1008732.

Grilli, J. Macroecological laws describe variation and diversity in microbial communities. Nature Communications, 11(1):1–11, 2020. ISSN 20411723. doi: 10.1038/s41467-020-18529-y. URL http://dx.doi.org/10.1038/s41467-020-18529-y.

Grime, J. P. Competitive Exclusion in Herbaceous Vegetation. Nature, 242(5396):344–347, 3 1973. ISSN 0028-0836. doi: 10.1038/242344a0. URL http://www.nature.com/doifinder/10.1038/242344a0.

Grime, J. P. Evidence for the Existence of Three Primary Strategies in Plants and Its Relevance to Ecological and Evolutionary Theory. The American Naturalist, 111(982):1169–1194, 1977. ISSN 0003-0147. doi: 10.1086/283244.

Guittar, J., Koffel, T., Shade, A., Klausmeier, C. A., and Litchman, E. Resource Competition and Host Feedbacks Underlie Regime Shifts in Gut Microbiota. The American Naturalist, 198(1):000–000, 5 2021. ISSN 0003-0147. doi: 10.1086/714527. URL https://www.journals.uchicago.edu/doi/10.1086/714527.

Hengge-Aronis, R. Survival of hunger and stress: The role of rpoS in early stationary phase gene regulation in E. coli. Cell, 72(2):165–168, 1 1993. ISSN 00928674. doi: 10.1016/0092-8674(93)90655-A. URL https://linkinghub.elsevier.com/retrieve/pii/009286749390655A.

Himeoka, Y. and Mitarai, N. Dynamics of bacterial populations under the feast-famine cycles. arXiv, pages 1–39, 10 2019. URL http://arxiv.org/abs/1910.05673.

Hoek, T. A., Axelrod, K., Biancalani, T., Yurtsev, E. A., Liu, J., and Gore, J. Resource Availability Modulates the Cooperative and Competitive Nature of a Microbial Cross-Feeding Mutualism. PLOS Biology, 14(8):e1002540, 8 2016. ISSN 1545-7885. doi: 10.1371/journal.pbio.1002540. URL http://dx.plos.org/10.1371/journal.pbio.1002540.

Horsthemke, W. and Lefever, R. Noise-Induced Transitions, volume 15 of Springer Series in Synergetics. Springer Berlin Heidelberg, 2nd edition, 4 2006. ISBN 978-3-540-11359-1. doi: 10.1007/3-540-36852-3. URL http://link.springer.com/10.1007/3-540-36852-3.

Huang, Q., Parshotam, L., Wang, H., Bampfylde, C., and Lewis, M. A. A model for the impact of contaminants on fish population dynamics. Journal of Theoretical Biology, 334:71–79, 2013. ISSN 00225193. doi: 10.1016/j.jtbi.2013.05.018. URL http://dx.doi.org/10.1016/j.jtbi.2013.05.018.

Jost, L. PARTITIONING DIVERSITY INTO INDEPENDENT ALPHA AND BETA COMPONENTS. Ecology, 88(10):2427–2439, 10 2007. ISSN 0012-9658. doi: 10.1890/06-1736.1. URL http://doi.wiley.com/10.1890/06-1736.1.

Kalyuzhny, M., Kadmon, R., and Shnerb, N. M. A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities. Ecology Letters, 18(6):572–580, 6 2015. ISSN 1461023X. doi: 10.1111/ele.12439. URL http://doi.wiley.com/10.1111/ele.12439.

Kamenev, A., Meerson, B., and Shklovskii, B. How Colored Environmental Noise Affects Population Extinction. Physical Review Letters, 101(26):268103, 12 2008. ISSN 0031-9007. doi: 10.1103/PhysRevLett.101.268103. URL https://link.aps.org/doi/10.1103/PhysRevLett.101.268103.

Larsen, H. Halophilic and halotolerant microorganisms-an overview and historical perspective. FEMS Microbiology Reviews, 39(April):3–7, 1986.

Leigh, E. G. The average lifetime of a population in a varying environment. Journal of Theoretical Biology, 90(2):213–239, 5 1981. ISSN 00225193. doi: 10.1016/0022-5193(81)90044-8. URL https://linkinghub.elsevier.com/retrieve/pii/0022519381900448.

León, J. A. and Tumpson, D. B. Competition between two species for two complementary or substitutable resources. Journal of Theoretical Biology, 50(1):185–201, 1975. ISSN 10958541. doi: 10.1016/0022-5193(75)90032-6.

Letten, A. D., Dhami, M. K., Ke, P. J., and Fukami, T. Species coexistence through simultaneous fluctuation-dependent mechanisms. Proceedings of the National Academy of Sciences of the United States of America, 115(26):6745–6750, 2018a. ISSN 10916490. doi: 10.1073/pnas.1801846115.

Letten, A. D., Dhami, M. K., Ke, P. J., and Fukami, T. Species coexistence through simultaneous fluctuation-dependent mechanisms. Proceedings of the National Academy of Sciences of the United States of America, 115(26):6745–6750, 2018b. ISSN 10916490. doi: 10.1073/pnas.1801846115.

Li, G., Xie, C., Lu, S., Nichols, R. G., Tian, Y., Li, L., Patel, D., Ma, Y., Brocker, C. N., Yan, T., Krausz, K. W., Xiang, R., Gavrilova, O., Patterson, A. D., and Gonzalez, F. J. Intermittent Fasting Promotes White Adipose Browning and Decreases Obesity by Shaping the Gut Microbiota. Cell Metabolism, 26(4):672–685, 2017. ISSN 19327420. doi: 10.1016/j.cmet.2017.08.019. URL http://dx.doi.org/10.1016/j.cmet.2017.08.019.

Lin, Y. H., Bayrock, D. P., and Ingledew, W. M. Evaluation of Saccharomyces cerevisiae grown in a multistage chemostat environment under increasing levels of glucose. Biotechnology Letters, 24(6):449–453, 2002. doi: 10.1023/A:1014501125355.

Mackey, R. L. and Currie, D. J. The diversity-disturbance relationship: Is it generally strong and peaked? Ecology, 82(12):3479–3492, 2001. ISSN 00129658. doi: 10.1890/0012-9658(2001)082[3479:TDDRII]2.0.CO;2.

Marsland III, R., Cui, W., Goldford, J., Sanchez, A., Korolev, K., and Mehta, P. Available energy fluxes drive a transition in the diversity, stability, and functional structure of microbial communities. PLOS Computational Biology, 15(2):e1006793, 2 2019. doi: 10.1371/journal.pcbi.1006793. URL http://dx.plos.org/10.1371/journal.pcbi.1006793.

Merritt, J. and Kuehn, S. Frequency- and Amplitude-Dependent Microbial Population Dynamics during Cycles of Feast and Famine. Physical Review Letters, 121(9):098101, 8 2018. ISSN 0031-9007. doi: 10.1103/PhysRevLett.121.098101. URL https://link.aps.org/doi/10.1103/PhysRevLett.121.098101.

Miller, A. D., Roxburgh, S. H., and Shea, K. How frequency and intensity shape diversity-disturbance relationships. Proceedings of the National Academy of Sciences of the United States of America, 108(14):5643–5648, 2011. ISSN 10916490. doi: 10.1073/pnas.1018594108.

Molinero, N., Ruiz, L., Sánchez, B., Margolles, A., and Delgado, S. Intestinal bacteria interplay with bile and cholesterol metabolism: Implications on host physiology. Frontiers in Physiology, 10(MAR):1–10, 2019. ISSN 1664042X. doi: 10.3389/fphys.2019.00185.

Mougi, A. and Kondoh, M. Diversity of Interaction Types and Ecological Community Stability. Science, 337 (6092):349–351, 2012. ISSN 0036-8075. doi: 10.1126/science.1220529. URL http://www.sciencemag.org/cgi/doi/10.1126/science.1220529.

Muturi, E. J., Donthu, R. K., Fields, C. J., Moise, I. K., and Kim, C.-H. Effect of pesticides on microbial communities in container aquatic habitats. Scientific Reports, 7(1):44565, 4 2017. ISSN 2045-2322. doi: 10.1038/srep44565. URL http://www.nature.com/articles/srep44565.

Nguyen, J., Fernandez, V., Pontrelli, S., Sauer, U., Ackermann, M., and Stocker, R. A distinct growth physiology enhances bacterial growth under rapid nutrient fluctuations. bioRxiv, page 2020.08.18.256529, 2020a. URL https://www.biorxiv.org/content/10.1101/2020.08.18.256529v1.

Nguyen, J., Lara-Gutiérrez, J., and Stocker, R. Environmental fluctuations and their effects on microbial communities, populations, and individuals. FEMS Microbiology Reviews, 12 2020b. ISSN 0168-6445. doi: 10.1093/femsre/fuaa068. URL https://academic.oup.com/femsre/advance-article/doi/10.1093/femsre/fuaa068/6041721.

Novick, A. and Szilard, L. Experiments with the Chemostat on Spontaneous Mutations of Bacteria. Proceedings of the National Academy of Sciences, 36(12):708–719, 12 1950. ISSN 0027-8424. doi: 10.1073/pnas.36.12.708. URL http://www.pnas.org/cgi/doi/10.1073/pnas.36.12.708.

Novozhilov, A. S., Karev, G. P., and Koonin, E. V. Biological applications of the theory of birth-and-death processes. Briefings in Bioinformatics, 7(1):70–85, 3 2006. ISSN 1467-5463. doi: 10.1093/bib/bbk006. URL https://academic.oup.com/bib/article/7/1/70/263777.

Oren, A. Microbial life at high salt concentrations: Phylogenetic and metabolic diversity. Saline Systems, 4(1):1–13, 2008. ISSN 17461448. doi: 10.1186/1746-1448-4-2.

Pérez, S., Eichhorn, P., and Aga, D. S. Evaluating the biodegradability of sulfamethazine, sulfamethoxazole, sulfathiazole, and trimethoprim at different stages of sewage treatment. Environmental Toxicology and Chemistry, 24(6):1361, 2005. ISSN 0730-7268. doi: 10.1897/04-211R.1. URL http://doi.wiley.com/10.1897/04-211R.1.

Piccardi, P., Vessman, B., and Mitri, S. Toxicity drives facilitation between 4 bacterial species. Proceedings of the National Academy of Sciences, 116(32):15979–15984, 8 2019. ISSN 0027-8424. doi: 10.1073/pnas.1906172116. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1906172116.

Ratzke, C., Barrere, J., and Gore, J. Strength of species interactions determines biodiversity and stability in microbial communities. Nature Ecology and Evolution, 4(3):376–383, 2020. ISSN 2397334X. doi: 10.1038/s41559-020-1099-4. URL http://dx.doi.org/10.1038/s41559-020-1099-4.

Rodríguez-Verdugo, A., Vulin, C., and Ackermann, M. The rate of environmental fluctuations shapes ecological dynamics in a two-species microbial system. Ecology Letters, 22(5):838–846, 5 2019. ISSN 1461-023X. doi: 10.1111/ele.13241. URL http://doi.wiley.com/10.1111/ele.13241.

Roesch, L. F., Fulthorpe, R. R., Riva, A., Casella, G., Hadwin, A. K., Kent, A. D., Daroub, S. H., Camargo, F. A., Farmerie, W. G., and Triplett, E. W. Pyrosequencing enumerates and contrasts soil microbial diversity. ISME Journal, 1(4):283–290, 2007. ISSN 17517362. doi: 10.1038/ismej.2007.53.

Roughgarden, J. Theory of population genetics and evolutionary ecology : an introduction. Macmillan, New York, 1979. ISBN 0024031801.

Ruiz, L., Margolles, A., and Sánchez, B. Bile resistance mechanisms in Lactobacillus and Bifidobacterium. Frontiers in Microbiology, 4(DEC):1–8, 2013. ISSN 1664302X. doi: 10.3389/fmicb.2013.00396.

Schreiber, S., Levine, J., Godoy, O., Kraft, N., and Hart, S. Does deterministic coexistence theory matter in a finite world? bioRxiv, 2020. doi: 10.1101/290882.

Smits, S. A., Leach, J., Sonnenburg, E. D., Gonzalez, C. G., Lichtman, J. S., Reid, G., Knight, R., Manjurano, A., Changalucha, J., Elias, J. E., Dominguez-Bello, M. G., and Sonnenburg, J. L. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. Science, 357(6353):802–805, 2017. ISSN 10959203. doi: 10.1126/science.aan4834.

Spalding, C., Doering, C. R., and Flierl, G. R. Resonant activation of population extinctions. Physical Review E, 96(4):042411, 10 2017. ISSN 2470-0045. doi: 10.1103/PhysRevE.96.042411. URL http://dx.doi.org/10.1103/PhysRevE.96.042411.

Srinivasan, S. and Kjelleberg, S. Cycles of famine and feast: The starvation and outgrowth strategies of a marine Vibrio. Journal of Biosciences, 23(4):501–511, 1998. ISSN 02505991. doi: 10.1007/BF02936144.

Sun, Z., Koffel, T., Stump, S. M., Grimaud, G. M., and Klausmeier, C. A. Microbial cross-feeding promotes multiple stable states and species coexistence, but also susceptibility to cheaters. Journal of Theoretical Biology, 465:63–77, 2019. ISSN 00225193. doi: 10.1016/j.jtbi.2019.01.009. URL https://linkinghub.elsevier.com/retrieve/pii/S0022519319300104.

Sunya, S., Bideaux, C., Molina-Jouve, C., and Gorret, N. Short-term dynamic behavior of Escherichia coli in response to successive glucose pulses on glucose-limited chemostat cultures. Journal of Biotechnology, 164 (4):531–542, 2013. ISSN 01681656. doi: 10.1016/j.jbiotec.2013.01.014. URL http://dx.doi.org/10.1016/j.jbiotec.2013.01.014.

Taitelbaum, A., West, R., Assaf, M., and Mobilia, M. Population Dynamics in a Changing Environment: Random versus Periodic Switching. Physical Review Letters, 125(4):048105, 7 2020. ISSN 0031-9007. doi: 10.1103/PhysRevLett.125.048105. URL https://link.aps.org/doi/10.1103/PhysRevLett.125.048105.

Tecon, R., Mitri, S., Ciccarese, D., Or, D., van der Meer, J. R., and Johnson, D. R. Bridging the Holistic-Reductionist Divide in Microbial Ecology. mSystems, 4(1):17–21, 2019. ISSN 2379-5077. doi: 10.1128/msystems.00265-18.

Thaiss, C. A., Zeevi, D., Levy, M., Zilberman-Schapira, G., Suez, J., Tengeler, A. C., Abramson, L., Katz, M. N., Korem, T., Zmora, N., Kuperman, Y., Biton, I., Gilad, S., Harmelin, A., Shapiro, H., Halpern, Z., Segal, E., and Elinav, E. Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. Cell, 159(3):514–529, 2014. ISSN 10974172. doi: 10.1016/j.cell.2014.09.048. URL http://dx.doi.org/10.1016/j.cell.2014.09.048.

Vasi, F., Travisano, M., and Lenski, R. E. Long-term experimental evolution in Escherichia coli. II.Changes in life-history traits during adaptation to a seasonal environment. American Naturalist, 144(3):432–456, 1994. ISSN 00030147. doi: 10.1086/285685. URL https://www.jstor.org/stable/2462954.

West, R. and Mobilia, M. Fixation properties of rock-paper-scissors games in fluctuating populations. Journal of Theoretical Biology, 491:110135, 4 2020. ISSN 00225193. doi: 10.1016/j.jtbi.2019.110135. URL http://dx.doi.org/10.1016/j.jtbi.2019.110135.

Wienand, K., Frey, E., and Mobilia, M. Evolution of a Fluctuating Population in a Randomly Switching Environment. Physical Review Letters, 119(15):158301, 10 2017. ISSN 0031-9007. doi: 10.1103/PhysRevLett.119.158301. URL https://link.aps.org/doi/10.1103/PhysRevLett.119.158301.

Wienand, K., Frey, E., and Mobilia, M. Eco-evolutionary dynamics of a population with randomly switching carrying capacity. Journal of the Royal Society Interface, 15(145):20180343, 8 2018. ISSN 17425662. doi: 10.1098/rsif.2018.0343. URL https://royalsocietypublishing.org/doi/10.1098/rsif.2018.0343.

Xavier, J. B., Picioreanu, C., and Van Loosdrecht, M. C. M. A framework for multidimensional modelling of activity and structure of multispecies biofilms. Environmental Microbiology, 7(8):1085–1103, 2005. ISSN 14622912. doi: 10.1111/j.1462-2920.2005.00787.x.

Xu, B., Mao, D., Luo, Y., and Xu, L. Sulfamethoxazole biodegradation and biotransformation in the water-sediment system of a natural river. Bioresource Technology, 102(14):7069–7076, 2011. ISSN 09608524. doi: 10.1016/j.biortech.2011.04.086. URL http://dx.doi.org/10.1016/j.biortech.2011.04.086.

Zhao, Y. and Lin, Y. H. Growth of Saccharomyces cerevisiae in a chemostat under high glucose conditions. Biotechnology Letters, 25(14):1151–1154, 2003. ISSN 01415492. doi: 10.1023/A:1024577414157.

Zuñiga, C., Li, C.-T., Yu, G., Al-Bassam, M. M., Li, T., Jiang, L., Zaramela, L. S., Guarnieri, M., Betenbaugh, M. J., and Zengler, K. Environmental stimuli drive a transition from cooperation to competition in synthetic phototrophic communities. Nature Microbiology, 4(12):2184–2191, 12 2019. ISSN 2058-5276. doi: 10.1038/s41564-019-0567-6. URL http://www.nature.com/articles/s41564-019-0567-6.

Chapter 4

Spatial structures to maximize community functions

4.1 Brief summary

In this chapter, I return to the optimization of microbial community functions as in Chapter 2. While the previous two chapters analyze the effects of the environmental fluctuations (changing inflow media in a chemostat), this chapter investigates an alternative way to design optimal microbial communities: using networks of multi-stage chemostats, I consider the best spatial structures in which to organise microbial species. This chapter has two goals: (i) building algorithms that efficiently find the best allocations of microorganisms to multi-stage chemostats, and (ii) predicting the best allocation of microorganisms from experimental data. For the first goal, I introduce two simple algorithms whose computational costs are smaller than a brute force search. For the second goal, I show the best order of the following four bacterial species: *Agrobacterium tumefaciens*, *Comamonas testosteroni*, *Microbacterium saperdae*, and *Ochrobactrum anthropi* to degrade ampicillin.

4.2 Introduction

Co-culturing multiple species or strains has advantages over a mono-culture of a single species or a strain in biotechnology. For example, the production of biofuels is more efficient in co-cultures of multiple strains than in mono-cultures (Eiteman et al., 2008; Shin et al., 2010). In some cases, multi-species communities are suggested to have higher efficiency in bioremediation or productivity in biomass than mono-cultures of single species (Kazamia et al., 2014; Kang et al., 2021). For example, Demeter (2015) demonstrates that multi-species communities degrade more naphthenic acids than any mono-cultures of single species in the community. In addition, Dell'Anno et al. (2012) reported that the bioremediation efficiency of hydrocarbons positively correlates with species diversity.

However, the usage of multi-species[1] communities in biotechnology has a problem that species

---

1. Hereafter, I stick to the word "species" but the analyses in this chapter can also be used in multi-strain cases.

interactions can drive the communities to undesired states and decrease their community functions. In uranium bioremediation, for example, the efficiency of bioremediation decreases as time passes because *Geobacter* species, the main removers of uranium, are outcompeted by other species in three months (Anderson et al., 2003). Ecological theories have shown that the stable coexistence of many species is a non-trivial problem (May, 1972; Chesson, 2000a; Meszéna et al., 2006; Barabás et al., 2018; Landi et al., 2018; Amarasekare, 2019). One of the most important challenges in biotechnology is, therefore, how to stabilize the community dynamics while making community functions as high as possible.

There are three types of approaches to such optimization problems: modifying species compositions in communities, controlling dynamics by introducing environmental fluctuations, and allocating species into a pre-defined spatial structures, see Subsection 1.1.3 in Chapter 1 for more details. In this chapter, I use the third approach, which enables species to coexist even when they cannot coexist in a well-mixed structure (Kim et al., 2008). The potential obstacle of this approach is how to determine the positions of microbial species (i.e., microbial allocations). If we have N microbial species and allocate each species to one of σ positions, we have N^σ candidate allocations in total. Analyzing all of them (i.e., the brute force search) in experiments is time-consuming even with a small number of N and σ (e.g., $N = 4$ and $\sigma = 3$ gives 64 allocations). Running simulations would be more time-efficient than performing experiments, but the brute force search in simulations takes a long time with a large number of N or σ. In addition, simulations require specifying parameter values in the models but the estimated parameter values typically have some uncertainties which are given by confidence or credible intervals. In such cases, we need to run simulations many times with changing parameter values to include the uncertainties, which results in a large computational cost in total. Therefore, it is necessary to build an algorithm that can find the best microbial allocation efficiently.

In this chapter, I introduce an algorithm that can effectively find the best microbial allocations that maximize community functions. Here, I consider the multi-stage chemostats as an example of spatial structures. Multi-stage chemostats, chemostats connected by tubes to share the media, have been used in producing ethanol (Bayrock and Michael Ingledew, 2001; Lin et al., 2002) or proteases (Raninger and Steiner, 2003), and studying gut microbiota (Cinquin et al., 2006; Payne et al., 2012b). In this chapter, I do not consider species migration between chemostats assuming ideal filters that media can pass through but species cannot. I introduce a simple algorithm that can efficiently find the best microbial allocations that maximize target community functions in this chapter. In Appendix C.2, I improve this algorithm by introducing stochasticity. In the latter part of this chapter, I predict the best allocations of four microbial species, *A. tumefaciens*, *C. testosteroni*, *M. saperdae*, and *O. anthropi*, to degrade ampicillin using experimental data from Dos Santos (2019) and Oliveira Sudário (2022).
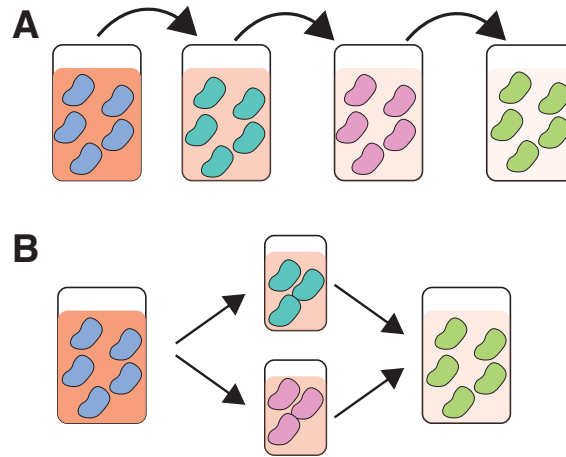
Figure 4.1 – Schematic representation of network structures

Schematic representations of acyclic networks of multi-stage chemostats. Different species are represented by different colors and each species are grown in each chemostat. In the main text, I analyze the acyclic networks without branches (A) while Appendix C.3 introduces the branches in the networks (B).

4.3 Model

4.3.1 Defining dynamics and community functions

To efficiently find the best allocations of microbial species that maximize targeted community functions, I first show a mathematical model of microbial species and chemical compounds in the multi-stage chemostats. In the main text, I assume that the network of the chemostats is non-branching and acyclic (i.e., linear or chain, Fig. 4.1A). This is because introducing branching (Fig. 4.1B) rarely improves community function but increases computational costs (see Appendix C.3). I do not consider chemostat networks with cycles because the motivation of introducing spatial structures is restricting species interactions; in acyclic networks, upstream species affect the growth of downstream species while downstream species cannot affect the growth of upstream species. This chapter explores two topics: (i) comparing the computational costs of my algorithms with a brute force search; i.e., analyzing all possible allocations, and (ii) predicting the best order of species combining experimental data from Dos Santos (2019) and Oliveira Sudário (2022).

I assume to allocate one of N species to stage (chemostat) $i = 1, \ldots, \sigma$ with allowing allocation of one species to multiple stages. The number of N is determined by how many species one can prepare to maximize the target community functions (i.e., N represents the number of species in the library). The number of stages σ can be increased to its maximum σ_m ($1 \leq \sigma \leq \sigma_m$). This is because there is no reason to always use as many chemostats as one has. In this system, M types of chemical compounds are considered. When we use σ chemostats, a community function can be defined by the difference in compounds' concentrations between inflow and outflow of the σ-stage

chemostat at a locally stable equilibrium:

$$\phi\left(\mathcal{A}; \alpha, \vec{w}\right) \equiv \alpha \sum_{j=1}^{M} w_j \left(C_{0j} - C_{\sigma j}^*\right), \tag{4.1}$$

where $\mathcal{A} = (A_1, \ldots, A_\sigma)$ represents a species allocation to each stage, $\vec{w} = (w_1, \ldots, w_M)$ represents a weight vector for each compound's concentration change ($w_j > 0$: aiming at degrading compound j, and $w_j < 0$: aiming at producing compound j), α is the dilution rate of the chemostats ($\alpha = 0.05$ for calculating computational costs of the algorithms, and $\alpha = 0.01$ for fitting experimental results), C_{0j} is compound j' concentration flowing into the first chemostat, and $C_{\sigma j}^*$ is compound j's outflow concentration from stage σ at an equilibrium state. Once an appropriate weight vector is given, the object function can deal with productions of biofuels, degradation of toxin, or combinations of them.

To evaluate the efficiency of the algorithms, I calculated the computational costs of them using the elapsed time of each simulations. The dynamics of chemical compounds and species k's abundance in stage i are given by the following consumer-resource model:

$$\frac{dC_{ij}}{dt} = \alpha \left(C_{i-1j} - C_{ij}\right) + f_{jk}\left(\vec{C}_i\right) S_k \tag{4.2a}$$

$$\frac{dS_k}{dt} = \left\{g_k\left(\vec{C}_i\right) - \alpha\right\} S_k \tag{4.2b}$$

where $f_{jk}\left(\vec{C}_i\right)$ is species k's impact vector on compound j, and $g_k\left(\vec{C}_i\right)$ is species k's growth function. Note that $\vec{C}_i = (C_{i1}, \ldots C_{iM})$ represent compound concentration vector in stage i. I implemented various forms of $f_{jk}\left(\vec{C}_i\right)$ and $g_k\left(\vec{C}_i\right)$ that would be applicable to many species-compound interactions in nature. Notably, I assume only species k is allocated to stage i, and allocated species cannot migrate from one chemostat to another because of filters between the chemostats in this chapters. The details of the mathematical model are explained in Appendix C.1.1.

Eqs (4.2a) and (4.2b) clarify that the dynamics at stage i depend on those at stage $i - 1$. To equilibriate the dynamics at stage i, it is necessary that the dynamics at stage $i - 1$ is equilibrated, see Appendix C.1.2 for more details. Therefore, it is convenient to assume that the initial compound concentration in stage i and the compound concentrations flowing into stage i are given by $\vec{C}_{i-1}^*$, the compounds' concentration vector flowing from the previous stage $i - 1$ at an equilibrium state. In short, I assume the following outflow

$$C_{i-1j} = \begin{cases} C_{0j} & (i = 1) \\ C_{i-1j}^* & \text{otherwise,} \end{cases} \tag{4.3}$$

and the initial condition

$$C_{ij}(0) = \begin{cases} C_{0j} & (i = 1) \\ C_{i-1j}^* & \text{otherwise,} \end{cases} \tag{4.4}$$

where $C_{0j} = 1$ if $w_j > 0$ otherwise $C_{0j} = 0$. This is an assumption for an ideal situation: a new species allocate to stage i after the dynamics in stage $i-1$ equilibriates. This assumption can be relaxed by modifying the dynamics (see Appendix C.4), which could affect the best allocations to maximize the community functions.

In the optimal allocation problem, we search the allocation that maximize the community function defined by Eq (4.1) at equilibrium using σ_m chemostats or less. In other words, the object function is written as follows:

$$\max \phi\left(\mathcal{A}; \alpha, \vec{w}\right)$$
$$\text{s.t. } \mathcal{A} = \left(A_1, \ldots, A_\sigma\right) \text{ stabilizes the dynamics Eqs(4.2a) and (4.2b)}$$
$$1 \leq \sigma \leq \sigma_m.$$

4.3.2 Computational cost of the brute force search

We pay attention to the computational costs of solving this optimization problem. If we use the brute force search, we allocate of each of N in stage $i = 1, \ldots, \sigma$, and the maximum number of stages is σ_m. Then, we need to analyze the following numbers of allocations:

$$\sum_{\sigma=1}^{\sigma_m} N^\sigma = \mathcal{O}\left(N^{\sigma_m}\right) \tag{4.5}$$

where $\mathcal{O}$ represents Landau notation. If we increase the number of species available N, the brute force search becomes experimentally and computationally unfeasible because the computational costs exponential increases over the maximum number of stages σ_m (see dashed lines in Fig. 4.2A).

4.3.3 Algorithm 1: the deterministic algorithm

Here, I introduce the deterministic algorithm, a kind of prune and search algorithm, into the optimal allocation problem (Algorithm 1). Suppose we have an allocation with length $i-1$ ($\mathcal{A} = (A_1, \ldots, A_{i-1})$) and want to add new species to stage i. Recall that the dynamics in stages $1, \ldots, i-1$ equbiliate before the species allocation to stage i. In stage i, we do not have to consider the allocation of species that is allocated to stage $i-1$ (i.e., $A_{i-1} \neq A_i$); the growth rate of this species in stage i is non-positive and therefore this species should go extinct. Then, there is no change in the concentrations of chemical compounds in stage i ($\vec{C}_i^* = \vec{C}_{i-1}^*$).

This rule can be expanded to other species: we do not have to analyze the allocation of species j to stage i if species j cannot survive in stage i at an equilibrium state. If allocated species goes extinct in stage i, we do not have to consider the allocations of species to stages $i+1, \ldots, \sigma$ and we can stop the search. In other words, it is sufficient to analyze the allocations that all-allocated species survive at a linearly stable equilibrium in the optimal allocation problems.

The upper boundary of the computational costs of this algorithm is easily derived. In the worst scenario, we can allocate N species in the first chemostat and $N-1$ species in the other

chemostats because we cannot allocate species j to stage i if species j is allocated to stage $i - 1$. In this case, the computational cost is given by $\sum_{\sigma=1}^{\sigma_m} N(N-1)^{\sigma-1}$. Therefore, the upper boundary of the computational cost of the deterministic algorithm exponentially increases over σ_m, although it is smaller than the computational cost of the brute force search.

4.3.4   Predicting best allocations from experimental data

Although I have considered the consumer-resource Eqs (4.2a) and (4.2b), this type of models is not always useful to fit experimental data. The consumer-resource models require information about how all species and all chemical compounds in the system interact. However, such data are not always available or easy to collect from experiments. For this reason, I used a generalized Lotka-Volterra (gLV) model to predict the best allocation of our four-species community, *A. tumefaciens*, *C. testosteroni*, *M. saperdae*, and *O. anthropi*to degrade ampicillin using three chemostats ($\sigma_m = 3$).

The gLV model needs to include the dynamics of species and chemical compounds that affect the target community functions ($|w_i| > 0$). I ignore the dynamic of chemical compounds that do not affect the community functions ($w_i = 0$) but can affect the growth of species. The effects from such compounds are summarized in a species interaction matrix. In this chapter, I consider the dynamics of the concentration of ampicillin and the four species:

$$\frac{dC_i}{dt} = \alpha \left(C_{i-1} - C_i \right) - \sum_{k=1}^{4} b_k \frac{C_i}{C_i + A_k} S_{ik} \tag{4.6a}$$

$$\frac{dS_{ik}}{dt} = r_k S_{ik} \left(K_k + \sum_{h=1}^{i-1} \sum_{l=1}^{4} a_{kl} S_{hl} - S_{ik} \right) - \alpha S_{ik} - d_k S_{ik} \frac{C_i}{C_i + A_k} \tag{4.6b}$$

where C_i represents the concentration of ampicillin in stage i, b_k represents species k's degradation efficiency of ampicillin, S_{ik} represent the abundance of species k in stage i (measured by colony forming unit, CFU), K_k represents the maximum population size of species k in mono-culture without ampicillin, A_k represents the concentration of ampicillin that gives the half-maximum effect of ampicillin on species k, r_k is the *re-parametarized* growth rate of species k so that $r_k K_k$ represents the maximum *per capita* growth rate of species k, a_{kl} represents the effect of species l on species k, and d_k represents the maximum death rate caused by ampicillin. Intraspecific interactions are given by $a_{kk} = -1$ for $k = 1, 2, 3, 4$.

To implement the model, I used the experimental data of mono-culture and pairwise co-culture data of the four species in batch culture (the dilution rate $\alpha = 0$), obtained by Andrea Dos Santos and Marina Oliveira Sudário, in Mitri group (Dos Santos, 2019; Oliveira Sudário, 2022). These species were grown in minimal media composed of glucose and citric acid, with or without ampicillin. The concentrations of ampicillin in the media was estimated by the growth of *A. tumefaciens* in the spent media. See Oliveira Sudário (2022) for the details of the experiments. The parameter values related to the mono-culture growth (i.e., intrinsic growth rates and carrying capacities) and interspecific interactions were estimated by Markov Chain Monte Carlo (MCMC)

Algorithm 1: Deterministic algorithm

---

1 Function Main($\vec{C}_0$, σ_m, N, α, $\vec{w}$):

2   Initialization: Let $\sigma = 0$ and $\mathcal{B}(\sigma) = \{\}$, which represents a list of allocations and outflow compound concentrations;

3 A_σ^* denotes the allocation whose community function is the largest in the allocations using σ stages;

4   **while** $\sigma \leq \sigma_m$ **do**

5 $\mathcal{B}(\sigma + 1) =$ADD_Species($\mathcal{B}(\sigma)$, σ);

6   Find $A_{\sigma+1}^*$ from $\mathcal{B}(\sigma + 1)$;

7 $\sigma = \sigma + 1$;

8   **end**

9   **return** $\max_\sigma \phi(A_\sigma^*)$;

10 Function ADD_Species($\mathcal{B}(\sigma)$, σ):

11   Initialize $\tilde{\mathcal{B}} = \{\}$;

12   **if** $\sigma = 0$ **then**

13   **for** $k = 1, \ldots, N$ **do**

14   Add species k to stage 1;

15   **if** *Species k persist* **then**

16   Add $\mathcal{A} = (k)$, its outflow compounds' concentrations, and its community function to $\tilde{\mathcal{B}}$;

17   **end**

18   **end**

19   **else**

20   **while** $\mathcal{B}$ *is not empty* **do**

21 $\tilde{\mathcal{A}}, \vec{C}_\sigma^* = \text{pop}(\mathcal{B})$;

22   Note that $\tilde{\mathcal{A}} = \left(\tilde{A}_1, \ldots, \tilde{A}_\sigma\right)$;

23   **for** $k = 1, \ldots, N$ **do**

24   **if** $k \neq \tilde{A}_\sigma$ **then**

25   Add species k to stage $\sigma + 1$;

26   **if** *Species k persist* **then**

27   Add new allocation $\left(\tilde{\mathcal{A}}, k\right)$, its outflow compounds' concentrations, and its community function to $\tilde{\mathcal{B}}$;

28   **end**

29   **end**

30   **end**

31   **end**

32   **end**

33   **return** return $\tilde{\mathcal{B}}$;

---

using pymc3 package (version 3.11.4) (Salvatier et al., 2016) in Python (version 3.8.8). One advantage of using MCMC is that it provides posterior probability distributions of the parameter values and the ampicillin concentrations in the media; we can thus intuitively calculate the uncertainty of the community function given a microbial allocation.

However, due to the limited number of growth data of species in the minimal media with ampicillin, I could not run MCMC to predict the posterior probability distributions of the parameters related to species-ampicillin interactions (i.e., resistance to the ampicillin and the degradation efficiency of ampicillin). To avoid this problem, I estimated the parameter values related to the species-ampicillin interactions using the least squares estimation (LSE) in lmfit package (version 1.0.2) (Newville et al., 2014) while fixing mono-culture and species interaction parameters as the means of their posterior probability distributions. For this reason, I could not estimate the probability distributions of community functions in this study.

4.4   Results

4.4.1   Comparison of computational costs

Fig. 4.2A shows the computational costs of the brute force search and the deterministic algorithm with various numbers of species N over the maximum number of chemostats σ_m. As expected from Eq (4.5), the computational costs of the brute force search exponentially increase over σ_m and the computational costs have small variations given N and σ_m. In contrast, the deterministic algorithm has larger variations in the computational costs than in the brute force search for each N and σ_m. This is because the number of analyzed allocations depends on the growth functions and the impact vectors of the species. Nevertheless, the average computational costs of the deterministic algorithm are smaller than the brute force search except for $(N, \sigma_m) = (1, 8)$.

Although the upper boundary of the deterministic algorithm 's computational cost exponentially increases with σ_m, Fig. 4.2B suggests that such worst cases rarely happen. As the maximum number of chemostats increases, the increase of computational costs gets smaller and the computational costs saturate around $\sigma_m = 2$ (when $N = 2$), 8 (when $N = 4$), and 10 (when $N = 8$). This implies that this algorithm analyzes a few numbers of allocations that use many chemostats. As the deterministic algorithm takes less time to find the brute force search and saturate its computational costs with a large number of σ_m, this algorithm efficiently finds the best microbial allocation. In Appendix C.2, I improved this algorithm by introducing the stochasticity on the choices of species allocation. This stochastic algorithm has smaller computational costs than the deterministic algorithm and able to find the best or quasi-best microbial allocations in most cases.

4.4.2   Parameter estimation of four species

To validate the concept of the algorithm, I investigated the best order of the four species, *A. tumefaciens*, *C. testosteroni*, *M. saperdae*, and *O. anthropi* to degrade ampicillin, using the
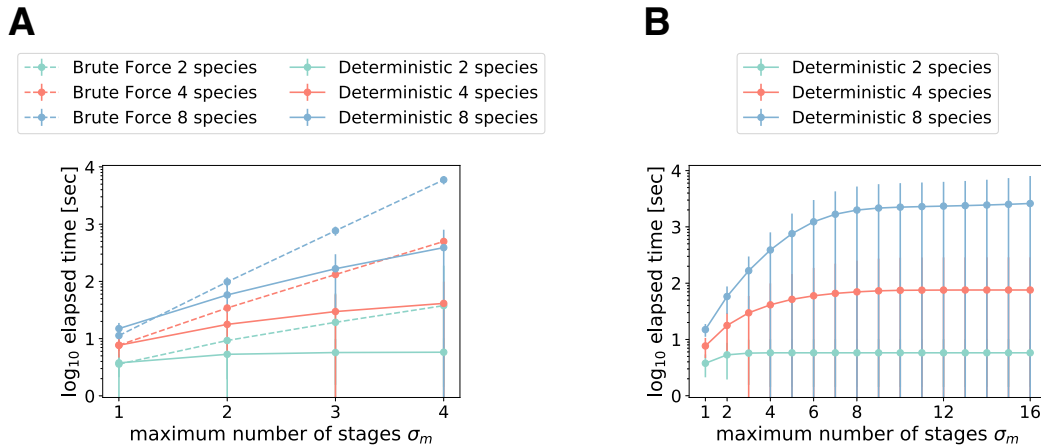
Figure 4.2 – Deterministic algorithm is more effective than the brute force search

Elapsed times of the brute force search and the deterministic search algorithms. A: Comparison of the brute force search and the deterministic search algorithms when $\sigma_m \leq 4$. Each dot and error bar represents the mean and the standard deviations of computational costs of 100 simulations (exceptionally, I used 30 simulations for the brute force search with 8 species due to the large computational costs) to find the best allocations. B: computational costs of the deterministic algorithm for larger numbers of chemostats $\sigma_m \leq 16$.

experimental data from Dos Santos (2019) and Oliveira Sudário (2022).

Fig 4.3 shows the co-culture data of all pairs (dashed lines) in the absence of ampicillin and the model prediction (solids lines) from the posterior probability distributions in batch culture. The model captures the pattern of the growth curves in the experimental data and the abundances of the four species at equilibrium. Interspecific interactions are summarized in Fig. 4.4 and Table 4.1, while the other parameters are summarized in Table 4.2. The dynamics of each species' mono-culture growth in the presence of ampicillin are shown in Fig 4.5. Except for *A. tumefaciens*'s dynamics, the model fits the mono-culture dynamics and the final estimated concentration of ampicillin. Two species, *C. testosteroni* and *O. anthropi* degrade ampicillin while *M. saperdae* cannot degrade at all. *A. tumefaciens* can also degrade ampicillin but its degradation efficiency is estimated lower than *C. testosteroni* and *O. anthropi*.
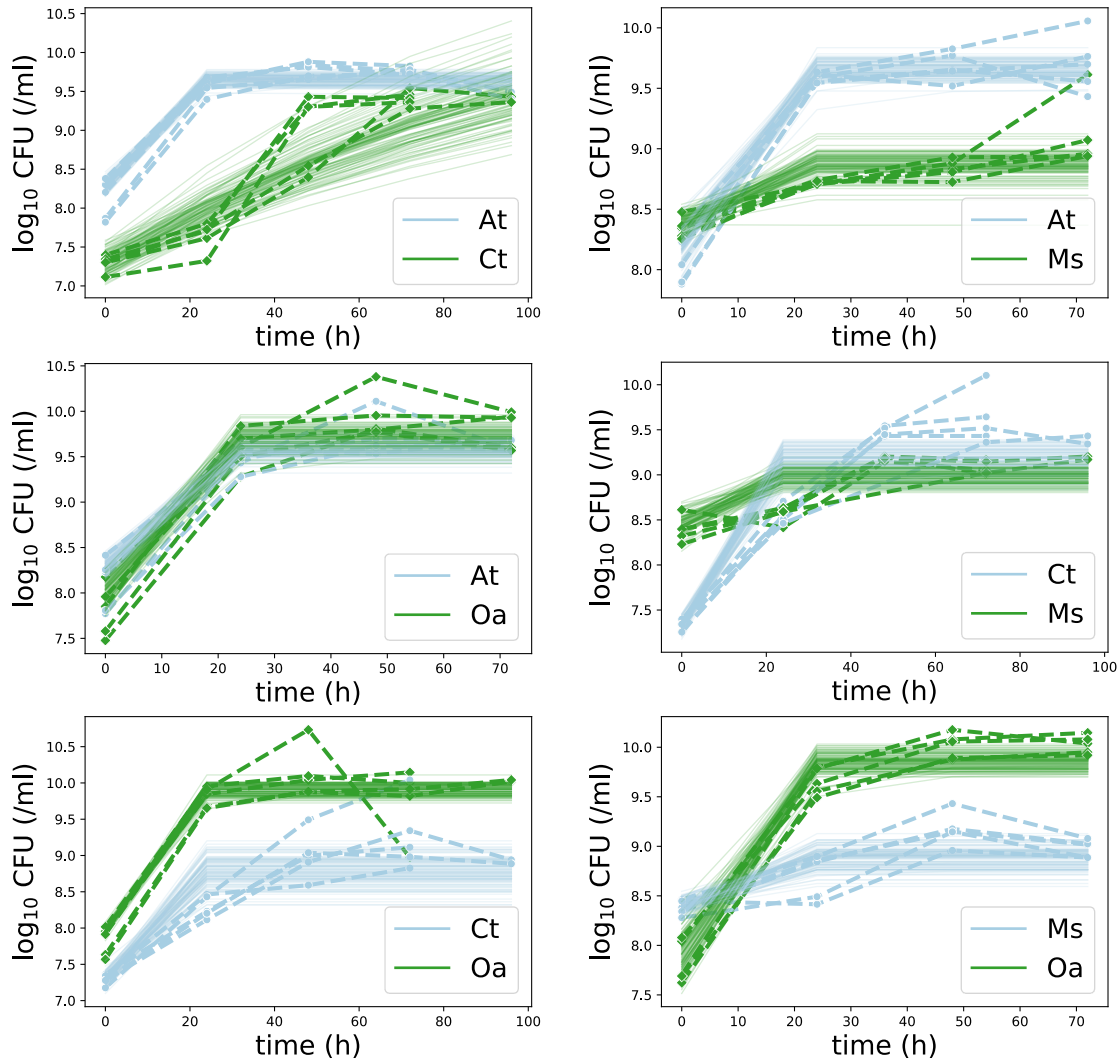
Figure 4.3 – Fitting to co-culture data without ampicillin

Each panel represents the fitting to pairwise co-culture data of our four bacteria. The dashed lines represent the experimental data (five replicates in each) from Dos Santos (2019) and Oliveira Sudário (2022) while the solid lines represent the simulations where the dynamics follow Eqs (4.6a) and (4.6b) with parameter values sampled from the posterior probability distributions (100 samples in each).

Table 4.1 – Estimated species interactions

to \from	A. tumefaciens	C. testosteroni	M. saperdae	O. anthropi
A. tumefaciens	-1	-0.007 $[-0.015, 0.002]$	-0.005 $[-0.017, 0.008]$	-0.010 $[-0.021, 0.002]$
C. testosteroni	0.180 $[0.124, 0.237]$	-1	0.052 $[0.024, 0.070]$	0.004 $[-0.020, 0.030]$
M. saperdae	0.048 $[0.035, 0.060]$	0.059 $[0.047, 0.072]$	-1	0.054 $[0.041, 0.069]$
O. anthropi	-0.029 $[-0.050, -0.012]$	-0.011 $[-0.023, 0.002]$	-0.014 $[-0.281, -0.001]$	-1

Element (i, j) represents interspecific interaction terms from species j to i. Each column represents the mean of the posterior probability distribution and 95% highest density interval (HDI) in off-diagonal elements. The diagonal elements are fixed -1 to scale the off-diagonal elements. See also Fig. 4.4.

Table 4.2 – Estimated parameter values except for species interactions

	r_i	K_i	b_i	d_i	M_i
A. tumefaciens	1.4	9.7	1.8	9.9	351.9
C. testosteroni	1.5	9.0	1.7	6.5	27.1
M. saperdae	1.6	8.4	0.0	0.4	146.9
O. anthropi	1.4	10.0	1.5	7.2	105.0

Parameter values estimated by LSE (b_i, d_i, and K_i) and mean of the posterior probability distributions (r_i and K_i).
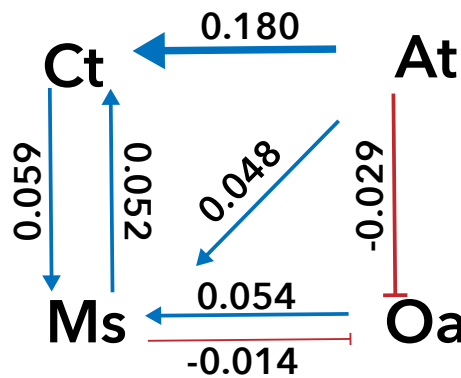
Figure 4.4 – Summary of species interactions

Species interactions estimated by our experiments without ampicillin. The width, color, and head type of each arrow are determined by the mean value of the posterior probability distribution of each parameter. I ignored the interactions whose 95% HDI includes zeros because the effects of these interactions are very weak and uncertain. See also Table 4.1 for details.At: *A. tumefaciens*, Ct: *C. testosteroni*, Ms: *M. saperdae*, and Oa: *O. anthropi*.

4.4.3   Predicting the best allocations for degrading ampicillin

Fig. 4.6 indicates the predicted concentrations of ampicillin left in the media given allocations when I assume the idealized setup (i.e., allocating species after the dynamics in the previous stage equilibrate). While the degradation efficiencies of *C. testosteroni* and *O. anthropi* are estimated similar in batch culture ($\alpha = 0$, Fig. 4.5), the degradation efficiency of *C. testosteroni* is higher than *O. anthropi* in 1-stage chemostat ($\alpha = 0.01$), probably due to the differences in d_k and M_k (Fig. 4.6A). In the 2-stage cases, the allocations using *C. testosteroni*and *O. anthropi* degrade ampicillin efficiently, but their orders of the species matter: allocating *O. anthropi* first degrades more ampicillin than allocating *C. testosteroni* first. This is because *C. testosteroni* has a negative effect on *O. anthropi* while *O. anthropi* has a positive one on *C. testosteroni* although these interspecific interactions are weak (Table 4.2). In the 3-stage cases, allocating *C. testosteroni*
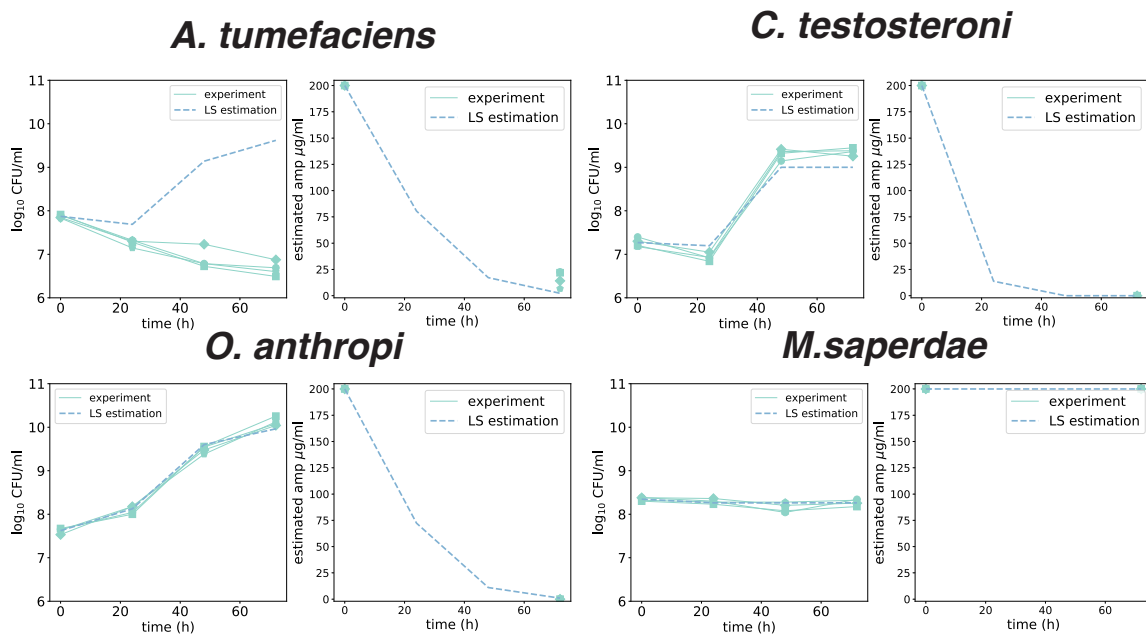
Figure 4.5 – Fitting to the mono-culture experiments with ampicilin

Top Left: *A. tumefaciens*, Top right: *C. testosteroni*, Bottom Left: *O. anthropi*, and Bottom right: *M. saperdae*. In each, left panels show the growth of focal species while right panels show the estimated ampicillin concentrations. Parameter values are summarized in Tables 4.2 and 4.1. The green solid lines represent experimental data, where each symbol corresponds to each experimental replicate. The dashed lines show the dynamics of the simulations with the parameter values estimated by LSE. I do not show the green solid lines in the right panels (ampicillin) to emphasize that we have missing data at 24 and 48 hours.

to the first and third stages, and *O. anthropi* to the second stage maximizes the degradation efficiency of ampicillin, which would be intuitive from the results in 1- and 2-stage scenarios.

Interestingly, the second-, third-, and fourth-best allocations in the 3-stage scenario use *A. tumefaciens*, and *C. testosteroni* is always allocated after *A. tumefaciens*. This would be because *A. tumefaciens* has a strong positive effect on *C. testosteroni*, which increases the degradation efficiency of *C. testosteroni* in the next stage. This strong positive interaction from *A. tumefaciens* to *C. testosteroni* would be why the second-, third-, and fourth-best allocations degrade ampicillin more than the sixth-best one, which allocates *O. anthropi* to the first and third chemostats and *C. testosteroni* to the second chemostat. This indicates that using efficient degradaters in mono-culture is not always a good strategy.

I also predicted the degradation efficiencies of three of the four species or all-four species co-cultures in a chemostat (Fig. 4.6D). In these cases, we cannot use two or more chemostats because no species can grow in the second chemostat. The model predicts that the degradation efficiencies of the mixtures of the three or four species are lower than the best allocations in 2- and 3-stage models (compare Fig 4.6D with Figs. 4.6B and C, respectively). These results indicate the advantages of introducing spatial structure over mixing species, although this does not hold in another condition (Appendix C.4).
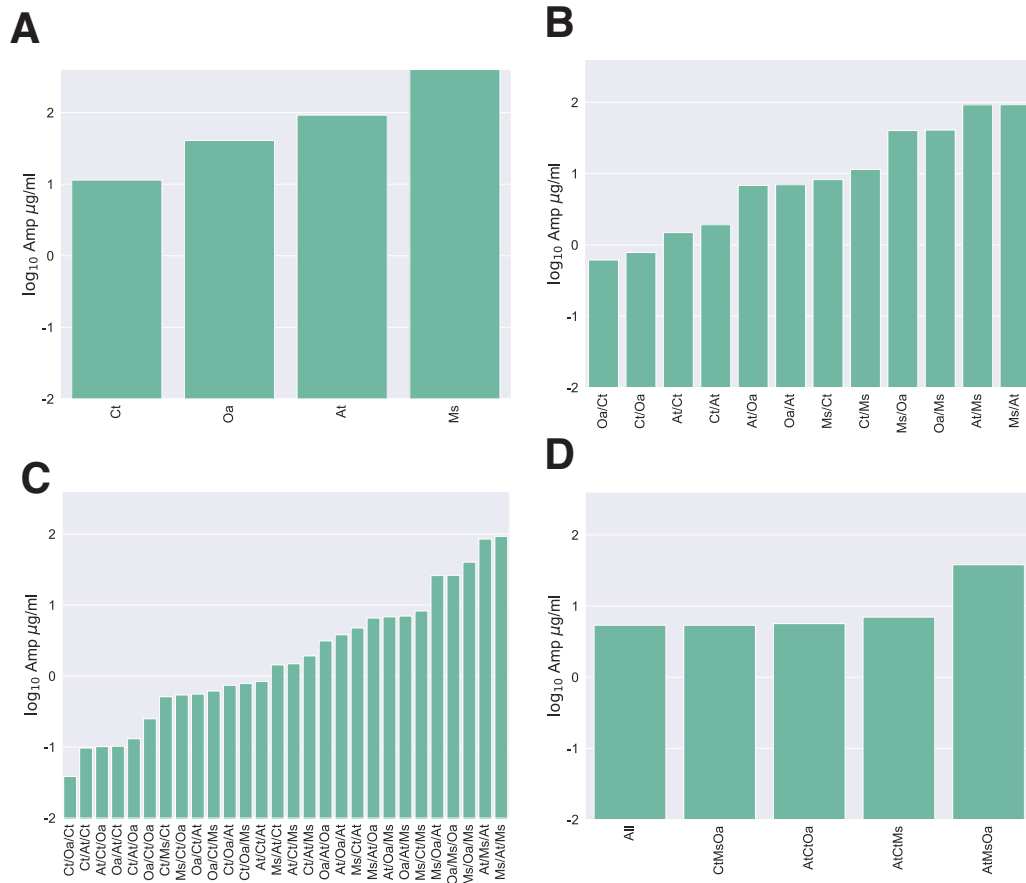
Figure 4.6 – Estimated ranks of detoxification in the idealized setup

The predicted concentrations of ampicillin left in the outflow media at equilibrium in the idealized setup. The number of stages are A: $\sigma = 1$, B: $\sigma = 2$, C: $\sigma = 3$. D: mixtures of three or four species with $\sigma = 1$; species cannot grow $\sigma \geq 2$ in this case. Allocation X/Y/Z means we allocate species X to the first chemostat, Y to the second, and Z to the third. At: *A. tumefaciens*, Ct: *C. testosteroni*, Ms: *M. saperdae*, Oa: *O. anthropi*.

4.5 Discussion

In this study, I first introduced the deterministic algorithm, which is a kind of prune and search algorithm. This algorithm can find the best microbial allocation effectively. The motivation behind is that the brute force search is not feasible both in experiments and simulations. In experiments, it is obvious that performing experiments in all (i.e., roughly N^{σ_m}) allocations is time-consuming. In simulations, the brute force search might be feasible when all parameter values are specified; however, the parameter values have some uncertainty in most cases and thus we need to repeat the brute force search with various parameter values to include such uncertainty, which is not feasible in the end. The deterministic algorithm can more efficiently find the optimal allocations than the brute force search and thus its total computational cost to include parameters' uncertainty would not be too large. In the latter part of this chapter, I used the deterministic algorithm to find the best allocation of four species – *A. tumefaciens*, *C. testosteroni*, *O. anthropi*, and *M. saperdae* – to degrade ampicillin using the experimental data from Dos Santos (2019) and Oliveira Sudário (2022). After statistical estimation of the parameter values, the best order of our four species is predicted as allocating species in the following order: *C. testosteroni*, *O. anthropi*, and *C. testosteroni*.

One advantage of the deterministic algorithm is that we can always find the best microbial allocation with this algorithm. For example, George and Korolev (2021) analyze the similar problem without spatial structures (i.e., finding the best species compositions) using a kind of greedy algorithm, but this algorithm sometime fails in finding the best compositions and suggests local optima. Indeed, the greedy algorithm does not always work in the optimal allocation problems. In Fig. 4.6B for example, the best allocation is allocating *O. anthropi* first and *C. testosteroni* next, while the greedy algorithm suggests allocating *C. testosteroni* first and *O. anthropi* next: *C. testosteroni* degrades ampicillin the most when $\sigma = 1$ and then the second degrading species, *O. anthropi* is allocated to the second chemostat as *C. testosteroni* cannot growth there. The problem of the greedy algorithms is that it cannot include the species interaction effects: even if species cannot directly contribute to community functions, they can indirectly increase the community functions by facilitating the growth of downstream species that directly contribute to community functions. The second-, third- and fourth-best allocations in Fig. 4.6C are such examples because they use *A. tumefaciens*, which degrades ampicillin less than *C. testosteroni* and *O. anthropi* but strongly facilitates the growth of *C. testosteroni*(Fig. 4.4). The deterministic algorithm, on the other hand, analyze all allocations where all-allocated species survive because the species that go extinct cannot contribute to either the community functions or the growth of downstream species. For this reason, the deterministic algorithm always enables us to find the global maximum of the community function.

Of course, this advantage reflects the disadvantage of the deterministic algorithm: the upper boundary of the computational costs exponentially increase over the number of chemostats, although the exact computational costs depend on the parameter values in the system. Fig. 4.2 shows that such worst cases rarely happen and the computational costs saturate over the

maximum number of stages. This is because one cannot grow many species in stages with a large σ. In Fig. 4.6C, for example, the deterministic algorithm analyzes only 28 allocations from $4^3 = 64$ possible allocations. From its computational costs and the ability to always find the best allocation, the deterministic algorithm would be useful in reality.

Although I used the ecological dynamics (Eqs (4.2a) and (4.2b), or Eqs (4.6a) and (4.6b)) in the deterministic algorithm, this algorithm can be used with other mathematical models. The deterministic algorithm requires a map from the compound concentrations in inflow media and allocated species to the compound concentrations in outflow media at equilibrium. If genome-scale data are available, for example, one can use dynamics flux balance analysis as Xu et al. (2019) did in stead of the ecological dynamics. Extremely speaking, we may not need a dynamical model to use the deterministic model. With massive amount of data, one might be able to build a statistical model that predict the outflow media given the inflow medium and allocated species. Therefore, the choice of the model in the deterministic algorithm depends on the amount and type of available data.

In the latter part of this chapter, I use the deterministic algorithm to predict the best of the four species to degrade ampicillin combining experimental data from Dos Santos (2019) and Oliveira Sudário (2022). Although these predictions should be experimentally tested in future, there are some statistical problems. First, the experimental data of the ampicillin concentration in the media is estimated by the growth of *A. tumefaciens* in the spend media of one of the four species (Oliveira Sudário, 2022). The estimated concentrations of ampicillin may be biased due to the species interaction effects from species cultured in the media to *A. tumefaciens*. Another statistical problem is the limited number of data of mono-culture data in the media with ampicillin. Currently, we have only data of estimated ampicillin concentrations at the beginning and the end of experiments. This makes it difficult to perform statistical estimations on the parameters related to species-ampicillin interactions. Indeed, I could not perform MCMC with these data and used LSE instead. The growth of *A. tumefaciens* with ampicillin (Fig. 4.5) seems overestimated, which may result in overestimating the community function of the second-, third-, and fourth-best allocations in Fig. 4.6C. These problems could be solved by using, for example, Ampicillin ELISA Kit (BioVision Inc., California,, USA), which enables us to get more time-series data of ampicillin concentrations with higher accuracy. In addition, we could use the consumer-resource model, Eqs (4.2a) and (4.2b), instead of the gLV model, Eqs (4.6a) and (4.6b), if we could obtain the time series-data of compound concentrations in the media. Such data may be obtained by Gas Chromatography, for example. One advantage of the consumer-resource model is to explicitly include the environmentally mediated species interactions, while the gLV model cannot. If ampicillin can mediate species interactions, the prediction based on the consumer-resource model would be more accurate than those based on the Lotka-Volterra model. I suggest therefore that more experiments are needed before the experimental validation of Fig. 4.6.

In conclusion, I introduced the the deterministic algorithm that can always find the best allocations of microbes effectively. This algorithm searches the spatial organization of microbes using a model (e.g., ordinary differential equations) that predict the compound concentrations in

the outflow medium given the inflow medium and allocated species. As an example, I predicted the best order of four species to degrade ampicillin. Thanks to the small computational costs and the ability to always find the best allocations, the deterministic algorithm would be useful in other cases as well.

Chapter 5

Community stability in the hierarchical structure

5.1 Brief summary

This chapter continues to analyze microbial communities in a hierarchical spatial structure. This spatial structure can be seen as a simplification of a gastrointestinal tract. Although many studies in ecology assume that spatial structures affect species migration, we can also consider cases where the spatial structures affect species interactions. In a hierarchical network, upstream species can affect the dynamics of downstream species if species interactions are mediated by chemical compounds and the compounds flow from the upstream to the downstream. In this case, we can ask how upstream species affect the stability of downstream communities. I simulated meta-community dynamics composed of two patches in the hierarchical network and performed statistical analysis. Here, I defined stability as the probability that the downstream species composition is maintained after a species invasion either in the upstream community or the downstream one. The statistical analyses suggest that (i) assembled downstream communities tend to be more stable than randomly generated ones, (ii) migration from the outside of the meta-communities has the strongest effect on the downstream stability, and (iii) positive effects from the upstream to the downstream stabilize the downstream communities. These results can be used to design gut microbiota that are beneficial to hosts and stable over time.

5.2 Author contribution

This study was conceived by S.S. and S.M. The simulations and statistical analyses were performed by S.S. The figures were made by S.S. and S.M., partially using Biorender. The first draft was written by S.S, and revised by S.S. and S.M. The manuscript of this project appears in the next pages and will be submitted to a journal and bioRxiv once it is ready.

Stability analysis of gut microbiota in a chain network model

Shota Shibasaki[1] and Sara Mitri[1]

[1]Department of Fundamental Microbiology, University of Lausanne, Switzerland

July 15, 2022

- Short running title: Stability of gut microbiota

- Number of figures: 4

- Number of tables: 2

- Number of boxes: 0

- Keywords: ecological stability, causal inference, species interactions, meta-community dynamics

Abstract

Spatial structures are essential for meta-community dynamics and the stability of local communities. While many studies of meta-community dynamics assume species interactions within each patch, species could interact with those in connected patches if species interactions are mediated by chemical compounds that flow into the connected patches. A gut microbiota within a host is such an example with a hierarchical spatial structure. In this example, upstream species can interact with species in the same and downstream patches. By analyzing a mathematical model, we found that (i) assembled downstream communities are more stable than randomly generated ones, that (ii) species migration from the outside of the meta-communities has the strongest effect on the downstream stability, and that (iii) positive interactions from upstream to downstream species stabilize the downstream communities. These results can be used to design a gut microbiota that is beneficial to its host and stable over time.

1 Introduction

Spatial structures are important for local community dynamics in meta-community (Leibold et al., 2004). Previous studies show that spatial structures of meta-communities affect species invasions and diversity in local communities (Economo and Keitt, 2007; Chisholm et al., 2011; Seymour et al., 2015; Holenstein et al., 2022). In other words, the stability of local communities against invasion depends on the spatial networks. Although many studies of meta-community dynamics assume that species interact within each local community and that the spatial structures affect species migration, species could affect the growth of species in other connected communities, which can affect the stability of these communities. Let us consider microbial species interactions with environmental mediation: e.g., microbial species indirectly interact with one another by, for example, consuming resources or producing toxins (Estrela et al., 2019). If these environmental factors can flow into connected patches, microbial species can affect the growth of species that exist in the connected patches. These flows of the environmental factors are considered in the context of meta-ecosystem studies (Loreau et al., 2003; Gravel et al., 2010; Massol et al., 2011; Guichard, 2019). If the flows of the environmental factors are one-directional (i.e., hierarchical spatial structures), upstream microbial species can affect the growth of upstream and downstream species, while downstream species interact only with other downstream species. In this case, we can ask whether and how interactions from upstream species to downstream species affect the stability of downstream communities.

A gastrointestinal tract is an example of such a hierarchical spatial structure, which can be simplified as patches in a chain network (Cinquin et al., 2006; Payne et al., 2012). Although the temporal stability of human gut microbiota has been investigated (Costello et al., 2009; Dethlefsen and Relman, 2011; Lozupone et al., 2012; Byrd et al., 2020; Fassarella et al., 2021), the effects of the spatial structure on the stability of gut microbiota have been overlooked. While many recent studies suggest the relationship between gut microbiota and human health (Vrieze et al., 2012; Xu et al., 2013; Lee and Hase, 2014; Marchesi et al., 2016; Selber-Hnatiw et al., 2017; Sánchez et al., 2017; Ley et al., 2006; Sanna et al., 2019; Fan and Pedersen, 2021), these studies have inspired researchers to design gut microbiota to improve human health (Pham et al., 2017; Swann et al., 2020; Clark et al., 2021). As the designed gut microbiota should be beneficial to the hosts and stable over time, we need to consider how the spatial structure of a gut affects the stability of microbial communities there.

In this manuscript, we focus on the stability of the species compositions rather than the species abundances. This is because we can expect that the functions of the designed communities would not collapse when the species compositions are maintained, although the changes in the abundances of species may quantitatively affect the community functions. In ecology, many criteria of stability have been used (Landi et al., 2018) and we consider the stability by combining two criteria in this manuscript. One is the resistance to invasion (Post and Pimm, 1983): i.e., whether a new species (an invader) can colonize or not. We care this stability because the invaders may decrease the community functions or drive one or more species in the community to extinction. Previous studies suggest that species richness (i.e., the number of species in a community) increases the resistance to invasion both in theory (Case, 1990) and experiments (Stachowicz et al., 1999; Bonanomi et al., 2014; Hromada

et al., 2021). The mechanism behind is considered that a species-rich community exploits the resources and decreases the resource availability for invaders in the environment (Mallon et al., 2015).

The second stability criterion in this study is the resistance to the environmental changes caused by the upstream communities. As the upstream and downstream patches are connected, changes in the upstream community caused by species' invasion there lead to the environmental changes in the downstream patch, which may lead to the extinction of one or more species in the downstream community. This idea is similar to structural stability. In the context of ecology, structural stability refers to the volume of parameter space, typically that of species' growth rates given a species interaction matrix, where species composition does not change (Rohr et al., 2014; Butler and ODwyer, 2018; Song and Saavedra, 2018; Saavedra et al., 2017; Cenci and Saavedra, 2018; OSullivan et al., 2019). We can understand structural stability as the stability against perturbations in parameters that can be caused by, for example, environmental changes. However, our resistance to the environmental changes differs from structural stability in the sense that we consider gradual changes of the growth rates in the downstream patch over time, not sudden changes.

In this study, we investigate the stability of the downstream communities using a mathematical model. Although we do not pay attention to the stability of the upstream communities in the main text, we can easily examine the upstream stability by the analysis without the spatial structure: species richness increases the stability in this case. We simplified the spatial structure of a gastrointestinal tract by considering the chain network composed of two patches (the upstream and downstream patches). First, we introduce the definition of the stability and two scenarios from which we sampled communities: whether the meta-communities are naturally assembled or artificially generated. If the assembled communities tend to be stable, we can justify transferring microbiota from healthy hosts to unhealthy hosts (Vrieze et al., 2012). However, the data from the assembled meta-communities are difficult to statistically analyze because the data are not randomly sampled. Instead, we used the data of randomly generated meta-communities so that we can perform simple statistical analyses and infer the relationships between community features and the downstream stability. We also analyze how parameters related to species migration affect the downstream stability. Finally, we show that increasing the strength of positive interactions from an upstream community to a downstream community increases the downstream stability.

2 Results

2.1 Stability with no spatial structure

In this manuscript, we analyze the stability of downstream communities in the chain network with two patches. Before introducing this spatial structure, we first show the results without the spatial structure (i.e, we have only one patch) so that reader can understand the definition of the stability and two sampling scenarios. We define the stability in this manuscript as the probability that species composition (i.e., presence or absence of each species) does not change after an invasion of a randomly chosen species: see Eq (2). For any species
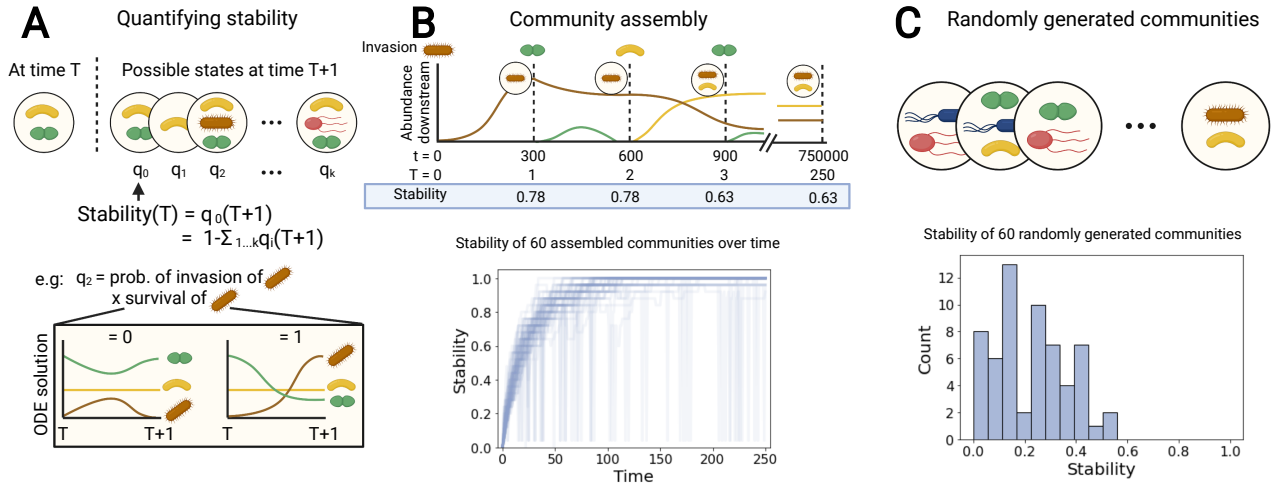
Figure 1: Stability with no spatial structure

(A) We define the stability of a community at time step T as the probability that the identical species composition is observed at time step $T+1$ (i.e., after invasion of one species and the stabilization of the dynamics), $q_0(T+1)$. Species composition can change if an invader establishes and/or one or more resident species go extinct. For example, the probability that a new species establishes ($q_2(T+1)$) is decomposed into the probability that the focal species invades the community , and whether the invader establishes or not (given by solving Eq (1b)). See Section 4.2 for more details. (B)We allowed 60 communities to assemble to simulate microbial assembly as it could occur in nature through 250 sequential invasion events, where stability is measured at each time step T. Between time steps T, community dynamics play out over 300 time steps t. The plot shows the changes in 60 community stability over the 250 time steps T. (C)In the designing scenario, we generated locally stable communities whose species composition is randomly chosen. The plot shows the distribution of randomly generated 60 communities. Created with BioRender.com.

composition, the stability at time step T is given by the probability ($q_0(T+1)$ in Fig. 1A) that the community has the identical species composition when the community dynamics equilibriate (i.e., solving ordinary differential equations, ODEs, until $t = 300$) after the invasion. Without the spatial structure, the stability represents the probability that an invader does not ether establish or remove the resident species.

We sampled communities' stability using two scenarios. One is the assembly scenario, where communities are assembled over time. Without the spatial structure, the species composition would not change once the stability reaches at one. Fig. 1B shows that the stability tends to increase over time except for some cases, where the stability suddenly decreases but recovers at the following time steps. In these cases, the dynamics equilibriate very slowly and $t = 300$ is not enough long to remove species that go extinct at equilibrium. While we can sample many communities in the assembly scenario (we sampled one community at each time step in each simulation), these samples are not independent: species composition at T depends on that at $T-1$, which makes it difficult to perform statistical analyses.

To overcome this problem, we used the the designing scenario (Fig. 1C): we randomly generated communities where species coexistence is locally stable. We used the data from this scenario to perform statistical analyses. In SI 2, we show that species richness increases stability in the absence of the spatial structure.
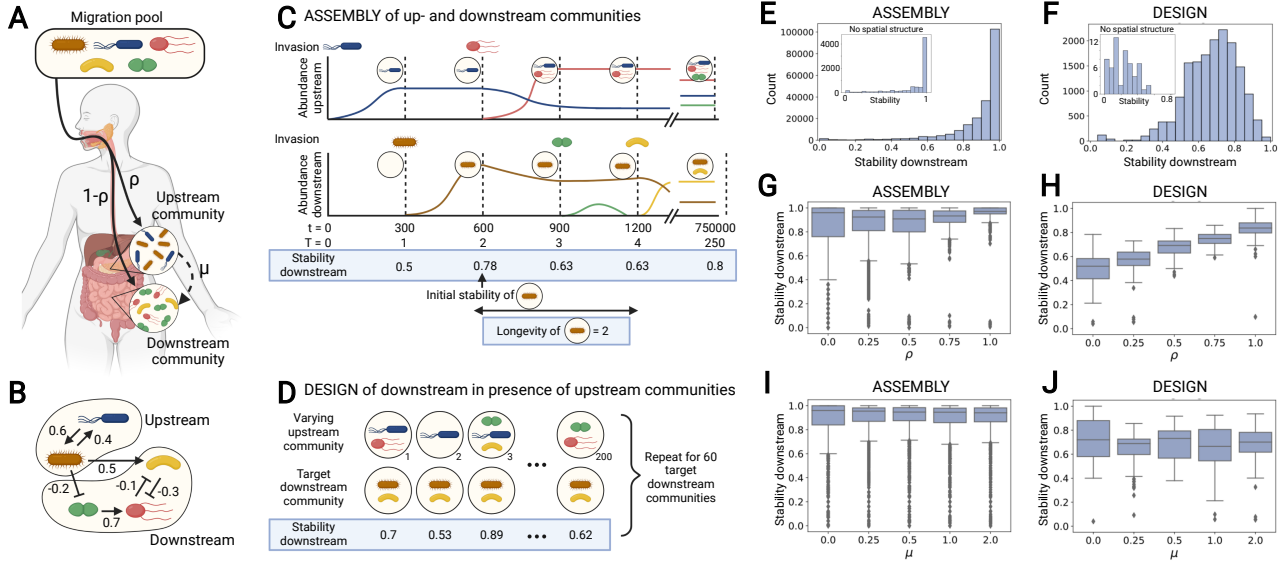
Figure 2: Adding hierarchical spatial structure under two scenarios

(A) In our hierarchical spatial structure, we consider the meta-community dynamics with two patches: upstream and downstream. Species can either migrate from the migration pool to (i) the upstream patch, (ii) or the downstream patch with likelihood ρ or $\rho - 1$, respectively, or (iii) from the upstream patch to the downstream patch with likelihood scaled by μ. (B) An example of upstream and downstream communities and their features (see also Table 1). Here, species richness is two in the upstream and three in the downstream, respectively. The upstream community has only positive interspecific interactions and thus their total and mean strength are 1.0 and 0.5, respectively. The downstream community has one positive interspecific interaction and two negative ones. The total and mean strength of positive interactions there are 0.7, while the total and mean strength of negative interactions are 0.4 and 0.2, respectively. In addition, there are a positive and a negative interaction between the two communities: total (and mean) strength of positive interaction from the upstream community to the downstream community is 0.5 and that of negative interaction is 0.2. The mean degrees within the upstream community and the downstream one are 1.0 (upstream: $(1+1)/2 = 1$, downstream: $(1+2+0)/3 = 1$). The mean degree from the upstream to the downstream is $0.667(= (1+1+0)/3)$. (C) The assembly scenario is as Fig. 1, but with an upstream community that can affect the downstream community (because of interactions as show in panel B). At each time step, one species migrates into either the up- or downstream community according to parameters ρ and μ. Stability is calculated for the downstream community only, and can change if the composition of the upstream community changes. Community longevity is calculated for a given community composition by counting the number of time steps T in which it persisted: see also Fig. S5. We assembled 60 such communities for each parameter set (ρ, μ). (D) In the design scenario, 60 target downstream communities were generated for each parameter set (ρ, μ), and for each target, we generated 200 upstream communities and analyzed the stability of the downstream community using logistic regression. (E-J) The stability of the downstream communities sampled from the assembly (E, G and I) and designing (F, H and J) scenarios. The top panels (E and F) represent the whole distributions of the stability under hierarchical spatial structure (with insets showing the distribution with no spatial structure), while the middle (G and H) and bottom panels (I and J) show the distributions given migration parameters ρ (migration from the migration pool to the upstream patch versus to the downstream) and μ (migration from the upstream patch to the downstream patch), respectively. Created with BioRender.com.

6

2.2 Migration parameters affect the stability in the structured model

Next, we consider the meta-community dynamics in the chain network composed of two patches, which can be seen as simplified gut microbiota. In the structured model, we focused on the stability of the downstream patch: the upstream stability can be analyzed as in the cases without the spatial structure because the spatial structure does not affect the dynamics in the upstream patch (see Sections 2.1, 4.2, and SI 2). The spatial structure requires us to classify three types of migration: (i) from the outside of the system (i.e., the migration pool) to the upstream patch, (ii) from the migration pool to the downstream patch, and (iii) from the upstream patch to the downstream patch. we introduce two migration parameters ρ and μ to implement these migrations. While μ represents the migration from the upstream patch to the downstream patch, ρ and $1 - \rho$ represent the migration from the migration pool to the upstream and downstream patches, respectively (Fig. 2A). If μ is larger (smaller) than 1, migration from the upstream to the downstream is more (less) likely to occur than migration from the migration pool. See Section 4.2 for more details.

Importantly, the spatial structure also affects species interactions: upstream species can affect the growth of downstream species (Fig. 2B and Eq (1b)): for example, resources produced by upstream species can flow into the downstream patch, which downstream species may consume. This type of interactions were implemented by modifying a generalized Lotka-Volterra model in the downstream patch: see Eq (1b) for mathematical representation.

To investigate the stability of downstream communities, we again used the two scenarios to sample the meta-communities. In the assembly scenario, species i invades either the upstream or the downstream with probabilities p_{i1} and p_{i2}, respectively. We sampled 60 time-series of meta-communities until $T = 250$ given a parameter set (ρ, μ). In contrast to the case without spatial structure, the stability of downstream communities depends both on upstream and downstream communities: if upstream species composition changes, the downstream stability can change (Fig. 2C) because upstream species can affect the downstream dynamics. This suggests that the downstream species composition can change even after the downstream community's stability reaches at one. However, Fig. S5 shows that the downstream communities are maintained longer when the initial stability (i.e., the downstream stability when the focal community is generated) is larger. This indicates the benefit of designing initially stable downstream communities.

To statistically analyze how we can design stable downstream communities (see also Section 2.3), we also sampled meta-communities from the designing scenario. In this scenario, we generated 60 randomly target downstream communities given a parameter set (ρ, μ), and we randomly generated 200 upstream communities for each target downstream community (fig. 2D). However, it is difficult to generate locally stable meta-communities: many randomly generated meta-communities lose species in either or both the upstream and downstream patches before the dynamics stabilize (Fig. S2). SI 3 discusses conditions for the coexistence in the meta-community.

Figs. 2E and F show the distributions of the stability sampled from the assembly ($n = 209260$) and designing scenarios ($n = 16160$), respectively. In the assembly scenarios, we removed the data where invasion does not

change any species abundances to avoid oversampling stable communities. In the designing scenario, on the other hand, we removed the data where one or more species in the generated meta-communities went extinct. Figs. 2E and F indicate that downstream communities in the assembly scenario are more stable than in the designing scenario (one-sided Wilcoxon rank-sum test: $T = 161.888$ and $p < 10^{-3}$). In the assembly scenario, the sampled stability of the downstream community is typically close to one (median: 0.949) and the distribution is biased toward the high stability. The stability also correlates with the time step (Spearman correlation coefficient:0.671 and $p < 10^{-3}$, see also Fig. S1). As these pattern also appear with no spatial structure (see Fig. 1B and the inset of Fig. 2E), the community assembly is likely to generate stable communities regardless of the spatial structure. In the designing scenario, on the other hand, the distribution of the stability is close to symmetric and its median is 0.692. Because the stability with no spatial structure (Fig. 1C or the inset of Fig. 2F) has smaller stability (median is 0.24 and the maximum stability is lower than 0.6), we can expect that the spatial structure and or the upstream species can increase the stability of downstream communities.

We then investigated how the stability changes over the two migration parameters, ρ and μ (Figs. 2G–J). These parameters change the effects of two types of the resistance on the stability: see Eq (2). For example, if $\mu = 0$ and $\rho > 0.5$, we expect that the resistance to invasion has a smaller effect than the resistance to the environmental changes because species migrate more frequently into the upstream patch than the downstream patch. In both two sampling scenarios, ρ increased the stability while μ decreased the stability (the coefficient in the logistic regression are $\rho : 0.439$ and $\mu : -0.0426$ in the assembly scenario while $\rho : 0.557$ and $\mu : -0.0585$ in the designing scenario. All p-values are smaller than 10^{-3}). As the absolute values of the coefficients of ρs are about 10 times larger than μs, the migration from the migration pool has stronger effects on the stability than the migration from the upstream to the downstream. Indeed, the causal inference suggests that increasing ρ stabilizes the downstream communities more than any analyzed communities' features and μ (Table S9).

2.3 Positive interactions increase the stability

To design stable downstream communities, we measures community features listed in Table 1 in the designing scenario. Assuming the causal diagram shown in Fig. 3, we considered the manipulations of the upstream community to increase the stability of the downstream community (but see SI 6 for the analyses on other community features). This corresponds to the cases where we have a target downstream community that is beneficial to the host and we want to stabilize the downstream community by designing an upstream community. Then, we can control the five features: species richness in the upstream, total strength of positive or negative interactions within the upstream community, and total strength of positive or negative interactions from the upstream to the downstream. Although we can also manipulate the mean degree within the upstream and the mean degree between the upstream and downstream communities, we do not consider these two features because Fig. 3 does not include causal relations from the degrees to the stability. The causal effects from the five features to the stability were estimated from different logistic regression models using the designing scenario data (see also SI 6 for more details). Table 2 summarizes the effect of each feature on the stability of
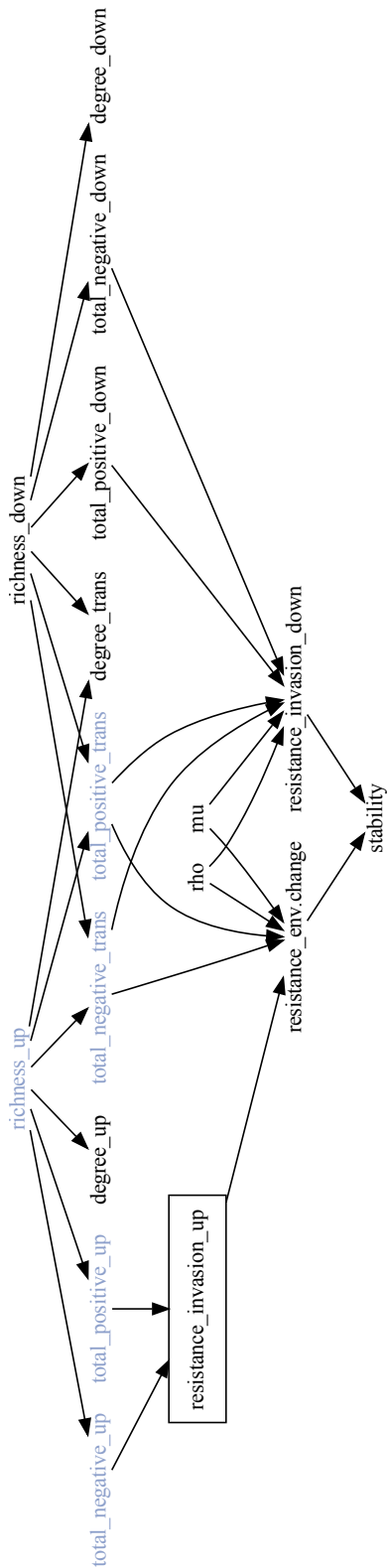
Figure 3: A full model of the causal diagram

The assumed causal diagram is represented by a directed acyclic graph. The features in blue are expected to be manipulated. The unobserved factor in this study, resistance to invasion in the upstream community, is written in a box. See the main text for the rationale of this model. The explanations of the features are summarized in Table 1.

Table 1: Summary of analyzed features

Feature or symbol	Explanation
Total_positive_up	total strength of positive interactions in the upstream
Mean_positive_up	mean strength of positive interactions in the upstream
Total_negative_up	total strength of negative interactions in the upstream
Mean_negative_up	mean strength of negative interactions in the upstream
Degree_up	mean degree in the upstream
Richness_up	species richness in the upstream
Total_positive_down	total strength of positive interactions in the downstream
Mean_positive_down	mean strength of interactions in the downstream
Total_negative_down	total strength of negative interactions in the downstream
Mean_negative_down	mean strength of negative interactions in the downstream
Degree_down	Mean degree in the downstream
Richness_down	species richness in the downstream
Total_positive_trans	total strength of positive interactions from the upstream to the downstream
Mean_positive_trans	mean strength of positive interactions from the upstream to the downstream
Total_negative_trans	total strength of negative interactions from the upstream to the downstream
Mean_negative_trans	mean strength of negative interactions from the upstream to the downstream
Degree_trans	Mean degree from the upstream to the downstream
ρ	migration from the migration pool to the upstream ($1 - \rho$: to the downstream)
μ	migration from the upstream to the downstream

Note: some features are only used for predicting the stability. See SI 5.

downstream communities: only the total strength of positive interactions from the upstream to the downstream significantly increases the stability, and the coefficient of this feature is the largest within the five community features.

We tested the effect of the positive interactions from the upstream communities to the downstream communities with the following three additional analyses. First, we compared the total strength of positive interactions from the upstream communities to the downstream communities in the assembly scenario and the designing scenario. This type of positive interactions are stronger in the assembly scenario (median: 4.764) than in the designing one (median: 1.497, Wilcoxon rank-sum test: $U = -75.621$, $p < 10^{-3}$). This can be why the downstream communities in the assembly scenario are more stable than in the designing scenario (Fig. 2E and F).

Second, we manipulated the total strength of positive interactions from the upstream to the downstream communities by adding or removing one species in the data of the assembly scenario ($n = 20308$, and 4023, respectively). Fig. 4A shows that introducing species tends to slightly increase the stability (median of changes in the logit function of the stability: 0.018, one-sided Wilcoxon signed-rank test: $T = 117072109.5$ and $p = 4.8 \times 10^{-88}$) while removing species slightly decrease the stability (median of changes in the logit function of the stability: -0.031, one-sided Wilcoxon signed-rank test: $T = 3199368$ and $p = 7.9 \times 10^{-31}$). The changes in the total strength of positive interactions from the upstream to the downstream communities positively correlate with the changes in the stability but the correlation is weak (Spearman correlation: correlation coefficient is 0.15, and $p < 3.0 \times 10^{-127}$): Fig. 4B does not show a clear pattern. This would be because adding or removing species changes community features other than the total strength of positive interactions from the upstream to the downstream communities (e.g., the total strength of positive and negative interactions within the upstream communities).

10

The third analysis overcomes this problem by generating 110 meta-communities where species compositions in the upstream and downstream do not overlap at all. In this case, we can change the total strength of positive interactions from the upstream to the downstream communities without changing the other community features. Figs. 4C and D show that the downstream communities become more stable by increasing the total strength of positive interactions from the upstream to the downstream communities. In this analysis, however, interactions from downstream species to upstream species are assumed negative. This assumption obviously increases the resistance to invasion because the upstream species receive many negative interactions when they invade the downstream patch. In other words, the increase of the stability can be because of these negative interactions, not because of the positive interactions we are interested in. To remove this bias, we also performed similar analysis while assuming species interaction effects from downstream species to upstream species are zeros (i.e., $b_{i,k}^- = 0$ for all (i, k)). In this case, we can also find that the strength of positive interactions from upstream species to downstream species tends to positively correlate with the stability. See SI 7 for more details. From these results, we can conclude that positive interactions from upstream species to downstream species stabilize the downstream communities.

Table 2: Summary of estimated causation on stability

Feature	Coefficient	Standard error	P-value
Richness_up	0.0057	0.016	0.726
Total_positive_up	0.0194	0.029	0.500
Total_negative_up	0.0319	0.029	0.265
Total_positive_trans	0.0487	0.022	0.029
Total_negative_trans	0.0193	0.022	0.387

3 Discussion

In this manuscript, we built a mathematical model of the meta-community dynamics which simplified gut microbiota and analyzed the stability of the downstream communities. Because many previous studies suggest the relationship between gut microbiota and hosts' health, current research focuses on designing beneficial gut microbiota (Pham et al., 2017; Swann et al., 2020; Clark et al., 2021). However, we need to care about the stability of the designed microbiota so that the beneficial effects from the microbes are maintained for a long time (Fig. S5). Our stability is composed of the resistance to the invasion in the downstream community and the resistance to the environmental changes caused by the upstream community. With the motivation of stabilizing the gut microbiota, we investigated the stability of the downstream community where the spatial structure has an effect. For the stability of the upstream community, we could analyze the stability in the absence of spatial structure (SI 2). In such cases, the species richness increases the stability.

First, we investigated the two sampling scenarios: assembly and designing. The assembly scenario simulates the natural process that a gut microbiota can be assembled. The designing scenario, on the other hand, randomly generated the meta-communities so that we can easily perform the statistical analysis. Figs. 2E and F show that the downstream communities in the assembly scenario are more stable than in the designing
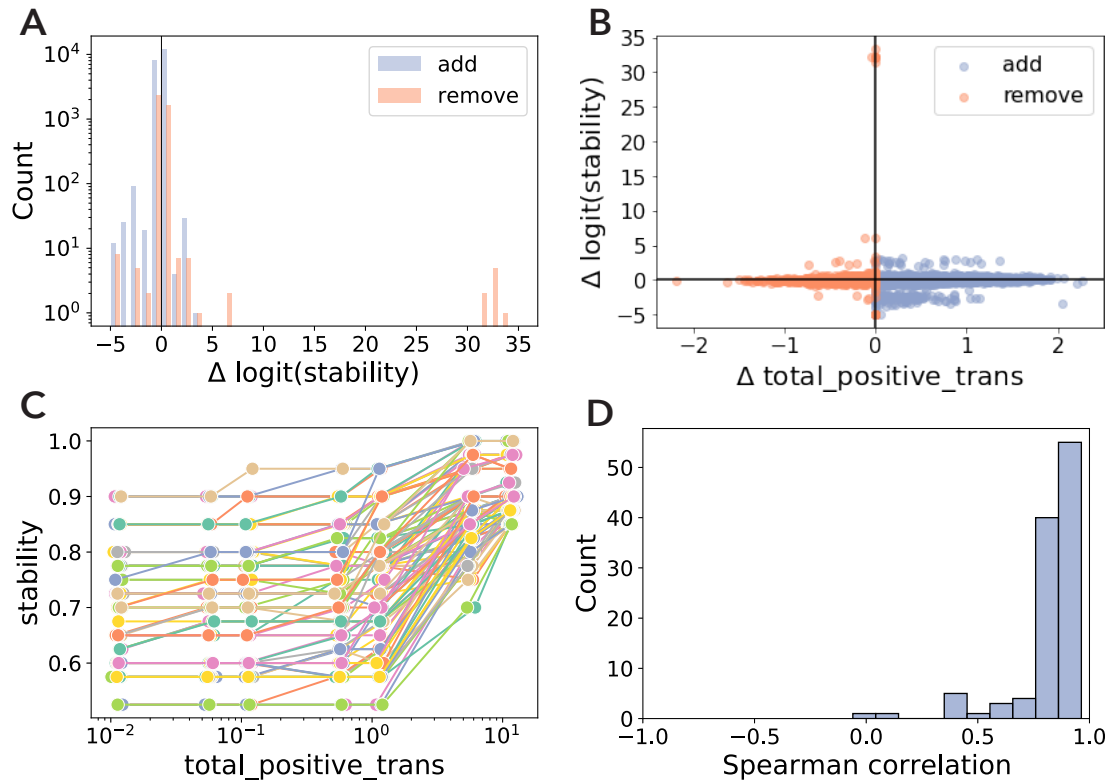
Figure 4: Manipulating positive interactions from the upstream to the downstream

We manipulated the total strength of positive interactions from the upstream to the downstream communities by adding/removing one species to/from the upstream communities in the assembly scenario (A and B), or by changing only the interactions between the communities (C and D). A: Introducing species tends to increases the stability ($n = 20308$, median of changes in the logit function of the stability: 0.018, one-sided Wilcoxon signed-rank test: $U = 117072109.5$ and $p = 4.8 \times 10^{-88}$) while removing species decrease the stability ($n = 4023$, median of changes in the logit function of the stability: -0.031, one-sided Wilcoxon signed-rank test: $U = 3199368$ and $p = 7.9 \times 10^{-31}$)). B: By adding or removing species, we can also manipulate the total strength of positive interactions from the upstream and the downstream communities, which positively correlate with the changes in the stability (Spearman correlation: correlation=0.15, and $p < 3.0 \times 10^{-127}$). However, this correlation is very weak, and increasing/decreasing the positive interactions does not necessarily increase/decrease the stability, respectively. C: We manipulated only the total strength of positive interactions from the upstream to the downstream communities while remaining other community features. Each line represents a different meta-community ($n = 110$). We calculated Spearman correlation coefficient between the stability and the total strength of positive interactions from the upstream community to the downstream community in each meta-community. D: We summarized the Spearman correlations calculated in panel C in the histogram. Except for one meta-community, the downstream stability positively correlates with the total strength of positive interactions and the correlation coefficients are large (> 0.8) in many cases.

scenario. Especially, the downstream communities in the assembly scenario are very stable (median stability: 0.949), which is consistent with the previous studies that show temporal stability of gut microbiota within hosts (Costello et al., 2009; Byrd et al., 2020). In addition, Fig. S2 and SI 3 suggest the difficulty in designing an upstream community that does not collapse the target downstream communities. From these results, we suggest that imitating meta-communities of healthy hosts would be more reasonable than artificially generating gut-microbiota for designing gut-microbiota.

Second, we analyzed the effects of the two migration parameters, ρ (migration from the outside the meta-community to the upstream community comparing the migration to the downstream community) and μ (migration from the upstream to the downstream). These parameters relate the downstream stability because they determine whether migration to the upstream (and then affecting resistance to the environmental changes) or to the downstream (and affecting resistance to invasion) more frequently occurs. Figs. 2G – J show that ρ has a stronger effect on the stability than μ in both assembly and designing scenarios. Indeed, the causal inference suggests that ρ has the strongest effect on the downstream stability in the measured features (Table S9). Intuitively, large ρ indicates that species migration to the downstream is unlikely and thus the downstream species composition is unlikely to change due to the migration. In addition, when an invader colonizes the upstream patch, this does not necessarily change the downstream species composition. As ρ does not affect the coexistence of the species in the upstream and downstream communities, increasing ρ would increase the stability without collapsing the meta-community in the gut. However, we do not know how easy it is to modify species migration in nature because the migration parameters would depend on both hosts and microbes.

Instead, we can consider the manipulation of the upstream species compositions to stabilize target downstream communities. The causal inference (Table 2) suggests that positive effects from the upstream to the downstream increase the downstream stability. Intuitively, increasing such positive effects prevent the extinction of the downstream species when an invader species arrives either in the upstream or the downstream communities. To increase the positive effects from the upstream to the downstream, one may increase the species richness in the upstream. However, we need to keep in mind that (i) increasing upstream species richness may collapse the coexistence in the upstream patch (May, 1972), (ii) introduction of species into the upstream patch may drive one or more species extinction in the downstream patch, (iii) species richness alone cannot increase the stability (Table 2), and (iv) other community features can change while changing species richness (Figs. 4A and B). To stabilize the downstream community, therefore, we carefully choose which species to introduce into the upstream community: we need to choose the species that (a) do not drive other species extinction either in upstream and downstream communities, and (b) have positive effects on the downstream species while affecting little on other community features (Figs. 4C and D, and Figs. S9A and B). As discussed in SI 3, deriving conditions to satisfy (a) in general is difficult. If we could increase the positive effects from the upstream to the downstream by increasing the upstream species richness, we can also expect that the upstream stability (i.e., the resistance to invasion in the upstream patch) increases as suggested by previous studies without the spatial structures (Stachowicz et al., 1999; Bonanomi et al., 2014; Hromada et al., 2021). See also SI 2.

In this study, we considered the abstract spatial structure composed of the upstream and downstream

patches. In reality, we need to define where we put these patches, which would affect the parameter values related to species migration. For example, we may define the two patches as an upstream and a downstream of a small intestine. Then, we could expect that species migration from the upstream to the downstream is frequent (large μ) because species migration would be easy when the distance between the patches is small. In this case, we can also assume that the frequencies of species migration from the migration pool to the upstream are similar to those from the migration pool to the downstream community (i.e., ρ is around 0.5). From the results of this manuscript, we argue that the downstream stability in the small intestine increases when the upstream community in the small intestine has a strong positive effect on the downstream community. Of course, we may be able to define the upstream community as an oral microbiome and the downstream community as a gut microbiome in the small intestine, assuming large ρ and small μ. However, species interactions between the patches could differ from those within the patches, because there would be many intermediate communities between the two patches which mediate the species interactions from the upstream to the downstream. In this case, our assumption that species interactions are identical within and between patches is violated.

One limitation of our study is that the results are based on the gLV model. Although gLV models are widely used in ecology due to their simplicity, Momeni et al. (2017) show that these models fail in capturing diverse microbial pairwise interactions. This is because many microbial interactions are mediated by chemical compounds such as resources and toxins while gLV models represent phenomenological species interactions. Due to these chemically mediated interactions, the strength and the sign of microbial species interactions can change depending on the environmental conditions (Hoek et al., 2016; Estrela et al., 2019; Piccardi et al., 2019; Zuñiga et al., 2019). One way to overcome this problem is building a consumer-resource model, where the dynamics of species, resources, and toxins are explicitly included. According to the analysis of this manuscript, we expect that introducing the species in the upstream community that secrete resources or degrade/deactivate toxins for the downstream species would increase the downstream stability.

In summary, we investigated the stability of simplified gut microbiota using the meta-community dynamics. We propose three ideas to stabilize the downstream communities. First, imitating gut microbiota of healthy hosts is reasonable because naturally assembled communities are more stable than randomly generated ones. Second, if possible, we suggest the modification of migration from outside the meta-community so that species more frequently migrate into the upstream patch than the downstream patch. This increases the downstream stability most in the analyzed features. However, such modification may be difficult. Alternatively, we can stabilize the focal downstream community by designing an upstream community so that the upstream species have strong positive effects on the downstream species.

Acknowledgment

Author contributions

Conceptualization, S.S. and S.M. Methodology, S.S.; Formal Analysis, S.S.; Writing – Original Draft, S.S.; Writing – Review and Editing, S.S. and S.M.; Visualization, S.S. and S.M.; Supervision, S.M.; Funding Acquisition, S.S.and S.M.

Declaration of interests

The authors declare no competing interests.

References

Arif, S. and MacNeil, A. Predictive models aren't for causal inference. Ecology Letters, 6 2022. ISSN 1461-023X. doi: 10.1111/ELE.14033. URL https://onlinelibrary.wiley.com/doi/10.1111/ele.14033.

Bonanomi, G., Capodilupo, M., Incerti, G., Gaglione, S. A., and Scala, F. Fungal diversity increases soil fungistasis and resistance to microbial invasion by a non resident species. Biological Control, 72:38–45, 2014. ISSN 10499644. doi: 10.1016/j.biocontrol.2014.02.005. URL http://dx.doi.org/10.1016/j.biocontrol.2014.02.005.

Butler, S. and ODwyer, J. P. Stability criteria for complex microbial communities. Nature Communications, 9(1):2970, 12 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05308-z. URL http://www.nature.com/articles/s41467-018-05308-z.

Byrd, A. L., Liu, M., Fujimura, K. E., Lyalina, S., Nagarkar, D. R., Charbit, B., Bergstedt, J., Patin, E., Harrison, O. J., Quintana-Murci, L., Mellman, I., Duffy, D., and Albert, M. L. Gut microbiome stability and dynamics in healthy donors and patients with non-gastrointestinal cancers. Journal of Experimental Medicine, 218(1), 2020. ISSN 15409538. doi: 10.1084/JEM.20200606.

Case, T. J. Invasion resistance arises in strongly interacting species-rich model competition communities. Proceedings of the National Academy of Sciences, 87(24):9610–9614, 12 1990. ISSN 0027-8424. doi: 10.1073/pnas.87.24.9610. URL https://pnas.org/doi/full/10.1073/pnas.87.24.9610.

Cenci, S. and Saavedra, S. Structural stability of nonlinear population dynamics. Physical Review E, 97(1): 012401, 1 2018. ISSN 2470-0045. doi: 10.1103/PhysRevE.97.012401. URL https://link.aps.org/doi/10.1103/PhysRevE.97.012401.

Chesson, P. Mechanisms of Maintenance of Species Diversity. Annual Review of Ecology and Systematics, 31(1):343–366, 11 2000. ISSN 0066-4162. doi: 10.1146/annurev.ecolsys.31.1.343. URL http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.31.1.343.

Chisholm, C., Lindo, Z., and Gonzalez, A. Metacommunity diversity depends on connectivity and patch arrangement in heterogeneous habitat networks. Ecography, 34(3):415–424, 6 2011. ISSN 09067590. doi: 10.1111/j.1600-0587.2010.06588.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1600-0587.2010.06588.x.

Cinquin, C., Le Blay, G., Fliss, I., and Lacroix, C. New three-stage in vitro model for infant colonic fermentation with immobilized fecal microbiota. FEMS Microbiology Ecology, 57(2):324–336, 8 2006. ISSN 01686496. doi: 10.1111/j.1574-6941.2006.00117.x. URL https://academic.oup.com/femsec/article-lookup/doi/10.1111/j.1574-6941.2006.00117.x.

Clark, R. L., Connors, B. M., Stevenson, D. M., Hromada, S. E., Hamilton, J. J., Amador-Noguez, D., and Venturelli, O. S. Design of synthetic human gut microbiome assembly and butyrate production. Nature

Communications, 12(1):3254, 12 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-22938-y. URL http://www.nature.com/articles/s41467-021-22938-y.

Costello, E. K., Lauber, C. L., Hamady, M., Fierer, N., Gordon, J. I., and Knight, R. Bacterial Community Variation in Human Body Habitats Across Space and Time. Science, 326(5960):1694–1697, 12 2009. ISSN 0036-8075. doi: 10.1126/science.1177486. URL https://www.science.org/doi/10.1126/science.1177486.

Dethlefsen, L. and Relman, D. A. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. Proceedings of the National Academy of Sciences of the United States of America, 108(SUPPL. 1):4554–4561, 2011. ISSN 10916490. doi: 10.1073/pnas.1000087107.

Economo, E. P. and Keitt, T. H. Species diversity in neutral metacommunities: a network approach. Ecology Letters, 11(1):071117033013001–???, 11 2007. ISSN 1461-023X. doi: 10.1111/j.1461-0248.2007.01126.x. URL http://doi.wiley.com/10.1111/j.1461-0248.2007.01126.x.

Estrela, S., Libby, E., Van Cleve, J., Débarre, F., Deforet, M., Harcombe, W. R., Peña, J., Brown, S. P., and Hochberg, M. E. Environmentally Mediated Social Dilemmas. Trends in Ecology & Evolution, 34(1):6–18, 1 2019. ISSN 01695347. doi: 10.1016/j.tree.2018.10.004. URL https://linkinghub.elsevier.com/retrieve/pii/S0169534718302490.

Fan, Y. and Pedersen, O. Gut microbiota in human metabolic health and disease. Nature Reviews Microbiology, 19(1):55–71, 2021. ISSN 17401534. doi: 10.1038/s41579-020-0433-9. URL http://dx.doi.org/10.1038/s41579-020-0433-9.

Fassarella, M., Blaak, E. E., Penders, J., Nauta, A., Smidt, H., and Zoetendal, E. G. Gut microbiome stability and resilience: Elucidating the response to perturbations in order to modulate gut health. Gut, 70(3):595–605, 2021. ISSN 14683288. doi: 10.1136/gutjnl-2020-321747.

Godwin, C. M., Chang, F., and Cardinale, B. J. An empiricist's guide to modern coexistence theory for competitive communities. Oikos, 129(8):1109–1127, 8 2020. ISSN 0030-1299. doi: 10.1111/oik.06957. URL https://onlinelibrary.wiley.com/doi/10.1111/oik.06957.

Gravel, D., Guichard, F., Loreau, M., and Mouquet, N. Source and sink dynamics in meta-ecosystems. Ecology, 91(7):2172–2184, 2010. ISSN 00129658. doi: 10.1890/09-0843.1.

Guichard, F. Recent advances in metacommunities and meta-ecosystem theories [version 1; peer review: 2 approved]. F1000Research, 8(May 2017):1–8, 2019. ISSN 1759796X.

Hanski, I. and Gyllenberg, M. Two General Metapopulation Models and the Core-Satellite Species Hypothesis. The American Naturalist, 142(1):17–41, 1993. URL https://www.jstor.org/stable/2462632.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W.,

Abbasi, H., Gohlke, C., and Oliphant, T. E. Array programming with NumPy. Nature, 585(7825):357–362, September 2020. doi: 10.1038/s41586-020-2649-2. URL https://doi.org/10.1038/s41586-020-2649-2.

Hernán, M. and Robins, J. CausalInference: What If. Chapman & Hall/CRC., 2020. URL https://cdn1.sph.harvard.edu/wp-content/uploads/sites/1268/2021/03/ciwhatif_hernanrobins_30mar21.pdf.

Hoek, T. A., Axelrod, K., Biancalani, T., Yurtsev, E. A., Liu, J., and Gore, J. Resource Availability Modulates the Cooperative and Competitive Nature of a Microbial Cross-Feeding Mutualism. PLOS Biology, 14(8): e1002540, 8 2016. ISSN 1545-7885. doi: 10.1371/journal.pbio.1002540. URL http://dx.plos.org/10.1371/journal.pbio.1002540.

Holenstein, K., Harvey, E., and Altermatt, F. Patch size distribution affects species invasion dynamics in dendritic networks. Oikos, 2022(1):1–11, 1 2022. ISSN 0030-1299. doi: 10.1111/oik.08679. URL https://onlinelibrary.wiley.com/doi/10.1111/oik.08679.

Hromada, S., Clark, R. L., Qian, Y., Watson, L., Safdar, N., and Venturelli, O. S. Species richness determines C. difficile invasion outcome in synthetic human gut communities. bioRxiv, page 2021.03.23.436677, 2021. doi: 10.1101/2021.03.23.436677. URL https://www.biorxiv.org/content/10.1101/2021.03.23.436677v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.03.23.436677v1.abstract.

James, G., Witten, D., Hastie, T., and Tibshirani, R. An Introduction to Statistical Learning - with Applications in R. Springer, 2013. ISBN 9781461471370. URL https://www.springer.com/gp/book/9781461471370%0Ahttp://www.springer.com/us/book/9781461471370.

Landi, P., Minoarivelo, H. O., Brännström, ., Hui, C., and Dieckmann, U. Complexity and stability of ecological networks: a review of the theory. Population Ecology, 60(4):319–345, 10 2018. ISSN 1438-3896. doi: 10.1007/s10144-018-0628-3. URL https://onlinelibrary.wiley.com/doi/10.1007/s10144-018-0628-3.

Lee, W. J. and Hase, K. Gut microbiota-generated metabolites in animal health and disease. Nature Chemical Biology, 10(6):416–424, 2014. ISSN 15524469. doi: 10.1038/nchembio.1535.

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Holt, R. D., Shurin, J. B., Law, R., Tilman, D., Loreau, M., and Gonzalez, A. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters, 7(7):601–613, 6 2004. ISSN 1461023X. doi: 10.1111/j.1461-0248.2004.00608.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2004.00608.x.

Ley, R. E., Turnbaugh, P. J., Klein, S., and Gordon, J. I. Human gut microbes associated with obesity. Nature, 444(7122):1022–1023, 12 2006. ISSN 0028-0836. doi: 10.1038/4441022a. URL http://www.nature.com/articles/4441022a.

Loreau, M., Mouquet, N., and Holt, R. D. Meta-ecosystems: A theoretical framework for a spatial ecosystem ecology. Ecology Letters, 6(8):673–679, 2003. ISSN 1461023X. doi: 10.1046/j.1461-0248.2003.00483.x.

Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., and Knight, R. Diversity, stability and resilience of the human gut microbiota. Nature, 489(7415):220–230, 2012. ISSN 00280836. doi: 10.1038/nature11550.

Mallon, C. A., Van Elsas, J. D., and Salles, J. F. Microbial invasions: The process, patterns, and mechanisms. Trends in Microbiology, 23(11):719–729, 2015. ISSN 18784380. doi: 10.1016/j.tim.2015.07.013.

Marchesi, J. R., Adams, D. H., Fava, F., Hermes, G. D., Hirschfield, G. M., Hold, G., Quraishi, M. N., Kinross, J., Smidt, H., Tuohy, K. M., Thomas, L. V., Zoetendal, E. G., and Hart, A. The gut microbiota and host health: A new clinical frontier. Gut, 65(2):330–339, 2016. ISSN 14683288. doi: 10.1136/gutjnl-2015-309990.

Massol, F., Gravel, D., Mouquet, N., Cadotte, M. W., Fukami, T., and Leibold, M. A. Linking community and ecosystem dynamics through spatial ecology. Ecology Letters, 14(3):313–323, 2011. ISSN 1461023X. doi: 10.1111/j.1461-0248.2011.01588.x.

May, R. M. Will a Large Complex System be Stable? Nature, 238(5364):413–414, 8 1972. ISSN 0028-0836. doi: 10.1038/238413a0. URL https://www.nature.com/articles/238413a0.

Momeni, B., Xie, L., and Shou, W. Lotka-Volterra pairwise modeling fails to capture diverse pairwise microbial interactions. eLife, 6:1–34, 2017. ISSN 2050084X. doi: https://doi.org/10.7554/eLife.25051.001. URL https://elifesciences.org/articles/25051#content.

OSullivan, J. D., Knell, R. J., and Rossberg, A. G. Metacommunityscale biodiversity regulation and the selforganised emergence of macroecological patterns. Ecology Letters, 22(9):1428–1438, 9 2019. ISSN 1461-023X. doi: 10.1111/ele.13294. URL https://onlinelibrary.wiley.com/doi/10.1111/ele.13294.

Payne, A. N., Zihler, A., Chassard, C., and Lacroix, C. Advances and perspectives in in vitro human gut fermentation modeling. Trends in Biotechnology, 30(1):17–25, 2012. ISSN 01677799. doi: 10.1016/j.tibtech.2011.06.011. URL http://dx.doi.org/10.1016/j.tibtech.2011.06.011.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12:2825–2830, 2011.

Pham, H. L., Ho, C. L., Wong, A., Lee, Y. S., and Chang, M. W. Applying the design-build-test paradigm in microbiome engineering. Current Opinion in Biotechnology, 48:85–93, 2017. ISSN 18790429. doi: 10.1016/j.copbio.2017.03.021. URL http://dx.doi.org/10.1016/j.copbio.2017.03.021.

Piccardi, P., Vessman, B., and Mitri, S. Toxicity drives facilitation between 4 bacterial species. Proceedings of the National Academy of Sciences, 116(32):15979–15984, 8 2019. ISSN 0027-8424. doi: 10.1073/pnas.1906172116. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1906172116.

Portillo, José, R., Soler-Toscano, F., and Analysis, N. Cooperation enhances structural stability in mutualistic systems. bioRxiv, pages 1–26, 2021. URL http://biorxiv.org/cgi/content/short/2021.04.

23.441179v1?rss=1&utm_source=researcher_app&utm_medium=referral&utm_campaign=RESR_MRKT_ Researcher_inbound.

Post, W. and Pimm, S. Community assembly and food web stability. Mathematical Biosciences, 64(2):169–192, 6 1983. ISSN 00255564. doi: 10.1016/0025-5564(83)90002-0. URL https://linkinghub.elsevier.com/ retrieve/pii/0025556483900020.

Rohr, R. P., Saavedra, S., and Bascompte, J. On the structural stability of mutualistic systems. Science, 345(6195):1253497–1253497, 7 2014. ISSN 0036-8075. doi: 10.1126/science.1253497. URL http://www. sciencemag.org/cgi/doi/10.1126/science.1253497.

Saavedra, S., Rohr, R. P., Bascompte, J., Godoy, O., Kraft, N. J. B., and Levine, J. M. A structural approach for understanding multispecies coexistence. Ecological Monographs, 87(3):470–486, 8 2017. ISSN 00129615. doi: 10.1002/ecm.1263. URL http://doi.wiley.com/10.1002/ecm.1263.

Sánchez, B., Delgado, S., Blanco-Míguez, A., Lourenço, A., Gueimonde, M., and Margolles, A. Probiotics, gut microbiota, and their influence on host health and disease. Molecular Nutrition & Food Research, 61(1): 1600240, 1 2017. ISSN 16134125. doi: 10.1002/mnfr.201600240. URL https://onlinelibrary.wiley.com/ doi/10.1002/mnfr.201600240.

Sanna, S., van Zuydam, N. R., Mahajan, A., Kurilshikov, A., Vich Vila, A., Vsa, U., Mujagic, Z., Masclee, A. A., Jonkers, D. M., Oosting, M., Joosten, L. A., Netea, M. G., Franke, L., Zhernakova, A., Fu, J., Wijmenga, C., and McCarthy, M. I. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. Nature Genetics, 51(4):600–605, 2019. ISSN 15461718. doi: 10.1038/s41588-019-0350-x. URL http://dx.doi.org/10.1038/s41588-019-0350-x.

Seabold, S. and Perktold, J. statsmodels: Econometric and statistical modeling with python. In 9th Python in Science Conference, 2010.

Selber-Hnatiw, S., Rukundo, B., Ahmadi, M., Akoubi, H., Al-Bizri, H., Aliu, A. F., Ambeaghen, T. U., Avetisyan, L., Bahar, I., Baird, A., Begum, F., Ben Soussan, H., Blondeau-Éthier, V., Bordaries, R., Bramwell, H., Briggs, A., Bui, R., Carnevale, M., Chancharoen, M., Chevassus, T., Choi, J. H., Coulombe, K., Couvrette, F., D'Abreau, S., Davies, M., Desbiens, M.-P., Di Maulo, T., Di Paolo, S.-A., Do Ponte, S., dos Santos Ribeiro, P., Dubuc-Kanary, L.-A., Duncan, P. K., Dupuis, F., El-Nounou, S., Eyangos, C. N., Ferguson, N. K., Flores-Chinchilla, N. R., Fotakis, T., Gado Oumarou H D, M., Georgiev, M., Ghiassy, S., Glibetic, N., Grégoire Bouchard, J., Hassan, T., Huseen, I., Ibuna Quilatan, M.-F., Iozzo, T., Islam, S., Jaunky, D. B., Jeyasegaram, A., Johnston, M.-A., Kahler, M. R., Kaler, K., Kamani, C., Karimian Rad, H., Konidis, E., Konieczny, F., Kurianowicz, S., Lamothe, P., Legros, K., Leroux, S., Li, J., Lozano Ro- driguez, M. E., Luponio-Yoffe, S., Maalouf, Y., Mantha, J., McCormick, M., Mondragon, P., Narayana, T., Neretin, E., Nguyen, T. T. T., Niu, I., Nkemazem, R. B., O'Donovan, M., Oueis, M., Paquette, S.,

Patel, N., Pecsi, E., Peters, J., Pettorelli, A., Poirier, C., Pompa, V. R., Rajen, H., Ralph, R.-O., Rosales-Vasquez, J., Rubinshtein, D., Sakr, S., Sebai, M. S., Serravalle, L., Sidibe, F., Sinnathurai, A., Soho, D., Sundarakrishnan, A., Svistkova, V., Ugbeye, T. E., Vasconcelos, M. S., Vincelli, M., Voitovich, O., Vrabel, P., Wang, L., Wasfi, M., Zha, C. Y., and Gamberi, C. Human Gut Microbiota: Toward an Ecology of Disease. Frontiers in Microbiology, 8(JUL), 7 2017. ISSN 1664-302X. doi: 10.3389/fmicb.2017.01265. URL http://journal.frontiersin.org/article/10.3389/fmicb.2017.01265/full.

Seymour, M., Fronhofer, E. A., and Altermatt, F. Dendritic network structure and dispersal affect temporal dynamics of diversity and species persistence. Oikos, 124(7):908–916, 2015. ISSN 16000706. doi: 10.1111/oik.02354.

Song, C. and Saavedra, S. Structural stability as a consistent predictor of phenological events. Proceedings of the Royal Society B: Biological Sciences, 285(1880):20180767, 6 2018. ISSN 0962-8452. doi: 10.1098/rspb.2018.0767. URL https://royalsocietypublishing.org/doi/10.1098/rspb.2018.0767.

Stachowicz, J. J., Whitlatch, R. B., and Osman, R. W. Species Diversity and Invasion Resistance in a Marine Ecosystem. Science, 286(5444):1577–1579, 11 1999. ISSN 0036-8075. doi: 10.1126/science.286.5444.1577. URL https://www.science.org/doi/10.1126/science.286.5444.1577.

Swann, J. R., Rajilic-Stojanovic, M., Salonen, A., Sakwinska, O., Gill, C., Meynier, A., Fança-Berthon, P., Schelkle, B., Segata, N., Shortt, C., Tuohy, K., and Hasselwander, O. Considerations for the design and conduct of human gut microbiota intervention studies relating to foods. European Journal of Nutrition, 59 (8):3347–3368, 2020. ISSN 14366215. doi: 10.1007/s00394-020-02232-1. URL https://doi.org/10.1007/s00394-020-02232-1.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

Vrieze, A., Van Nood, E., Holleman, F., Salojärvi, J., Kootte, R. S., Bartelsman, J. F., Dallinga-Thie, G. M., Ackermans, M. T., Serlie, M. J., Oozeer, R., Derrien, M., Druesne, A., Van Hylckama Vlieg, J. E., Bloks, V. W., Groen, A. K., Heilig, H. G., Zoetendal, E. G., Stroes, E. S., De Vos, W. M., Hoekstra, J. B., and Nieuwdorp, M. Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. Gastroenterology, 143(4):913–916, 2012. ISSN 15280012. doi: 10.1053/j.gastro.2012.06.031. URL http://dx.doi.org/10.1053/j.gastro.2012.06.031.

Xu, X., Xu, P., Ma, C., Tang, J., and Zhang, X. Gut microbiota, host health, and polysaccharides. Biotechnology

Advances, 31(2):318–337, 2013. ISSN 07349750. doi: 10.1016/j.biotechadv.2012.12.009. URL http://dx.doi.org/10.1016/j.biotechadv.2012.12.009.

Zuñiga, C., Li, C.-T., Yu, G., Al-Bassam, M. M., Li, T., Jiang, L., Zaramela, L. S., Guarnieri, M., Betenbaugh, M. J., and Zengler, K. Environmental stimuli drive a transition from cooperation to competition in synthetic phototrophic communities. Nature Microbiology, 4(12):2184–2191, 12 2019. ISSN 2058-5276. doi: 10.1038/s41564-019-0567-6. URL http://www.nature.com/articles/s41564-019-0567-6.

4 STAR Methods

4.1 Resource availability

4.1.1 Lead contact

Further information should be directed to and will be fulfilled by the lead contact, Sara Mitri (sara.mitri@unil.ch).

4.1.2 Materials availability

This study did not generate any new materials.

4.1.3 Data and code avaiolability

All original codes with Jupyter notebook and csv files, have been deposited at Zenod and is publicly available as of the date of submission. The DOI appears in the key resources table.

4.2 Method details

Dynamics within each patch

In this manuscript, we consider meta-community dynamics in the simplest chain network composed of two patches: an upstream and a downstream patch. While species interact within each patch, species in the upstream patch can also affect the dynamics in the downstream community because chemical compounds flow from the upstream patch into the downstream patch. For example, if an upstream species consumes a resource that a downstream species needs, the upstream species will have a negative effect on the growth of that species in the downstream patch. In contrast, if species in the upstream community produce beneficial byproducts, they have positive effects on the species' growth in the downstream patch. Without migration of species, we can represent the dynamics of species i's abundance $(i = 1, \ldots, N)$ in patch j (where $j = 1, 2$ represents upstream and downstream patch, respectively) using a generalized Lotka-Volterra (gLV) model:

$$\frac{dx_{i1}}{dt} = x_{i1} \left(r_{i1} + \sum_{k=1}^{N} a_{ik} x_{k1} \right) \tag{1a}$$

$$\frac{dx_{i2}}{dt} = x_{i2} \left\{ r_{i2} + \sum_{k=1}^{N} a_{ik} (x_{k1} + x_{k2}) \right\}, \tag{1b}$$

where x_{ij} represents species i's abundance in patch j, r_{ij} is the intrinsic growth rate of species i in patch j, and a_{ik} represents the net interaction from species k to i. Here, we assume that species' intrinsic growth rates can differ between the two patches due to environmental conditions that species cannot change (e.g., temperature), while species interactions remain constant between patches. Although gLV models represent species interactions phenomenologically, we assume that all species interact with one another in indirect ways: secreting or absorbing chemical compounds which are not explicitly written in this model. For simplicity, we scaled the interaction parameters so that $a_{ii} = -1$ for all i so that the abundance of species i in patch j is $x_{ij}^* = r_{ij}$ at equilibrium in mono-culture. See SI 1 for more details and how we could fit this model to empirical data. For simulating cases without the spatial structure, we used Eq (1b) while $x_{k1} = 0$ for all k.

Defining stability

This manuscript investigates the effect of hierarchical spatial structure on the stability of downstream communities. By assuming rare species migration, our model has two time scales: the short time scale t, when we consider the growth of species in the ordinary differential equations Eqs (1a) and (1b), and the long time scale T when we consider species invasion and stability of the communities. We ran the simulations given by Eqs (1a) and (1b) for a long time ($t = 300$) using solve_ivp function with LSODA in scipy packages (Virtanen et al., 2020), after which a new species migrates into either upstream and downstream communities.

We define the stability at T as the probability that downstream species composition (i.e., presence/absence of each species) does not change after one species invades either the upstream or downstream patch (Fig. 1A). Without loss of generality, we can denote $q_0(T+1)$ as the probability that the downstream species composition is identical to that at T. Then,

$$\text{Stability}(T) \equiv q_0(T+1)$$
$$= 1 - \sum_{i \neq 0} q_i(T+1), \tag{2}$$

where $q_i(T+1), i \neq 1$ represents the probability that the downstream species composition changes to state i at time step $T+1$.

Our stability can be divided into two criteria: resistance to invasion and resistance to environmental changes. Resistance to invasion is the probability that an invader species fails to colonize the downstream community ad do not exclude any resident species. We use $\text{Ri}(i, T)$ to represents whether downstream species composition at time step T is maintained or not when species i invades the downstream community:

$$\text{Ri}(i, T) = \begin{cases} 1 & \text{(maintained)} \\ 0 & \text{(otherwise)}. \end{cases} \tag{3}$$

Resistance to environmental change represents the maintenance of the downstream community when a species invades the upstream community, changing the environment, i.e. the growth rates of downstream species. Such

upstream changes can thereby cause the extinction of one or more downstream species. This effect can be clarified by rewriting Eq (1b) as follows:

$$\frac{dx_{i2}}{dt} = x_{i2}\left\{\hat{r}_{i2}\left(\vec{x}_1\right) + \sum_{k=1}^{N} a_{ik}x_{k2}\right\},$$ (4)

where

$$\hat{r}_{i2}\left(\vec{x}_1\right) \equiv r_{i2} + \sum_{k=1}^{N} a_{ik}x_{k1}$$ (5)

and $\vec{x}_1 = (x_{11}, x_{21}, \ldots, x_{N1})$. In other words, the realized intrinsic growth rate of species i in the downstream patch $\hat{r}_{i2}$ is decomposed into its baseline r_{i2} (given by the growth rate when no species exist in the upstream patch) and the effects from the upstream community $\sum_{k=1}^{N} a_{ik}x_{k1}$. As the downstream species cannot affect the upstream dynamics, changes in the upstream community can be seen as environmental changes from the perspective of the downstream communities. We use $\mathrm{Re}(i,T)$ to represent whether the invasion of species i to the upstream drive one or more downstream species at time step T extinction or not:

$$\mathrm{Re}(i,T) = \begin{cases} 0 & \text{(extinction)} \\ 1 & \text{(otherwise)} \end{cases}$$ (6)

We emphasize that the resistance to environmental changes differs from structural stability, although both can consider the perturbation in $\vec{r}_2$. Structural stability considers the sudden jump of $\vec{r}_2 \to \vec{r}_2 + \vec{\Delta}$ where $\vec{\Delta}$ represents the changes in the growth rates. On the other hand, resistance to the environmental changes indicates the gradual changes of $\vec{r}_2$ because such changes are caused by the dynamics of the upstream community.

Species migration

To calculate stability, we need to know the probability that each species migrates into each patch. We consider three types of migration: (i) from outside the meta-community (hereafter, called the migration pool) to the upstream patch, (ii) from the migration pool to the downstream patch, and (iii) from upstream to downstream patch (Fig. 2A). We emphasize that an invader should exist in the upstream patch in migration type (iii) while the existence of this species in the upstream patch is not necessary in migration type (ii). In all migration types, the abundances of invader species are fixed at 0.01.

The probabilities of species migration in each patch can be written as follows:

$$p_{i1} = \frac{\rho}{N + \mu \sum_{i=1}^{N} \delta\left(x_{i1}\right)}$$ (7a)

$$p_{i2} = \frac{(1-\rho) + \mu\delta(x_{i1})}{N + \mu \sum_{i=1}^{N} \delta\left(x_{i1}\right)},$$ (7b)

where p_{ij} represents the probability of species i's migration to patch j, ρ and $1-\rho$ represent the frequencies of migration events from the migration pool to the upstream and downstream patches, respectively, μ scales the

frequencies of migration events of species i from the upstream to downstream patch, and $\delta(x_{i1})$ is a function to give the frequency of species i's migration from the upstream to the downstream patch, depending on the abundance of the focal species in the upstream patch. In this manuscript, we consider the simplest form of this function: species abundances do not affect the frequency of migration from the upstream to the downstream patch if species exist in the upstream patch. This form of δ clarifies the interpretation of μ's effect: we can scale the migration from the upstream to the downstream against the migration from the migration pool. If δ has another functional form, the interpretation of μ gets more complicated. The form of δ in this manuscript is written as follows:

$$
\delta(x) = \begin{cases} = 1 & (x > 0) \\ = 0 & (x \leq 0). \end{cases} \tag{8}
$$

Some parameter sets of (ρ, μ) correspond to intuitive scenarios of meta-community dynamics. For example, $(\rho, \mu) = (1, 0)$ and $\rho = 0$ are the cases where species migrate only to upstream and downstream patches, respectively. $(\rho, \mu) = (0.5, 0)$ represents the mainland-island model (Hanski and Gyllenberg, 1993), except that the upstream community affects the downstream dynamics. $\rho = 1$ with $\mu > 0$ represents the cases where the migration is perfectly hierarchical: migration only occurs from the migration pool to the upstream patch and from the upstream to the downstream, although μ changes which types of migration are more likely to occur.

The downstream stability at time point T (or $q_0(T + 1)$) is then written as follows:

$$
\begin{aligned}
q_0(T + 1) &= 1 - \sum_{i \neq 0} q_i(T + 1) \\
&= 1 - \sum_{i=1}^{N} p_{i1} \left(1 - \mathrm{Re}(i, T)\right) - \sum_{i=1}^{N} p_{i2} \left(1 - \mathrm{Ri}(i, T)\right) \\
&= \sum_{i=1}^{N} \left(p_{i1} \mathrm{Re}(i, T) + p_{i2} \mathrm{Ri}(i, T)\right).
\end{aligned} \tag{9}
$$

This equation clarifies that the migration parameters, ρ and μ, weight the effects of the two types of resistance on the downstream stability.

Two sampling scenarios

We analyzed the stability of downstream communities in two scenarios. For each scenario, we generated 60 sets of $N = 25$ species. A, $\vec{r}_1$, and $\vec{r}_2$ and varied the migration parameters for each set. See SI 1.2 for more details. , while fixing the number of species in the migration pool to

In the first, the assembly scenario (Fig. 2C), we simulate how a community might assemble in nature: starting from two empty patches, we simulate 250 migration events ($T = 1, 2, \ldots, 250$) given the values of $A = (a_{ik})$, $\vec{r}_1 = (r_{11}, \ldots, r_{N1})$, $\vec{r}_2 = (r_{21}, \ldots, r_{N2})$, ρ, and μ. We then measure the longevity of downstream species composition and sample many meta-communities. However, this method does not randomly sample meta-communities as the sampled upstream and downstream communities depend on them at the previous

time step; the statistical analysis for this sampling method would be very difficult.

The second is the design scenario (Fig. 2D), which represents how one might want to design a target community with known properties: we define 60 target downstream communities composed of randomly chosen species (species richness is randomly chosen from 1 to 10). These target communities are feasible and locally stable in the absence of upstream communities. We calculated the stability of each target downstream community over randomly generated 200 upstream communities. In this scenario, we assume that we know the downstream species composition that is the most beneficial to the host and we want to stabilize this composition. Then, we can ask what are designs of upstream communities to increase the downstream stability. In contrast to the assemble scenario, we randomly sample the upstream and downstream communities in the designing scenario: we can use these data in regression analysis. However, as some upstream communities collapse the target downstream communities, this sampling method is not effective to collect data (Fig. S2). Although it is difficult to derive the general conditions where the upstream community does not collapse the target downstream species composition, we can derive what kinds of upstream communities would not collapse a target downstream community when all interactions in the downstream community is negative by using the results of Chesson (2000) and Saavedra et al. (2017). See SI 3 for more details.

Manipulating upstream species richness

The causal inference suggested that the total strength of positive interactions from the upstream community to the downstream community increases the stability (Table 2). To validate this result, we performed two additional analyses.

One intuitive method to manipulate the positive interactions between the communities is changing the upstream species richness: adding species increases the total strength of the positive interactions while removing species decreases it. We added/removed one species to/from the upstream communities which experienced 10 migration events in the assemble scenario because the variation of the stability at the time step is large and the stability is unlikely to be one (see Fig. S1). Then, we calculated how adding or removing species changed the stability and the total strength of positive interactions from the upstream to the downstream communities. We also calculated Spearman correlation coefficients between the stability and the total strength of positive interactions.

Manipulating positive interactions

The above manipulation can change community features other than the total strength of positive interactions from the upstream to the downstream communities. To avoid this problem, we generated meta-communities where species compositions in the upstream and the downstream communities do not overlap (e.g., species 1– 8 existed in the upstream patch while species 9 – 16 existed in the downstream patch). To generate such meta-communities, we first generated 11 eight-species communities where species can coexist in the absence of the spatial structure. The coexistence was tested given two growth vectors, which represent the species' growth vector in the upstream patch and downstream patch, respectively. This means that these eight species should

coexist (i) in the upstream patch and (ii) in the downstream patch when no species interactions between the communities exist. The growth rates and the species interaction matrices were sampled as explained in SI 1.2.

Then, we allocated one of the 11 eight-species communities in the upstream patch and another to the downstream patch (in total, we have $11 \times 10 = 110$ meta-communities). As we assumed that species compositions did not overlap, we can manipulate the species interactions between communities without changing species interactions within upstream and downstream communities, respectively. To clarify the effect of positive interactions from the upstream to the downstream communities, we assumed that all interactions from the upstream species to the downstream species were positive. On the other hand, we assumed that all interactions from the downstream species to the upstream species were negative, which makes the mean of the off-diagonal elements of the interaction matrix zero. To do so, we sampled species interactions between communities from modified half-normal probability distributions. In short, a species interaction matrix is given by the following 16×16 block matrix:

$$
\left(
\begin{array}{c|c}
A_{\text{up}} & B^- \\
\hline
B^+ & A_{\text{down}}
\end{array}
\right)
\tag{10}
$$

where A_{up} and A_{down} represents species interactions within the upstream and the downstream patches, respectively, while $B^\pm$ represents interactions from the upstream species to the downstream species and vice versa. (i, k) element of either A_{up} or A_{down} (we use a_{ik} for simplicity) is give as below:

$$
a_{ik}
\begin{cases}
= -1 & \text{if } i = k \\
\sim \mathcal{N}(0, 0.25^2) & \text{otherwise.}
\end{cases}
\tag{11}
$$

On the other hand, (i, k) element of matrix $B^\pm$, $b_{ik}^\pm$, follows the following probability distribution with parameter σ:

$$
b_{ik}^\pm \sim \pm \frac{h(y)}{\sigma}
\tag{12}
$$

where

$$
y = \frac{b_{ik}^\pm}{\sigma} - 1
\tag{13}
$$

$$
h(y) = \sqrt{\frac{2}{\pi}} \exp\left(\frac{-y^2}{2}\right) \quad \text{(half-normal distribution).}
\tag{14}
$$

We manipulated the total strength of positive interactions from the upstream community to the downstream community by charging σ.

We fixed the migration parameter values as $(\rho, \mu) = (0.5, 0.5)$ to exclude the effects of these parameters on the stability.

4.3 Statistical analysis

Causal inference

After calculating the stability of the downstream communities, we statistically analyzed how features of communities (Table 1) relate to the stability of the downstream communities. As the stability is the probability that the downstream composition is maintained, we performed the logistic regression analysis using logit function in statsmodels.formula.api (Seabold and Perktold, 2010). The community features were calculated from realized species interactions within or between the two patches, the number of species in each patch, and the two migration parameters (Table 1). These features were standardized so that means are 0s and the standard deviations are 1s using preprocessing.StandardScaler.fit().transform() function in package scikit-learn (Pedregosa et al., 2011). This enabled us to compare the effects of each feature without considering the difference in their scales.

In the main text, we analyzed the causal effects of the following five features related to the upstream communities: species richness in the upstream, total strength of positive or negative interactions within the upstream community, and total strength of positive or negative interactions from the upstream to the downstream. The rationale of this analysis is as follows: if we want to stabilize a certain downstream community that is beneficial to a host but the manipulation of the downstream community is difficult, we may be able to change the stability of the downstream community by manipulating the upstream community. For the causal inference in the absence of the spatial structure or the effects of other features, see SI 2 and SI 6, respectively.

To perform the causal inference (Arif and MacNeil, 2022), we assumed the causal relationships between the community features and the stability metrics represented by Fig. 3. First, we can expect that species richness increases the total strength of positive/negative interactions as well as the mean degree within upstream and downstream communities, respectively, because the number of species interactions increases with the number of species. Species richness in both communities should increase the total strength of positive/negative interactions and a degree from the upstream community to the downstream community for the same reason. According to the previous studies (Stachowicz et al., 1999; Bonanomi et al., 2014; Hromada et al., 2021), species richness increases the resistance to invasion because of resource competition. Although our model (Eqs (1a) and (1b)) does not explicitly include resource competition, species interaction in the model can relate to resource competition (i.e., positive interactions: providing resources, and negative interactions: competing for resources). Then, we can expect that the total strength of interactions affects the resistance to invasion within upstream and downstream communities, respectively. To our best knowledge, there is no study that suggests the causal relationship between the degrees and the two types of resistance, and Eqs (1a) and (1b) do not include the degrees. For these reasons, we do not consider the causal relationships from the degrees to the stability metrics. We do not consider the causal effects of mean strength of positive or negative interactions on the stability metrics for similar reasons. The resistance to invasion in *the upstream community* would increase the resistance to the environmental change because the downstream environment would not change if no species can colonize the upstream community. In addition, we can expect that the total strength of positive and negative interactions from the upstream to the downstream communities affect the resistance to the environmental change and resistance to invasion: the

strong interactions indicate that the changes in the upstream community propagate the downstream dynamics, and the interactions from the upstream to the downstream affect the growth rates of invader species in the downstream community (see Eq (5)). Finally, the two migration parameters ρ and μ affect the resistance to the environmental change and resistance to invasion as these migration parameters determine whether species are more frequently invade the upstream communities or the downstream communities. We changed which features should be included in the logistic regression depending on whose effects we want to estimate (Table S9) to satisfy the backdoor criterion (Hernán and Robins, 2020) in the causal diagram of Fig 3.

Statistical tests

In this manuscript, we performed Wilcoxon rank-sum test, Wilcoxon signed-rank test, and Spearman correlation analysis using scipy version 1.6.2 (Virtanen et al., 2020) with Python version 3.8.8.

4.3.1 Programming

All mathematical and statistical analyses were performed in Python (version 3.8.8) with the following packages: numpy version 1.20.1 (Harris et al., 2020), scipy version 1.6.2 (Virtanen et al., 2020), statsmodels version 0.13.1 (Seabold and Perktold, 2010), and scikit-learn version 0.24.1 (Pedregosa et al., 2011).

Chapter 6

General discussion

6.1 Synthesis of the results

In this thesis, I investigated two design approaches to optimize microbial community functions and analyzed how these approaches relate to the fundamental aspects of microbial ecology. The first approach is introducing environmental fluctuations to control the dynamics without feedback. The second approach is, on the other hand, allocating species in acyclic (or hierarchical) networks to restrict species interactions. In Chapter1, I introduced the following research questions:

1. How can one maximize community function by fluctuating the environment without feedback?

2. How do environmental fluctuations affect species composition and diversity while affecting the intensity of demographic noise?

3. How can one maximize community function by introducing hierarchical spatial structures?

   (a) What is the best spatial structure to maximize community function?

   (b) What are efficient methods to find the best allocations of species to a given network structure?

4. How does an upstream community affect the stability of a downstream community in a hierarchical network?

From the studies in Chapters 2 – 5, the answers are as follows:

1. We can derive the optimal fluctuations of inflow toxin concentrations to maximize long-term detoxification efficiency.

2. The coupling effects of the environmental fluctuations and demographic noise enable slower-growing species to outcompete the faster-growing species, but the probability of such events shows complex patterns. These complex patterns also appear in the heterogeneity of communities (beta diversity).

3. I derive the conditions for efficiently maximizing the community functions in hierarchical (or acyclic) networks of chemostats.

   (a) The chain networks tend to maximize the community functions.

(b) It is sufficient to analyze the microbial allocations where all allocated species survive to maximize the community functions.

4. Positive interactions from upstream species to downstream species stabilize the downstream communities.

In Chapter 2, I addressed the optimization of detoxification in a chemostat model. Because detoxification can be seen as a kind of public good games (Samuelson, 1954; Broom et al., 2018), mutants that do not degrade the toxin (i.e., non-detoxifiers) can spread in the population (Ellis et al., 2007; O'Brien et al., 2014). However, Chapter 2 suggests that if individuals can also differ in the private resistance to the toxin (e.g., due to efflux pumps (Blair et al., 2015; Bottery et al., 2016; Rojo-Molinero et al., 2019)), we can recover the detoxification by re-inoculating detoxifiers that have a different private resistance level from the resident non-detoxifiers while manipulating the toxin concentration flowing into the chemostat. In addition, we can calculate the optimal frequencies of the manipulation of the inflow toxin concentration and inoculation of detoxifiers without observing what kind of individuals exist in the population.

Although Chapter 2 assumed the deterministic dynamics (i.e., the model does not include demographic noise), Chapter 3 suggested a non-trivial effect of demographic noise under the environmental fluctuations. Due to the demographic noise, species (or strains) that should go extinct in the deterministic dynamics can outcompete the competitors with some probability, and this probability depends on the rate of environmental changes (changes in the inflow resource and/or toxin concentrations, see Table 1 of Chapter 3) and species' sensitivities to the toxin (affecting species abundances and the intensity of demographic noise). In other words, the optimization protocol in Chapter 2 may fail due to the demographic noise coupled with the environmental fluctuations. Therefore, I add two remarks to the optimization protocol of Chapter 2. First, inoculation sizes of detoxifiers should be enough large so that demographic noise can be ignored. Second, we should have a long duration between the inoculations (or environmental fluctuations) so that non-detoxifiers should go extinct before the next inoculation; as we should change environments when inoculating detoxifiers, demographic noise and the environmental fluctuations might help the non-detoxifies.

In Chapter 4, I addressed optimization of general microbial community functions in acyclic networks of multi-stage chemostats. Spatial structures enable species to coexist by restricting species interactions (Kim et al., 2008). Acyclic networks prevent the species interactions from the downstream chemostats to the upstream ones. Although there are many possible microbial allocations to the chemostats, I implemented two algorithms to find the best allocation efficiently (Algorithms 1 and 2). In addition, Appendix C.3 shows that chain networks tend to maximize microbial community functions and that we do not have to analyze other acyclic networks.

Although the model in Chapter 4 considers the local stability, other stability metrics should be considered. For example, upstream species may contaminate in the downstream chemostats, which can affect the community functions. In the model of Chapter 4, I assumed an ideal filter: species cannot move between the chemostats due to the filters but the media can flow between the chemostats. However, in the master project of Oliveira Sudário (2022), making such filters was

difficult: species were clogged in the chemostats. Chapter 5 provides solutions to this problem. For example, we can introduce resistance to invasion into the model while removing the filters between the chemostats: we consider only allocations that upstream species have smaller growth rates than the dilution rate in the downstream chemostats. In this case, the upstream species cannot colonize downstream chemostats. Another way is to consider resistance to the environmental changes defined in Chapter 5. If we know which species contribute to the community functions (e.g., *A. tumefaciens*, *C. testosteroni*, and *O. anthropi* for degrading ampicillin in Fig. 4.6), we may allow the invasion of species to the downstream chemostats unless the target species do not go extinct. From these points of view, we can combine the results of Chapters 4 and 5. According to Chapter 5, we can increase resistance to invasion and/or the environmental changes when we have strong positive interactions from the upstream to the downstream. Although the model in Chapter 5 assumes only the two patches (or chemostats), we can speculate that this may also be the case with more patches. Fig. C.6 suggests that the optimal allocations allocate species so that positive interactions from the upstream to the downstream tend to occur. Then, resistance to invasion and/or the environmental changes in the optimal allocations in Chapter 4 may be high.

6.2 Outlook and future directions

6.2.1 Comparing three approaches

In this thesis, I investigated two of the three approaches (i.e., optimal compositions, optimal control, and optimal allocation) to optimize microbial community functions. I have not, however, compared the performances of these three approaches given a microbial community and a target community function. Although the optimal composition approach would be the easiest to experimentally implement, we can implement environmental fluctuations in a chemostat and build multi-stage chemostats (at least in a chain network) using Chi.Bio. (Steel et al., 2019, 2020). Then, we need to consider which optimization approach we should use to maximize the community function given by Eqs (1.1) or (1.2).

Let us consider the example of degrading ampicillin with the four microbial species in Chapter 4 to compare the optimal allocation approach with the optimal composition approach. Fig. 4.6 indicates that cultivating two or three species in the multi-stage chemostats can have higher community functions than those when three or four species are co-cultured without the spatial structure. However, this might not be the case if all interspecific interactions were positive. Species abundances would be lower in the hierarchical structure than in the well-mixed co-culture because the hierarchical spatial structures remove species interactions from downstream species to upstream species. This would decrease the degradation efficiency of ampicillin because the species can degrade ampicillin more with larger abundances. I speculate that species-species and/or species-compounds interaction networks affect whether the optimal composition approach or the optional allocation approach gives higher community functions.

We can also compare the optimal control approach with the optimal composition. One advantage of the controlling approach is that we can generate species coexistence that could not

happen without environmental fluctuations. However, this does not necessarily imply that the optimal control approach provides a higher community function than the optimal composition approach. If a target community function requires the coexistence of species that cannot coexist without environmental fluctuations, and/or if species richness positively correlates with the community function (Dell'Anno et al., 2012; Demeter, 2015), the control approach would be a better option than the optimal composition approach.

Comparison between the optimal allocation and the optimal control is difficult because the optimal allocation approach partially includes the optimal composition and control problems. In the optimal allocation, we need to choose species to culture as in the optimal composition problem and we can manipulate the medium flowing into a focal chemostat by changing the order of species allocated to the upstream chemostats. The manipulation of the medium flowing into stage i in the optimal allocation problem then depends on the possible allocations with length $i-1$: if we have only a few allocations with length $i-1$, we have a small capacity of manipulating the medium flowing into stage i. In addition, the number of allocations with length $i-1$ depends on the species coexistence from stage 1 to stage $i-1$, which would depend on the species-species or species-compound interactions. I expect that we first need to compare the optimal composition approach with the optimal control and the optimal allocation approaches, respectively; then we can hypothesize when the optimal allocation approach gives the higher community functions than the optimal control approach.

From these considerations, I emphasize the connections between maximizing microbial community functions and the fundamental microbial ecology. To understand which is the best approach to maximize a community function, species-species or species-compound interaction networks, and species diversity would be useful. Understanding these fundamental aspects could help researchers to decide how to maximize a target community function.

6.2.2   Computational costs of algorithms

We may also care about the computational costs of each approach and algorithm. If the difference in the maximum community functions is similar among the optimal composition, control, and, allocation approaches, we may choose the approach with the smallest computational cost. However, there are three obstacles to compare the computational costs. Firstly, we cannot always explicitly write the (upper boundaries of) computational costs of the algorithms. This means that we need to evaluate the computational costs in simulations. Secondly, some studies do not report their computational costs (Angulo et al., 2019; Wang et al., 2020; George and Korolev, 2021). Finally, the measures for the computational costs can differ depending on the mathematical backgrounds of the algorithms. In the control approach with the reinforcement learning, Treloar et al. (2020) reported the number of episodes, as many studies of the neural network did. In Fig. 4.2 in Chapter 4, on the other hand, I show the elapsed time of the deterministic algorithm (Algorithm 1). Reporting elapsed time may be a good option because we can report this measure for any simulations and analyses. However, we need to keep in mind that the elapsed time depends both on the algorithm and the details of programming codes such as used computer

languages and libraries. We, therefore, need to standardize the details of the code to compare the computational costs.

6.2.3 Reformulating community functions

In addition, we may reformulate the objective function given by Eqs (1.1). In Chapters 2 and 4, I evaluate the community functions with Eq (1.1) at linearly stable equilibria. With this stability, I did not consider the invasion of species (or mutants in Chapter 2) from the inflow media. In reality, contamination in the inflow media or non-ideal filter in the multi-stage chemostats would cause such species invasion, which can change community functions. Then, we can introduce the resistance to invasion and resistance to environmental changes (see Chapter 5) into the objective function. One simple objective function with the resistance is given by

$$\phi\left(\vec{C}(t)\right) \propto \vec{w}\vec{C}^{\mathrm{T}}(t) + WR, \tag{6.1}$$

where $R \geq 0$ represents the resistance to invasion and resistance to environmental changes, and $W \geq 0$ is a weight for the resistance. Instead of R, we can use $f(R) > 0$ with $df/dR \geq 0$ (i.e., high resistance increases the objective function value) to include non-linear evaluations on the resistance. For example, we may neglect the slight loss of the resistance when the resistance is enough high.

Another idea to include these types of resistance in the objective function is to evaluate the mean of the community function at each time step while considering the stochastic migration of species as in Chapter 5. Similar to Eq (1.2), we can evaluate the long-term community functions:

$$\Phi(T) \equiv \int_0^T dt \mathbb{E}\left[\phi\left(\vec{C}(t)\right)\right], \tag{6.2}$$

where $\mathbb{E}\left[\phi\left(\vec{C}(t)\right)\right]$ represents the mean community function value at time t. If the resistance of the designed community is high, $\mathbb{E}\left[\phi\left(\vec{C}(t)\right)\right]$ would be close to the designed value at each time step because invades rarely colonize the community and the community function value is unlikely to change. Indeed, we used such an objective function to evaluate the long-term detoxification efficiency in Chapter 2, although we considered the invasion of mutants and a discrete time scale in that paper. In this example, the resistance is small because the mutants that differ in the resistance to the toxin but do not detoxify can always invade the population. See Box 1 of Chapter 2 or S7 in Appendix A for more details.

6.2.4 Modelling the dynamics

All my analyses in this thesis assumed that the governing dynamics of the systems are given. However, formulating the dynamics is also challenging. Here, I discuss (i) how to estimate the parameter values of the equation, and (ii) how to determine the equations to represent the dynamics.

Parameter estimation

Specifying the parameter values is essential to predict the dynamics and their equilibrium. In addition, we need to consider the uncertainty of parameter values because experimental data typically have noise. I used the two methods to estimate the parameter values from the experimental data (Dos Santos, 2019; Oliveira Sudário, 2022) in Chapter 4: the Bayesian estimation and least squares estimation (LSE). In the Bayesian statistics, we can calculate the posterior probability distributions of the parameter values. One advantage of the Bayesian estimation is that we can intuitively understand the results: using the parameter values sampled from the posterior probability distributions, we can calculate the probability distributions of community functions. These probability distributions of the community functions are useful when we compare the community functions. Let us consider the case where we want to know whether allocation 1 or allocation 2 has a higher community function. The probability distributions of the two allocations' community functions tell us the probability distribution of the difference in the two allocations' community functions. Then, we can calculate how likely allocation 1 is to have a higher community function than allocation 2. As in Chapter 4, however, we need a certain amount of data to perform the Bayesian estimation with Markov Chain Monte Carlo.

The other way to estimate models' parameter values in Chapter 4 is LSE. LSE provides a set of parameter values but does not show the uncertainty of the parameter values. To calculate the uncertainty of the parameter values (and community functions), we need to perform a bootstrap analysis. We keep in mind, however, that we deal with the time-series data, where each observation is not independent. This leads the technical difficulty in bootstrapping time-series data (Buhlman, 2002). Of course, if we have many time-series data, we may not have to consider the uncertainty of the parameter values and it would be sufficient to perform LSE. We could collect many time-series data by, for example, automatically measuring species abundances using Chi.Bio (Steel et al., 2019, 2020).

Formulating model

It is also challenging to determine an appropriate model to predict microbial community dynamics. As discussed in Subsection 1.1.3, there are two frameworks to represent the dynamics: generalized Lotka-Volterra (gLV) models or consumer-resource (CR) models. gLV models consider phenomenological interactions typically in linear equations while CR models explicitly represent species interactions via chemical compounds. Momeni et al. (2017) show that gLV models sometimes quantitatively or qualitatively fail in predicting microbial dynamics, which suggests the advantage of CR models over gLV models. However, time-series data of chemical compounds in media are not always available (see Chapter 4). In such cases, using gLV models with the dynamics of target compounds would be a good option. To avoid the problem suggested by Momeni et al. (2017), we may need to formulate gLV models with non-linear interaction terms and to choose the best formulations by calculating information criteria (Hooten and Hobbs, 2015) such as Akaike information criterion (using likelihood functions) or Watanabe-Akaike information

criterion (with posterior probability distributions) (Watanabe, 2010).

Alternatively, we may be able to infer governing equations from time-series data (i.e., model discovery). For example, Bongard and Lipson (2007) and Schmidt and Lipson (2009) use symbolic regression to perform model discovery. Brunton et al. (2016) proposed sparse identification of nonlinear dynamics (SINDy), which predicts the governing dynamics when the time-series data of all variables are given. As it is difficult to obtain the time series data of chemical compounds in media, SINDy would be used to predict the dynamics of gLV models but not CR models. Recently, Bakarji et al. (2022) expand SINDy so that we can infer the dynamics with hidden variables. This means that we may be able to infer CR models from the time-series data of species abundances (and some chemical compounds that can be measured). In addition, Daniels and Nemenman (2015) suggest another method to infer phenomenological models (e.g., gLV models) with hidden variables. Using these methods, we may be able to infer ordinary differential equations from time-series data, and to optimize microbial community functions.

6.3   Conclusion

In my thesis, I have analyzed two design approaches to maximize microbial community functions: the optimal control approach via environmental fluctuations, and the optimal allocation approach in hierarchical spatial structures. In addition, I investigated the effects of these approaches on the fundamental aspects of ecology such as diversity and stability. In Chapter 2, I showed the optimization of the detoxification efficiency over the evolutionary time scale by modifying the inflow toxin concentration into a chemostat. These environmental fluctuations can change the magnitude of demographic noise, and I analyzed the coupling effects of the environmental fluctuations with demographic noise on species composition and diversity (Chapter 3). In Chapter 4, I introduced two algorithms that find the best microbial allocations in a hierarchical network of multi-stage chemostats and showed the predicted allocation of microbial species to degrade ampicillin using experimental data. I also investigated how upstream communities affect the stability of downstream communities in the simple hierarchical spatial structure in Chapter 5. The positive interactions from the upstream community to the downstream community stabilize the downstream community. These results showcase the importance of environmental fluctuations and spatial structures in both fundamental and applied microbial ecology and evolution, which are strongly connected.

Appendix A

Appendices of Chapter 2

The following appendices are also available at `https://onlinelibrary.wiley.com/doi/10.1111/eva.13050`

Supplementary information

Appendix S1 Mono-culture analysis

In this section, we analyze the mono-culture of each of the four strategies. From Eqs (1) and (5) in the main text, the null-clines for strategy i are written as follows:

$$T = \frac{\alpha T_{\text{in}}}{\alpha + f(x_{\text{Co}})} \tag{S.1a}$$

$$x_i = 1 - \frac{\delta_i(T) + \alpha}{r_i} \tag{S.1b}$$

$$x_i = 0. \tag{S.1c}$$

Eq (S.1a) represents a decreasing function of x_{Co} (the density of the cooperators; i.e., in the mono-culture $x_{\text{Co}} = x_i$ where $i \in \{\text{sCo}, \text{rCo}\}$, otherwise $x_{\text{Co}} = 0$) and this function converges to $T = \alpha T_{\text{in}}/(\alpha + f_{\text{max}})$ in the limit $x_{\text{Co}} \to \infty$. Eq (S.1b) is, on the other hand, a decreasing function of T. The last null-cline, Eq (S.1c) is the T axis.

There exist two equilibria. One is the trivial equilibrium $(T, x) = (T_{\text{in}}, 0)$, which is an intersection of Eqs (S.1a) and (S.1c). The other is a non-trivial one $(T, x) = (T^*, x_i^*)$ where T^* and x_i^* are defined in Eqs (9a) and (9b). Such an equilibrium is the intersection of Eqs (S.1a) and (S.1b). To investigate the linear stability of this equilibrium, we evaluate the 2×2 Jacobian matrix:

$$J = \begin{pmatrix} -\{f(x_{\text{Co}}) + \alpha\} & -Tf'(x_{\text{Co}}) \\ -\delta_i'(T)x_i & r_i(1 - 2x_i) - \delta_i(T) - \alpha \end{pmatrix}. \tag{S.2}$$

An equilibrium is linearly stable if and only if the maximum value of the real part of the eigenvalue of the Jacobian matrix is negative. From the Routh-Hurwitz criteria (Murray, 2002, Appendix B), the maximum value of the real part of the eigenvalue of the 2×2 Jacobian matrix is negative if and only if the following two inequalities are satisfied:

$$\text{tr}J < 0, \tag{S.3a}$$

$$\det J > 0. \tag{S.3b}$$

At the trivial equilibrium in the mono-culture of cooperators $i \in \{\text{sCo}, \text{rCo}\}$, the trace and the determinant of the Jacobian matrix are

$$\text{tr}J = r_i - \delta_i(T_{\text{in}}) - 2\alpha, \tag{S.4a}$$

$$\det J = -\alpha \{r_i - \delta_i(T_{\text{in}}) - \alpha\}. \tag{S.4b}$$

If $r_i > \delta_i(T_{\text{in}}) + \alpha$, the trivial equilibrium is unstable because $\det J < 0$. On the other hand, the trivial equilibrium is stable if $r_i < \delta_i(T_{\text{in}}) + \alpha$ as the trace of the Jacobian matrix is negative and the determinant of the Jacobian matrix is positive. When the former condition holds, cooperators can invade from rare.

At the non-trivial equilibrium for cooperators $i \in \{\text{sCo}, \text{rCo}\}$, the trace of the Jacobian matrix is always negative because

$$\text{tr}J = -\{f(x_i^*) + \alpha\} + \underbrace{\{r_i(1 - x_i^*) - \delta_i(T^*) - \alpha\}}_{=0 \quad \because \text{Eq (9b)}} - r_i x_i^* < 0. \tag{S.5}$$

Therefore, the non-trivial equilibrium is stable if and only if

$$
\begin{aligned}
\det J &> 0 \\
\Leftrightarrow \{f(x_i^*) + \alpha\} r_i x_i^* - T^* f'(x_i^*) \delta_i'(T^*) x_i^* &> 0 \\
\Leftrightarrow r_i \{f(x_i^*) + \alpha\} &> T^* f'(x_i^*)\delta_i'(T^*) \\
\Leftrightarrow -\frac{r_i}{\delta_i'(T^*)} &< -\frac{\alpha T_{\text{in}} f'(x_i^*)}{\{\alpha + f(x_i^*)\}^2} \quad (\because T^* = \alpha T_{\text{in}}/\{\alpha + f(x_i^*)\}) \tag{S.6}
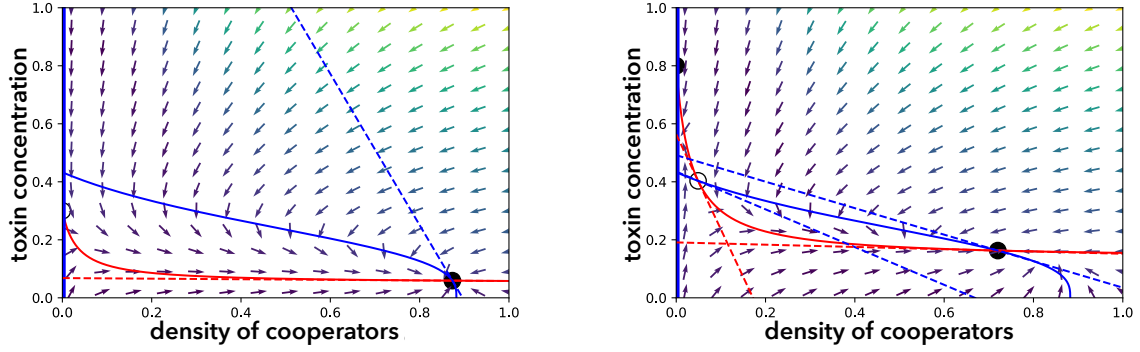\end{aligned}
$$

1

Figure S.1: Phase plane analysis

Two examples of the phase plane analysis of a mono-culture of sensitive cooperators are shown ($i = $ sCo). The red solid line represents the null-clines defined by Eq (S.1a) and the two blue solid lines show the null-clines of Eqs (S.1b) and (S.1c). The dashed lines represent the tangents of the null-clines at the non-trivial equilibria given by Eqs (S.7a) and (S.7b), respectively. Left: When $r_i > \delta_i (T_{\mathrm{in}}) + \alpha$, the dynamics converge to the non-trivial stable equilibrium (black circle). Right: When $r_i < \delta_i (T_{\mathrm{in}}) + \alpha$, the non-trivial equilibrium closest to the trivial equilibrium is unstable (white circle), although there exists another non-trivial equilibrium that is stable. Notice that the trivial equilibrium $(T_{\mathrm{in}}, 0)$ is unstable if $r_i > \delta_i (T_{\mathrm{in}}) + \alpha$ (left) but stable if $r_i < \delta_i (T_{\mathrm{in}}) + \alpha$ (right). The parameter values are: left $T_{\mathrm{in}} = 0.3$, $n = 3$, $c_d = 0.15$, $r = 1$, $K_d = 0.2$, $K_s = 0.3$, $d_{\max} = 1.0$, $f_{\max} = 0.5$, and $\alpha = 0.1$; right $T_0 = 0.8$ and the other parameter values are the same as in the left panel. Notice that in the case of the mono-culture of resistant cooperators, the parameters should be changed as follows: $K_s \to K_r$ and $c_d \to c_d + c_r$.

Inequality (S.6) represents the relationship between the slope of the tangent lines of the two null-clines of Eqs (S.1a) and (S.1b) at the non-trivial equilibrium. The two tangent lines are given as below:

$$T = -\frac{\alpha T_{\mathrm{in}} f'(x^*)}{\{\alpha + f(x^*)\}^2} (x - x^*) + T^*, \tag{S.7a}$$

$$T = -\frac{r_i}{\delta_i' (T^*)} (x - x^*) + T^*. \tag{S.7b}$$

If inequality (S.6) is satisfied, the null-cline Eq (S.1a) should exist under (above) the null-cline Eq (S.1b) at $x_i = x_i^* + \epsilon$ (or, $x_i = x_i^* - \epsilon$) where $0 < \epsilon \ll 1$.

When $r_i > \delta_i (T_{\mathrm{in}}) + \alpha$ is satisfied, there exists at least one stable and non-trivial equilibrium point. In a mono-culture of cooperators, the null-cline Eq (S.1a) is a decreasing function of the density of the cooperators x_i and converges to $\alpha T_{\mathrm{in}} / (\alpha + f_{\max})$ in the limit of $x_i \to \infty$ while the null-cline Eq (S.1b) is a decreasing function of the toxic concentration T and converges to $1 - \alpha/r_i$ in the limit of $T \to 0$. Notice that $r_i > \alpha$ is a necessary condition for $r_i > \delta_i (T_{\mathrm{in}}) + \alpha$. When $r_i > \delta_i (T_{\mathrm{in}}) + \alpha$ is satisfied, $0 < 1 - \alpha/r_i < 1$, and therefore, there exists at least one intersection of the two null-clines. If a unique intersection exists, this point is a stable and non-trivial equilibrium point because the null-cline Eq (S.1b) exists above the null-cline Eq (S.1a) within $x \in [0, x_i^*)$ (Figure S.1 left). If instead multiple intersections exist, all those that satisfy inequality (S.6) are stable and the dynamics from (x_0, T_{in}) converge to the intersection with the smallest value of x_i^*. It should be noted that when $r_i < \delta_i (T_{\mathrm{in}}) + \alpha$, the trivial equilibrium is stable, and the dynamics from (x_0, T_{in}) converge to it, even if a second stable non-trivial equilibrium exists (Figure S.1 right). In summary, cooperators can grow and the population dynamics converge to a non-trivial equilibrium if and only if and only if $r_i > \delta_i (T_{\mathrm{in}}) + \alpha$ where $i \in \{\mathrm{sCo}, \mathrm{rCo}\}$.

In the case of a mono-culture of cheaters, the null-cline Eq (S.1a) is a line $T = T_{\mathrm{in}}$. If cooperator i satisfies $r_i > \delta (T_{\mathrm{in}}) + \alpha$, there exist a unique stable equilibrium for cheaters that have the same level of resistance as the cooperators, because the cheaters' intrinsic growth rate is larger than that of the cooperators. To sum up, we obtain inequality (8) in the main text.

Appendix S2   Conditions where cooperators invade cheaters

Here, we analyze the conditions where one strategy i can invade a population of strategy j. It is assumed here that the population of strategy j converges to the mono-culture equilibrium state prior to invasion (i.e., $(x_j, T) = (x_j^*, T_j^*)$). From Eq (9b), the following equation is satisfied:

$$r_j \left(1 - x_j^*\right) - \delta_j \left(T_j^*\right) - \alpha = 0. \tag{S.8}$$

Now, let us assume that a small number of cells in the population mutate into strategy i. Then, strategy i can invade the population if and only if

$$r_i \left(1 - x_j^*\right) - \delta_i \left(T_j^*\right) - \alpha > 0, \tag{S.9}$$

because the total cell density is still x_j^*.

For ease of following the analysis, let us define $W_i(T)$ as shown in Eq (2) in the main text. Note that strategy i can grow at toxin concentration T if $W_i(T) > 1/\left(1 - \sum_j x_j\right)$. However, $W_i(T) > W_j(T)$ implies that strategy i can increase faster or decrease slower than strategy j. From Eq (S.8) and inequality (S.9), strategy i invades the population of strategy j if and only if

$$W_i \left(T_j^*\right) > W_j \left(T_j^*\right). \tag{S.10}$$

Note that cooperators cannot invade a population of cheaters which have the same level of resistance $(i,j) \in \{(\mathrm{sCo}, \mathrm{sCh}), (\mathrm{rCo}, \mathrm{rCh})\}$ because cooperators have a lower growth rate than cheaters due to the cost of cooperation ($r_i < r_j$), while the death rates are the same at any toxin concentration ($\delta_i(T) = \delta_j(T)$):

$$W_i(T) < W_j(T), (i,j) \in \{(\mathrm{sCo}, \mathrm{sCh}), (\mathrm{rCo}, \mathrm{rCh})\}. \tag{S.11}$$

For the same reason, cheaters exclude a population of cooperators that have the same level of resistance once the cheaters appear.

However, it is possible that cooperators can invade a population of cheaters with a different level of resistance at certain toxin concentrations. In addition, resistant cooperators (cheaters) can invade sensitive cooperators (cheaters), respectively, and vice versa. For such cases, inequality (S.10) is simplified using Eqs (3) and (4) in the main text as follows:

$$W_i \left(T_j^*\right) > W_i \left(T_j^*\right)$$

$$\Leftrightarrow \quad \frac{r_i}{\delta_i \left(T_j^*\right) + \alpha} > \frac{r_j}{\delta_j \left(T_j^*\right) + \alpha}$$

$$\Leftrightarrow \quad \frac{r_i}{d_{\max} \frac{T_j^{*n}}{T_j^{*n} + K_i^n} + \alpha} > \frac{r_j}{d_{\max} \frac{T_j^{*n}}{T_j^{*n} + K_j^n} + \alpha}$$

$$\Leftrightarrow \quad \alpha \left(r_i - r_j\right) \left(\tau_j^* + K_i^n\right) \left(\tau_j^* + K_j^n\right) + d_{\max} \left\{r_i \tau_j^* \left(\tau_j^* + K_i^n\right) - r_j \tau_j^* \left(\tau_j^* + K_j^n\right)\right\} > 0$$

$$\Leftrightarrow \quad F_{ij} \left(\tau = \tau_j^*\right) \equiv \left(\alpha + d_{\max}\right) \left(c_j - c_i\right) \tau_j^{*2} + \left\{d_{\max}(1 - c_i) K_i^n - d_{\max}(1 - c_j) K_j^n + \alpha(c_j - c_i)(K_i^n + K_j^n)\right\} \tau_j^*$$
$$+ \alpha(c_j - c_i) K_i^n K_j^n > 0, \tag{S.12}$$

where $\tau = T^n$, and c_i and c_j represent the total cost of strategies i and j, respectively. Note that as $F_{ij}(\tau)$ is a quadratic function of τ, one can analytically find the range of τ_j^* where $F_{ij}\left(\tau = \tau_j^*\right)$ is satisfied once the parameter values used in inequality (S.12) have been obtained (Figure S.2).

At the point when a small number of resistant (sensitive) cooperators invade a population of sensitive (resistant) cheaters ($(i,j) \in \{(\mathrm{rCo}, \mathrm{sCh}), (\mathrm{sCo}, \mathrm{rCh})\}$), the toxin concentration is close to that which is flowing into the chemostat ($T_j^* = T_{\mathrm{in}}$). In these two cases, the conditions where cooperators successfully invade cheaters with a different level of resistance at a certain toxin level are analytically derived. First, it is necessary to derive the range of toxin concentrations where cheaters can be maintained in mono-culture. In other words, we need to derive the range of T_{in} where the density of cheaters j is positive at an equilibrium state. This condition is derived by starting from the population equilibrium x_j^* and solving the inequality for T_{in}:

$$x_j^* = 1 - \frac{\delta_j \left(T_{\mathrm{in}}\right) + \alpha}{r_j} > 0$$

$$\Leftrightarrow \quad r_j > \delta_j \left(T_{\mathrm{in}}\right) + \alpha$$

$$\Leftrightarrow \quad \left(d_{\max} - a_j\right) \tau_{\mathrm{in}} < a_j K_j^n$$

$$\Leftrightarrow \quad \begin{cases} \tau_{\mathrm{in}} \le 1 & d_{\max} - a_j < a_j K_j^n \\ \tau_{\mathrm{in}} < a_j K_j^n / (d_{\max} - a_j) & \text{otherwise} \end{cases} \tag{S.13}$$
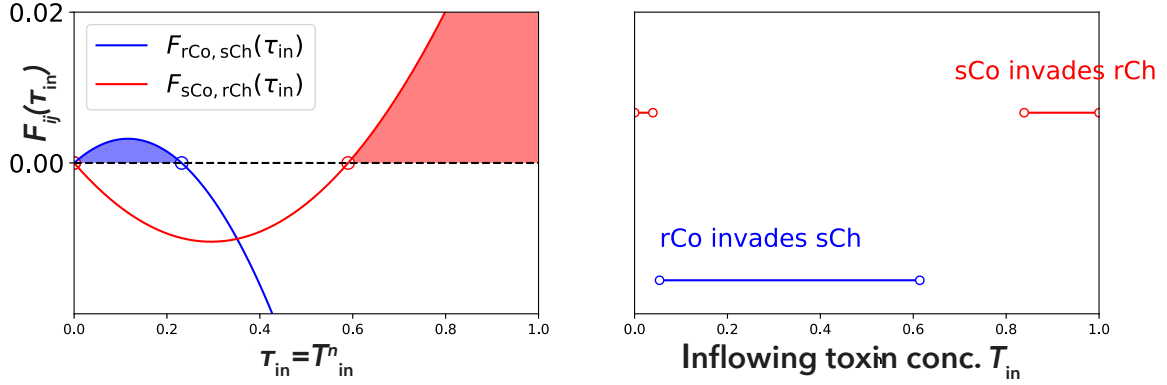
Figure S.2: Toxin concentration range where invasion of cooperators into a population of cheaters succeeds.

The range of τ_{in} (left) or T_{in} (right) where the invasion of resistant cooperators into a population of sensitive cheaters (blue) or the invasion of sensitive cooperators into a population of resistant cheaters (red) succeeds. In the left panel, one can analytically find the roots of $F_{ij}(\tau_{\text{in}})$ because $F_{ij}(\tau_{\text{in}})$ is a quadratic function. The shaded blue or red area corresponds to the value of τ_{in} where invasion succeeds. In the right panel, the ranges of T_{in} where the invasion succeeds in the left panel are shown; at low or high toxic concentrations, resistant cheaters are invaded by sensitive cooperators. On the other hand, at intermediate toxic concentrations, sensitive cheaters are invaded by resistant cooperators. Parameter values are $c_d = 0.1$, $c_r = 0.3$, $K_r = 0.6$, $K_s = 0.1$, $n = 3$, $\alpha = 0.1$, and $d_{\max} = 0.5$.

where $\tau_{\text{in}} = T_{\text{in}}^n$ and $a_j = r_j - \alpha$. Note that $a_j > 0$ is a necessary condition that strategy j can grow in mono-culture. Here, the maximum value of τ_{in} is denoted by $\hat{\tau}_{\text{in}}$:

$$\hat{\tau}_{\text{in}} = \begin{cases} 1 & \text{if } d_{\max} - a_j < a_j K_j^n \\ a_j K_j^n / (d_{\max} - a_j) & \text{otherwise} \end{cases} \tag{S.14}$$

When resistant cooperators can invade sensitive cheaters, the quadratic function $F_{\text{rCosCh}}(\tau_{\text{in}})$ should be positive:

$$\begin{aligned} F_{\text{rCosCh}}(\tau_{\text{in}}) &= -(\alpha + d_{\max})c\tau_{\text{in}}^2 + \{d_{\max}(1-c)K_r^n - d_{\max}K_s^n - \alpha c(K_s^n + K_r^n)\}\tau_{\text{in}} \\ &\quad -\alpha c K_s^n K_r^n > 0 \end{aligned} \tag{S.15}$$

$$\Leftrightarrow F_1(\tau_{\text{in}}) \equiv -F_{\text{rCosCh}}(\tau_{\text{in}}) < 0 \tag{S.16}$$

where $c = c_d + c_r$. The condition that $F_1(\tau_{\text{in}})$ is negative for different values of $\tau_{\text{in}} = [0., \hat{\tau}_{\text{in}}]$ is either (i) the existence of one real root of $F_1(\tau)$ within $\tau_{\text{in}} = [0, .\hat{\tau}_{\text{in}}]$

$$F_1(\hat{\tau}_{\text{in}}) < 0 \tag{S.17}$$

or (ii) the existence of two real roots of $F_1(\tau_{\text{in}})$ (Eq (S.15)) within $\tau_{\text{in}} = [0, \hat{\tau}_{\text{in}}]$

$$\begin{cases} F_1(\hat{\tau}_{\text{in}}) \geq 0, \\ D \equiv \{d_{\max}(1-c)K_r^n - d_{\max}K_s^n - \alpha c(K_r^n + K_s^n)\}^2 - 4(d_{\max} + \alpha)\alpha(K_r K_s)^n c^2 > 0 \\ \hat{\tau}_{\text{in}} > \frac{1}{2(d_{\max}+\alpha)c}\{d_{\max}(1-c)K_r^n - d_{\max}K_s^n - \alpha c(K_r^n + K_s^n)\} > 0, \end{cases} \tag{S.18}$$

because $F_1(0) > 0$. Note that the second and third inequalities of inequalities (S.18) represents the conditions where there exist two real roots of $F_1(\tau)$, and where the axis of symmetry exists within $\tau_{\text{in}} = [0, \hat{\tau}_{\text{in}}]$, respectively. Inequality (S.17) is simplified as below:

$$\begin{aligned} F_1(\hat{\tau}_{\text{in}}) &< 0 \\ \Leftrightarrow c &< \frac{d_{\max}(K_r^n - K_s^n)\hat{\tau}_{\text{in}}}{(d_{\max}+\alpha)\hat{\tau}_{\text{in}}^2 + \{dK_r^n + \alpha(K_r^n + K_s^n)\}\hat{\tau}_{\text{in}} + \alpha K_r^n K_s^n}. \end{aligned} \tag{S.19}$$

4

Inequalities (S.18) are, on the other hand, rewritten as below:

$$F_1(\hat{\tau}_{\mathrm{in}}) \geq 0$$

$$\Leftrightarrow \quad c \geq \frac{d_{\max}\left(K_r^n - K_s^n\right)\hat{\tau}_{\mathrm{in}}}{(d_{\max} + \alpha)\hat{\tau}_{\mathrm{in}}^2 + \{d_{\max}K_r^n + \alpha(K_r^n + K_s^n)\}\hat{\tau}_{\mathrm{in}} + \alpha K_r^n K_s^n}, \quad (\text{S.20a})$$

$$\hat{\tau}_{\mathrm{in}} > \frac{1}{2(d_{\max} + \alpha)c}\left\{d_{\max}(1 - c)K_r^n - d_{\max}K_s^n - \alpha c(K_r^n + K_s^n)\right\} > 0$$

$$\Leftrightarrow \quad \frac{d_{\max}\left(K_r^n - K_s^n\right)}{d_{\max}K_r^n + \alpha\left(K_r^n + K_s^n\right)} > c > \frac{d_{\max}\left(K_r^n - K_s^n\right)}{2(d_{\max} + \alpha)\hat{\tau}_{\mathrm{in}} + d_{\max}K_r^n + \alpha\left(K_r^n + K_s^n\right)}, \quad (\text{S.20b})$$

$$D > 0$$

$$\Leftrightarrow \quad \begin{cases} c < \frac{d_{\max}(K_r^n - K_s^n)}{d_{\max}K_r^n + \alpha(K_r^n + K_s^n) + 2\{\alpha(\alpha + d_{\max})K_r^n K_s^n\}^{1/2}} \quad \text{or} \\ (d_{\max} - c)K_r^n - d_{\max}K_s^n - \alpha c(K_r^n + K_s^n) < -2\{(1 + \alpha)\alpha\}^{1/2}(K_r K_s)^{n/2}c < 0. \end{cases} \quad (\text{S.20c})$$

Note that one can ignore the case of $(d_{\max} - c)K_r^n - d_{\max}K_s^n - \alpha c(K_r^n + K_s^n) < 0$ when analyzing $D > 0$ because of the left-hand side of inequality (S.20b). Then, inequalities (S.18) are summarized as follows:

$$\frac{d_{\max}\left(K_r^n - K_s^n\right)}{A + 2\{\alpha(\alpha + 1)K_r^n K_s^n\}^{1/2}} > c \geq \max\left\{\frac{d_{\max}\left(K_r^n - K_s^n\right)}{(d_{\max} + \alpha)\hat{\tau}_{\mathrm{in}} + A + \alpha K_r^n K_s^n/\hat{\tau}_{\mathrm{in}}}, \frac{d_{\max}\left(K_r^n - K_s^n\right)}{2(d_{\max} + \alpha)\hat{\tau}_{\max} + A}\right\} (\text{S.21})$$

where $A = d_{\max}K_r^n + \alpha\left(K_r^n + K_s^n\right)$. It should be noted that

$$\frac{d_{\max}\left(K_r^n - K_s^n\right)}{(d_{\max} + \alpha)\hat{\tau}_{\mathrm{in}} + A + \alpha K_r^n K_s^n/\hat{\tau}_{\mathrm{in}}} \quad \lessgtr \quad \frac{d_{\max}\left(K_r^n - K_s^n\right)}{2(d_{\max} + \alpha)\hat{\tau}_{\mathrm{in}} + A}$$

$$\Leftrightarrow \quad \frac{\alpha K_r^n K_s^n}{d_{\max} + \alpha} \quad \gtrless \quad \hat{\tau}_{\mathrm{in}}^2$$

$$\Leftrightarrow \quad \left(\frac{\alpha K_r^n K_s^n}{d_{\max} + \alpha}\right)^{1/2} \quad \gtrless \quad \hat{\tau}_{\mathrm{in}} \quad (\because \hat{\tau}_{\mathrm{in}} > 0) \quad (\text{S.22})$$

Especially when $(\alpha K_r^n K_s^n/(d_{\max} + \alpha))^{1/2} < \hat{\tau}_{\mathrm{in}}$, the condition where resistant cooperators can invade a population of sensitive cheaters is simplified because inequalities (S.19) and (S.21) are combined:

$$\frac{d_{\max}\left(K_r^n - K_s^n\right)}{A + 2\{\alpha(\alpha + 1)K_r^n K_s^n\}^{1/2}} > c. \quad (\text{S.23})$$

When sensitive cooperators invade a population of resistant cheaters $((i, j) = (\mathrm{sCO}, \mathrm{rCh}))$, on the other hand, the sign of $c_j - c_i = c_r - c_d = \Delta c$ determines the shape of $F_{\mathrm{sCorCh}}(\tau)$. Sensitive cooperators can invade a population of resistant cheaters if and only if

$$F_{\mathrm{sCorCh}}(\tau_{\mathrm{in}}) = \quad (d_{\max} + \alpha)\Delta c\tau_{\mathrm{in}}^2 \quad (\text{S.24})$$
$$+ \{d_{\max}(1 - c_d)K_s^n - d_{\max}(1 - c_r)K_r^n + \alpha\Delta c\left(K_r^n + K_s^n\right)\}\tau_{\mathrm{in}}$$
$$+ \alpha\Delta c K_r^n K_s^n > 0.$$

When $\Delta c = 0$, this inequality is simplified as $K_s^n > K_r^n$, which never holds. When instead $\Delta c > 0$, sensitive cooperators can invade resistant cheaters at least at a low toxin concentration because $F_{\mathrm{sCorCh}}(0) > 0$. However, one would be interested in whether sensitive cooperators can invade resistant cheaters at a high toxin concentration. The necessary and sufficient condition for the successful invasion at a high toxin concentration is given by

$$F_{\mathrm{sCorCh}}(\hat{\tau}_{\mathrm{in}}) \quad > \quad 0$$

$$\Leftrightarrow \Delta c \underbrace{\left\{(d_{\max} + \alpha)\hat{\tau}_{\mathrm{in}} + \alpha\left(K_r^n + K_s^n\right) + \alpha K_r^n K_s^n/\hat{\tau}_{\mathrm{in}}\right\}}_{=B(\hat{\tau}_{\mathrm{in}})} + d_{\max}K_r^n c_r + d_{\max}\left(K_s^n - K_r^n\right) \quad > \quad d_{\max}K_s^n c_d$$

$$\Leftrightarrow \{B(\hat{\tau}_{\mathrm{in}}) + d_{\max}K_r^n\}c_r + d_{\max}\left(K_s^n - K_r^n\right) \quad > \quad (B + d_{\max}K_s^n)c_d$$

$$\Leftrightarrow \frac{\{B(\hat{\tau}_{\mathrm{in}}) + d_{\max}K_r^n\}c_r + d_{\max}\left(K_s^n - K_r^n\right)}{B(\hat{\tau}_{\mathrm{in}}) + d_{\max}K_s^n} \quad > \quad c_d. \quad (\text{S.25})$$

Note that $\hat{\tau}_{\mathrm{in}}$ is a function of c_r in this case as $r_j = r(1 - c_r)$. If $\Delta c < 0 \Leftrightarrow c_r < c_d$, sensitive cooperators cannot invade resistant cheaters at any toxin concentration. In this case, $F_{\mathrm{sCorCh}}(\tau_0)$ is an upper convex
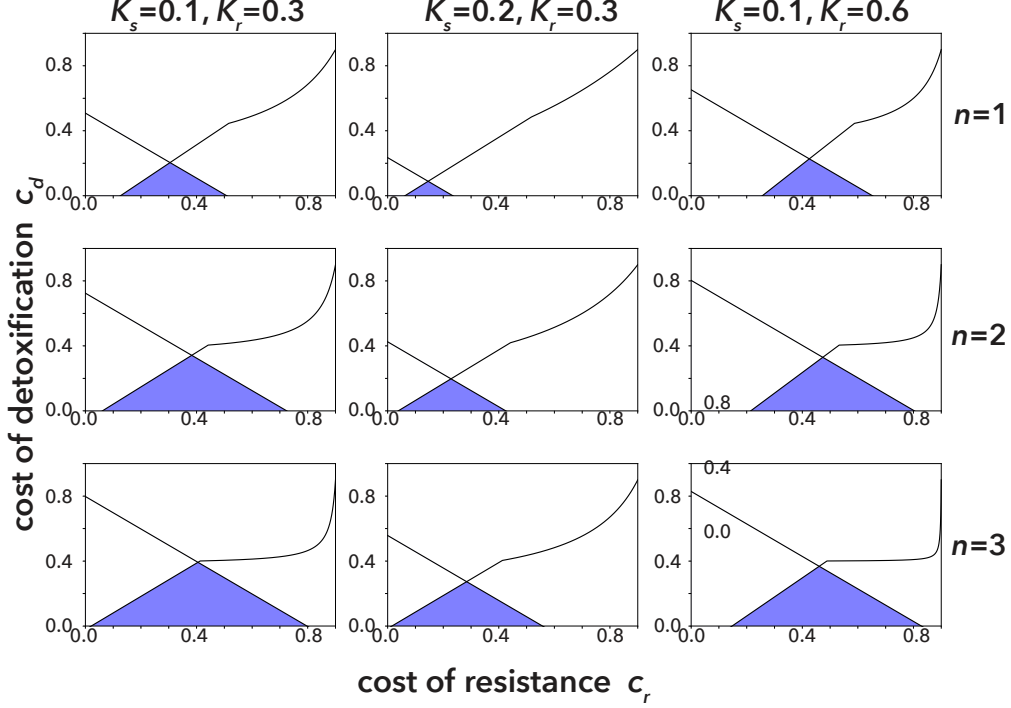
Figure S.3: The regions of the two costs where the two cooperators can invade the cheaters.

In each panel, the blue triangle area represents the region of c_r and c_d where the invasion of resistant cooperators to the sensitive cheaters, and the invasion of the sensitive cooperators to the resistant cheaters are successful. In the fist, second, and the third column, $(K_s, K_r) = (0.1, 0.3), (0.2, 0.3)$, and $(0.1, 0.3)$, respectively. In the first, second, and third row, $n = 1, 2$, and 3, respectively. The value of $d_{\max}$ is fixed as $d_{\max} = 0.5$ in all panels.

function and $F_{\mathrm{sCorCh}}(0) < 0$. In addition, the axis of the symmetry of $F_{\mathrm{sCorCh}}(\tau_{\mathrm{in}})$ is negative as below:

$$
\begin{aligned}
c_r &< c_d \\
\Rightarrow (1 - c_d)K_s^n &< (1 - c_r)K_r^n \quad (\because K_s < K_r) \\
\Rightarrow d_{\max}(1 - c_d)K_s^n - d_{\max}(1 - c_r)K_r^n + \alpha\Delta c\,(K_r^n + K_s^n) &< 0 \quad (\because \Delta c < 0). \quad (\text{S.26})
\end{aligned}
$$

In other words, $F_{\mathrm{sCorCh}}(\tau_0)$ is a decreasing function of τ within $\tau = [0., \hat\tau_0]$ when $c_r < c_d$. In summary, sensitive cooperators can invade resistant cheaters when both $c_r > c_d$ and inequality (S.25) hold. Figure S.3 shows the region of c_r and c_d and toxin concentrations where resistant cooperators can invade sensitive cheaters and sensitive cooperators can invade resistant cheaters given the values of K_r, K_s, n and $d_{\max}$.

Of course, it is possible to analyze the condition where resistant cheaters can invade sensitive cheaters, and vice versa. The resistant (sensitive) cheaters invade the sensitive (resistant) cheaters when

$$
W_{\mathrm{rCh}}\,(T_{\mathrm{in}}) \gtrless W_{\mathrm{sCh}}\,(T_{\mathrm{in}}) \quad (\text{S.27})
$$

holds, respectively. As F_{rChsCh} is a quadratic function, one can analytically find the range of τ_{in} where the invasion is successful once the parameter values are obtained. When resistant (sensitive) cooperators invade the sensitive (resistant) cooperators, on the other hand, it should be noted that it is impossible to analytically calculate the value of the toxin concentration at which invasion occurs (T_j^*). Once this toxin concentration when invasion occurs is numerically calculated by finding the equilibrium in the mono-culture, however, one can find whether the invasion is successful or not using inequality (S.12).

Appendix S3   Stability of the equilibria in evolutionary dynamics

As shown in Appendix S2, there can exist a range of toxin concentrations where resistant (sensitive) cooperators can invade a population of sensitive (resistant) cheaters, respectively. After the successful

6

invasion, however, it is unclear whether resistant (sensitive) cooperators sweep out sensitive (resistant) cheaters or they coexist because the invasion of cooperators decreases the toxin concentration, thereby decreasing the death rate of each strategy. In this section, we fist show the equilibrium where two strategies i and j coexist. Then, we derive the conditions for which the equilibria are stable and the invader i sweeps out the resident strategy j or where these two strategies coexist. Although these analyses can be performed using arbitrary pairs of the four strategies (sCo, rCo, sCh, and rCh), we mainly focus on the pairs of cooperators and cheaters whose resistance levels are different (i.e., $(i, j) \in \{(\mathrm{rCo}, \mathrm{sCh}), (\mathrm{sCo}, \mathrm{rCh})\}$), because they are the only cases where a pair of strains can stably coexist.

When strategy i coexists with strategy j at toxin concentration $T_{ij}^{\dagger}$ the following equation should be satisfied:

$$W_i\left(T_{ij}^{\dagger}\right) = W_j\left(T_{ij}^{\dagger}\right)$$

$$\Leftrightarrow \quad \frac{r_i}{\delta_i\left(T_{ij}^{\dagger}\right) - \alpha} = \frac{r_j}{\delta_j\left(T_{ij}^{\dagger}\right) - \alpha}$$

$$\Leftrightarrow \quad F_{ij}\left(\tau_{ij}^{\dagger}\right) = 0, \tag{S.28}$$

where $\tau_{ij}^{\dagger} = T_{ij}^{\dagger n}$. In other words, $\tau_{ij}^{\dagger}$ is a real root of $F_{ij}(\tau)$. Therefore, cooperators cannot coexist with cheaters which have the same level of resistance (i.e., $(i, j) \in \{(\mathrm{sCo}, \mathrm{sCh}), (\mathrm{rCo}, \mathrm{rCh})\}$) because the fitness of the cooperator is always lower than that of the cheaters as shown in inequality (S.11). When resistant (sensitive) cooperators coexist with sensitive (resistant) cheaters (i.e., $(i, j) \in \{(\mathrm{rCo}, \mathrm{sCh}), (\mathrm{sCo}, \mathrm{rCh})\}$), the densities of cooperators $x_i^{\dagger}$ and cheaters $x_j^{\dagger}$ are uniquely determined because the density of cooperators is determined by the toxin concentration $T_{ij}^{\dagger}$ at equilibrium (by setting Eq (5) to 0 and solving for $x_i^{\dagger}$ and $x_j^{\dagger}$):

$$T_{ij}^{\dagger} = \frac{\alpha T_{\mathrm{in}}}{\alpha + f\left(x_i^{\dagger}\right)}$$

$$\Leftrightarrow x_i^{\dagger} = \frac{\alpha K_d(T_{\mathrm{in}} - T_{ij}^{\dagger})}{f_{\max} T_{ij}^{\dagger} - \alpha(T_{\mathrm{in}} - T_{ij}^{\dagger})}, \tag{S.29a}$$

$$r_j\left(1 - x_i^{\dagger} - x_j^{\dagger}\right) = \delta_j\left(T_{ij}^{\dagger}\right) + \alpha$$

$$\Leftrightarrow x_j^{\dagger} = 1 - x_i^{\dagger} - \frac{\delta_j\left(T_{ij}^{\dagger}\right) + \alpha}{r_j}$$

$$= 1 - x_i^{\dagger} - \frac{\delta_i\left(T_{ij}^{\dagger}\right) + \alpha}{r_i} \tag{S.29b}$$

because $W_i\left(T_{ij}^{\dagger}\right) = W_j\left(T_{ij}^{\dagger}\right)$.

When two types of cooperators or cheaters coexist (i.e., $(i, j) \in \{(\mathrm{sCo}, \mathrm{rCo}), (\mathrm{sCh}, \mathrm{rCh})\}$), the densities of two strategies i and j are not uniquely determined. When two types of cooperators coexist, we can find the toxin density $T_{ij}^{\dagger}$ where they can coexist by calculating roots of $F_{i,j}(\tau)$. This toxin concentration is, however, determined only by the total density of cooperators, and therefore, the density of each cooperator type is not uniquely determined. When the two types of cheaters coexist, on the other hand, the toxin concentration remains T_{in}. In other words, the necessary condition that two types of cheaters coexist is that the relative fitness of each type of cheater is the same. As the toxin density is not affected by the density of cheaters, the densities of the two types of cheaters cannot be determined. Notice that $0 < x_i^{\dagger}, x_j^{\dagger} < 1$ should be satisfied; otherwise the dynamics never converge to the equilibrium $\left(T^{\dagger}, x_i^{\dagger}, x_j^{\dagger}\right)$.

These inequalities are rewritten as follows when $(i, j) \in \{(\mathrm{sCo}, \mathrm{rCo}), (\mathrm{sCh}, \mathrm{rCh})\}$:

$$
\left\{
\begin{array}{l}
0 < x_i^\dagger < 1 \\
0 < x_j^\dagger < 1
\end{array}
\right.
$$

$$
\Leftrightarrow \quad 0 < x_i^\dagger < 1 - \frac{\delta_j\left(T_{ij}^\dagger\right) + \alpha}{r_j}
$$

$$
\Leftrightarrow \quad
\left\{
\begin{array}{l}
0 < x_i^\dagger \Leftrightarrow \frac{\alpha T_{\mathrm{in}}}{\alpha + f_{\mathrm{max}}} < T_{ij}^\dagger < T_{\mathrm{in}} \\
0 < x_j^\dagger
\end{array}
\right.
\tag{S.30}
$$

When strategy i can invade the population of strategy j at toxin concentration $T = T_j^*$, there exist two types of equilibria. One is where strategy i excludes strategy j; i.e., $(T, x_i, x_j) = (T_i^*, x_i^*, 0)$ where (T_i^*, x_1^*) represents the stable equilibrium in the mono-culture of strategy i. The other is where the two strategies coexist; i.e., $(T, x_i, x_j) = \left(T_{ij}^\dagger, x_i^\dagger, x_j^\dagger\right)$, where $x_i^\dagger, x_j^\dagger$ is uniquely determined when $(i, j) \in \{(\mathrm{rCo}, \mathrm{sCh}), (\mathrm{sCo}, \mathrm{rCh})\}$. Although there exist at most two values of $T^\dagger$ in each pair of i, j as $F_{ij}(\tau)$ is a quadratic function, we shall find that only one value of $T^\dagger$ can provide a stable equilibrium.

To analyze the linear stability of each equilibrium, we evaluate 3×3 Jacobian matrix given as below:

$$
J = \left(
\begin{array}{ccc}
-\{\alpha + f(x_{\mathrm{Co}})\} & -T\partial f(x_{\mathrm{Co}})/\partial x_i & -T\partial f(x_{\mathrm{Co}})/\partial x_j \\
-x_i \delta_i'(T) & r_i(1 - 2x_i - x_j) - \delta_i(T) - \alpha & -r_i x_i \\
-x_j \delta_j'(T) & -r_j x_j & r_j(1 - x_i - 2x_j) - \delta_j(T) - \alpha
\end{array}
\right)
\tag{S.31}
$$

where x_{Co} is the total density of cooperators. The partial differentiation of $f(x_{\mathrm{Co}})$ is

$$
\frac{\partial f(x_{\mathrm{Co}})}{\partial x_i} =
\left\{
\begin{array}{ll}
f_{\mathrm{max}} \frac{K_d}{(x_{\mathrm{Co}} + K_d)^2} & i = \mathrm{sCo}, \mathrm{rCo} \\
0 & \text{otherwise.}
\end{array}
\right.
\tag{S.32}
$$

From the Routh-Hurwizt criteria, an equilibrium is linearly stable if and only if

$$
\mathrm{tr} J \quad < \quad 0
\tag{S.33a}
$$

$$
\det J \quad < \quad 0
\tag{S.33b}
$$

$$
\sum_{k=1}^{3} M_{kk} \quad > \quad 0.
\tag{S.33c}
$$

where M_{kk} is the (k, k) minor of the Jacobian matrix defined by Eq (S.31).

At the equilibrium where resistant (sensitive) cooperators exclude sensitive (resistant) cheaters $(T^*, x_i^*, 0)$, the trace, the determinant, and the minors of the Jacobian matrix are

$$
\mathrm{tr} J \;=\; -\{f(x_i^*) + \alpha\} - r x_i^* + \{r_j(1 - x_i^*) - \delta_j(T_i^*) - \alpha\},
\tag{S.34a}
$$

$$
\det J \;=\; \{r_j(1 - x_i^*) - \delta_j(T_i^*) - \alpha\} M_{33},
\tag{S.34b}
$$

$$
M_{11} \;=\; -r x_i^* \{r_j(1 - x_i^*) - \delta_j(T^*) - \alpha\},
\tag{S.34c}
$$

$$
M_{22} \;=\; -(f(x_i^*) + \alpha)\{r_j(1 - x_i^*) - \delta_j(T_i^*) - \alpha\},
\tag{S.34d}
$$

$$
M_{33} \;=\; \{\alpha + f(x_i^*)\} r_i x_i^* - T_i^* f'(x_i^*) x_i^* \delta_i'(T^*).
\tag{S.34e}
$$

It should be noted that M_{33} should be positive because strategy i persists in mono-culture. Therefore, the determinant of the Jacobian matrix is negative if and only if

$$
r_j(1 - x_1^*) - \delta_j(T^*) - \alpha < 0.
\tag{S.35}
$$

Then, M_{11} and M_{22} are positive and the trace of the Jacobian is negative. In short, the necessary and sufficient condition for the linear stability of the equilibrium where only strategy i exists while strategy j goes extinct is given by inequality (S.35). Note that at this equilibrium, the following equation should hold:

$$
r_i(1 - x_i^*) - \delta_i(T_i^*) - \alpha = 0.
\tag{S.36}
$$

Combining with this equation, inequality (S.35) is rewritten as below:

$$
W_i(T_i^*) \quad > \quad W_j(T_i^*)
$$

$$
\Leftrightarrow F_{ij}(\tau_i^*) \quad > \quad 0.
\tag{S.37}
$$

8

In other words, if the relative fitness of strategy i is larger than strategy j at toxin concentration T_i^*, the equilibrium where strategy i excludes j is stable.

At an equilibrium where two strategies i and j coexist $\left(T_{ij}^\dagger, x_i^\dagger, x_j^\dagger\right)$, the Jacobian matrix is simplified as below:

$$J = \begin{pmatrix} -\left\{\alpha + f(x_{\mathrm{Co}}^\dagger)\right\} & -T^\dagger \partial f\left(x_{\mathrm{Co}}^\dagger\right)/\partial x_i & -T^\dagger \partial f\left(x_{\mathrm{Co}}^\dagger\right)/\partial x_j \\ -x_i^\dagger \delta_i'\left(T^\dagger\right) & -r_i x_i^\dagger & -r_i x_i^\dagger \\ -x_j^\dagger \delta_j'\left(T^\dagger\right) & -r_j x_j^\dagger & -r_j x_j^\dagger \end{pmatrix}. \tag{S.38}$$

Obviously $\mathrm{tr}J < 0$ and $M_{11} = 0$, and therefore, we should evaluate $\det J$ and $M_{22} + M_{33}$. When $(i,j) = (\mathrm{sCo}, \mathrm{rCo})$, or $(\mathrm{sCh}, \mathrm{rCh})$, $J_{12} = J_{13}$, which leads to:

$$\det J = 0. \tag{S.39}$$

In other words, at least one eigenvalue of the Jacobian matrix is zero. Therefore, the coexistence of two types or cooperators or cheaters (i.e., $(i,j) \in \{(\mathrm{sCo}, \mathrm{rCo}), (\mathrm{sCh}, \mathrm{rCh})\}$) is not stable because the maximum value of the real part of the eigenvalues of the Jacobian matrix is zero or positive. Indeed, the coexistence of two types of cheaters is neutrally stable because $\sum_i M_{ii} = \alpha\left(r_i x_i^\dagger + r_j x_j^\dagger\right) > 0$ and $\mathrm{tr}J < 0$, suggesting that the other two eigenvalues of the Jacobian matrix are negative. Intuitively, this is because neutral selection works when sCh coexists with rCh; when one cheater changes its resistance level by mutation, there is no force to recover the densities of each type of cheater before the mutation. The coexistence of two types of cooperators, on the other hand, is neutrally stable if $\sum_i M_{ii} > 0$; otherwise this equilibrium is unstable because a positive eigenvalue exists.

When cooperators coexist with cheaters of different resistance levels (i.e., $(i,j) \in \{(\mathrm{rCo}, \mathrm{sCh}), (\mathrm{sCo}, \mathrm{rCh})\}$), however, coexistence can be stable. In this case $J_{13} = 0$, and the determinant of the Jacobian matrix is

$$\det J = -\underbrace{T^\dagger x_i^\dagger x_j^\dagger \partial f/\partial x_i}_{>0}\left\{r_i \delta_j'\left(T_{ij}^\dagger\right) - r_j \delta_i'\left(T_{ij}^\dagger\right)\right\} < 0$$

$$\Leftrightarrow \quad r_j \delta_i'\left(T_{ij}^\dagger\right) < r_i \delta_j'\left(T_{ij}^\dagger\right)$$

$$\Leftrightarrow \quad r_j \left\{\delta_i\left(T\right) + \alpha\right\}'\big|_{T=T_{ij}^\dagger} < r_i \left\{\delta_j\left(T\right) + \alpha\right\}'\big|_{T=T_{ij}^\dagger}$$

$$\Leftrightarrow \quad \frac{dG_{ij}}{dT}\bigg|_{T=T_{ij}^\dagger} < 0, \tag{S.40}$$

where $G_{ij}\left(T\right) = r_j\left\{\delta_i\left(T\right) + \alpha\right\} - r_i\left\{\delta_j\left(T\right) + \alpha\right\}$. Then, one can find

$$\frac{dG}{dT} = \frac{dG}{d\tau}\frac{d\tau}{dT} < 0 \tag{S.41}$$

$$\Leftrightarrow \quad \frac{dG}{d\tau} < 0 \quad \because \frac{d\tau}{dT} = nT^{n-1} > 0. \tag{S.42}$$

It should be noted that $G_{ij}\left(\tau\right)$ is a part of $F_{ij}\left(\tau\right)$ as below:

$$F_{ij}\left(\tau\right) = -G_{ij}\left(\tau\right)\left(\tau + K_i^n\right)\left(\tau + K_j^n\right)/r. \tag{S.43}$$

As $\tau_{ij}^\dagger$ is non-negative, the roots of $F_{ij}\left(\tau\right)$ (i.e., $\tau_{ij}^\dagger$) should be roots of $G_{ij}\left(\tau\right)$. Then,

$$\frac{dG_{ij}}{d\tau}\bigg|_{\tau=\tau_{ij}^\dagger} < 0$$

$$\Leftrightarrow \quad \frac{dF_{ij}}{d\tau}\bigg|_{\tau=\tau^\dagger} = -\frac{dG_{ij}}{d\tau}\bigg|_{\tau=\tau_{ij}^\dagger}\underbrace{\left(\tau_{ij}^\dagger + K_i^n\right)\left(\tau_{ij}^\dagger + K_j^n\right)/r}_{>0}$$

$$+ \underbrace{G\left(\tau_{ij}^\dagger\right)}_{=0}\frac{2\tau_{ij}^\dagger + K_r^n + K_s^n}{r} > 0. \tag{S.44}$$

$$\therefore \quad \det J > 0 \Leftrightarrow \frac{dF}{d\tau}\bigg|_{\tau=\tau_{ij}^\dagger} > 0. \tag{S.45}$$

9

As shown in Appendix S2, $F_{ij}(\tau)$ is a convex (concave) quadratic function when resistant (sensitive) cooperators coexist with sensitive (resistant) cheaters, and therefore, only the smaller (larger) root of $F_{ij}(\tau)$ satisfies inequality (S.45). From here, $\tau_{ij}^\dagger$ is assumed to represent the smaller (larger) root of F when resistant (sensitive) cooperators invade a population of sensitive (resistant) cheaters.

The last necessary and sufficient condition for the stable coexistence of two strategies ($(i,j) \in \{(\mathrm{rCo},\mathrm{sCh}),(\mathrm{sCo},\mathrm{rCh})\}$) is that the sum of the three minors of the Jacobian matrix given by Eqs (S.34c) - (S.34e) is positive:

$$\sum_i M_{ii} > 0 \quad \Leftrightarrow \quad \left\{\alpha + f\left(x_i^\dagger\right)\right\}\left(r_i x_i^\dagger + r_j x_j^\dagger\right) > T_{ij}^\dagger f'(x_i^\dagger) x_i \delta_i'(T_{ij}^\dagger),$$

$$\Leftrightarrow \quad -\frac{r_i x_i^\dagger + r_j x_j^\dagger}{x_i^\dagger \delta_i'\left(T_{ij}^\dagger\right)} < -\frac{\alpha T_{\mathrm{in}} f'\left(x_i^\dagger\right)}{\left\{\alpha + f\left(x_i^\dagger\right)\right\}^2}, \tag{S.46}$$

because $M_{11} = 0$. Notice that the sufficient condition of inequality (S.46) is:

$$-\frac{r_i}{\delta_i'\left(T_{ij}^\dagger\right)} < -\frac{\alpha T_{\mathrm{in}} f'\left(x_i^\dagger\right)}{\left\{\alpha + f\left(x_i^\dagger\right)\right\}^2} \tag{S.47}$$

because $x_j^\dagger > 0$. As we have shown in Appendix S1, inequality (S.47) represents the relationship between the slopes of the two tangent lines of the two null-clines below:

$$T = \frac{\alpha T_{\mathrm{in}}}{\alpha + f(x_i)}, \tag{S.48a}$$

$$x_i = 1 - \frac{\delta_i(T) + \alpha}{r_i} - x_j. \tag{S.48b}$$

By fixing $x_j = x_j^\dagger$, one can find that Eq (S.1b) is a special case of Eq (S.48b) when $x_j^\dagger = 0$. In other words, the null-cline given by Eq (S.48b) is obtained by parallel moving the null-cline Eq (S.1b) along the x_i axis, and the intersection(s) of the two null-clines given by Eqs (S.48a) and (S.48b) is an equilibrium $\left(T^\dagger, x_j^\dagger, x_j^\dagger\right)$ (Fig. S.4). When $x_j^\dagger = 0$, on the other hand, the intersection(s) of the two null-clines represents the equilibrium of $(T_i^*, x_i^*, 0)$. It should be noted that the dynamics never converge to the equilibrium where strategies i and j coexist when $T_i^* > T_{ij}^\dagger \Leftrightarrow \tau_i^* > \tau_{ij}^\dagger$, because $x_2^\dagger < 0$. As $F_{\mathrm{rCo},\mathrm{sCh}}(\tau)$ is a convex quadratic function, the equilibrium $(T_{\mathrm{rCo}}^*, x_{\mathrm{rCo}}^*, 0)$ is stable when the equilibrium where these two strategies coexist is not feasible. In other words, there exists at most one stable equilibrium when resistant cooperators invade sensitive cheaters. When sensitive cooperators invade resistant cheaters, on the other hand, there exist at most two stable equilibria as $F_{\mathrm{sCo},\mathrm{rCh}}(\tau)$ is a concave quadratic function (Fig. S.5). If the two equilibria are unstable, oscillations occur (Fig. S.6). However, Monte-Carlo simulations suggest that the parameter range which causes the oscillations is very small (Fig. S.7).
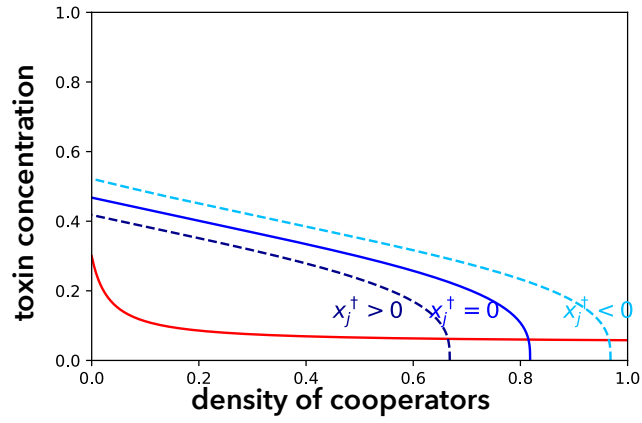
Figure S.4: Schematic illustration of the phase plane analysis in the evolutionary dynamics

Fixing $x_j = x_j^\dagger$, one can find the equilibria in $x_i - T$ phase plane. The equilibrium where cooperators exclude cheaters is the intersection of the null-clines given by Eqs (S.1a) and (S.1b) with $x_j^\dagger = 0$ (solid red and blue lines, respectively). If the equilibrium where cooperators and cheaters coexist exists, $x_2^\dagger$ should be positive (dashed navy line). In this case $T_{ij}^\dagger > T_i^*$. If $T_{ij}^\dagger < T_i^*$, on the other hand, $x_2^\dagger < 0$ (dashed sky blue line), meaning that no such equilibrium exists within $0 < T, x_i, x_j < 1$.
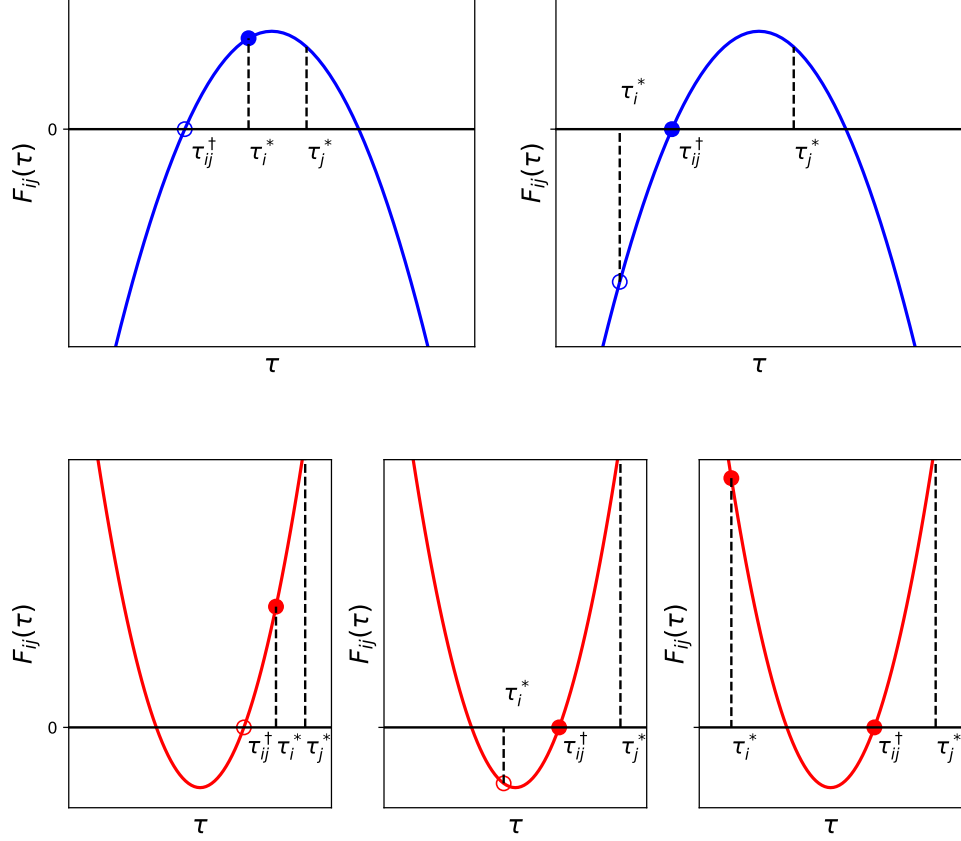
Figure S.5: Schematic illustrations of $F_{ij}(\tau)$ and the stability of two equilibria

Schematic illustrations of the quadratic function of $F_{ij}(\tau)$ and examples of two equilibria when $(i,j) = (\text{rCo}, \text{sCh})$ (top) or $(i,j) = (\text{sCo}, \text{rCh})$ (bottom) are shown. $F_{ij}\left(\tau_j^*\right) > 0$ is a necessary and sufficient condition for the invasion of strategy i into the population of strategy j in both cases. When $(i,j) = (\text{rCo}, \text{sCh})$, $F_{ij}(\tau)$ is a convex function and $\tau_{ij}^\dagger$ is the smaller root of $F_{ij}(\tau)$. If $\tau_i^* > \tau_{ij}^\dagger$ (top left), the equilibrium where the resistant cooperators exclude the sensitive cheaters is stable when $F_{ij}\left(\tau_{ij}^\dagger\right) > 0$ while the equilibrium of the coexistence does not exist as $x_j^\dagger < 0$. When $\tau_i^* < \tau_{ij}^\dagger$, on the other hand, the equilibrium where the resistant cooperators exclude the sensitive cheaters is unstable but the equilibrium of the coexistence exists. In the case of $(i,j) = (\text{sCo}, \text{rCh})$, $F_{ij}(\tau)$ is a concave function and $\tau^\dagger$ is the larger root of $F_{ij}(\tau)$. As in the case of $(i,j) = (\text{rCo}, \text{sCh})$, the equilibrium where the sensitive cooperators exclude the resistant cheaters is stable but the equilibrium of the coexistence does not exist (bottom left). However, $\tau_{ij}^\dagger > \tau_i^*$ does not mean the equilibrium of the exclusion of the resistant cheaters is unstable; when τ_i^* is larger than the smaller root of $F_{ij}(\tau)$, only the equilibrium of the coexistence can be stable (bottom center). On the other hand, if τ_i^* is smaller than the smaller root of $F_{ij}(\tau)$, the equilibrium of the exclusion is again stable, suggesting that both equilibria can be stable (bottom right).
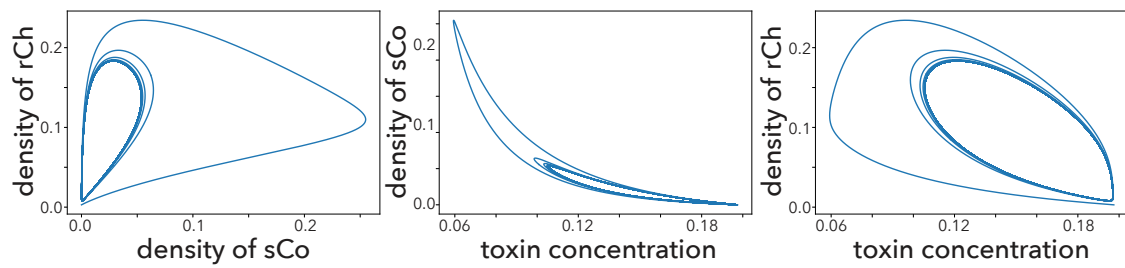
Figure S.6: Oscillation in evolutionary dynamics

Oscillation appears when the equilibria of exclusion and coexistence are unstable. These figures show an example of $(i, j) = (\text{sCo}, \text{rCh})$. Parameter values are $c_r = 0.672595$,, $c_d = 0.175532$, $K_s = 0.087769$, $K_r = 0.759002$, $n = 1$, $T_0 = 0.197481$, $\alpha = 0.120010$, $d_{\max} = 1$, $f_{\max} = 0.5$, $K_d = 0.2$, and $r = 1$.
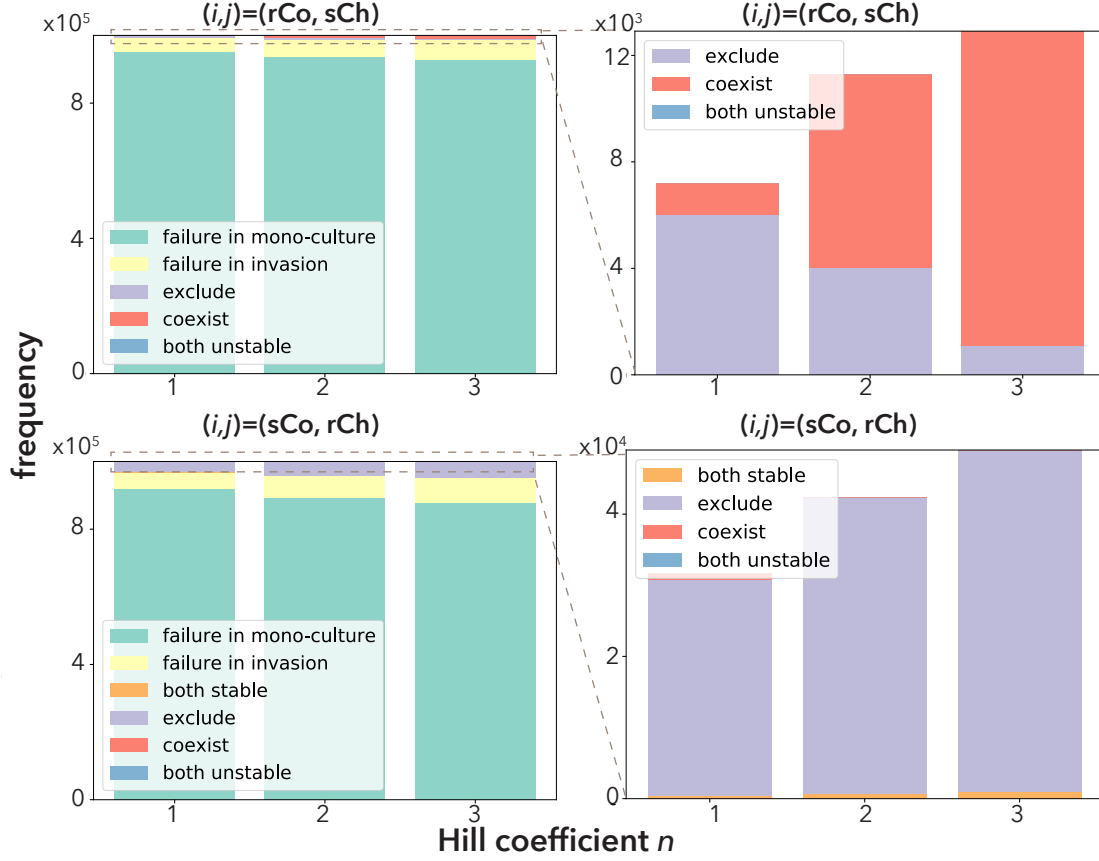
Figure S.7: Stability analysis with Monte Carlo simulation

To check whether at least one equilibrium is stable when cooperators invade cheaters, the linear stability of the equilibria of coexistence and exclusion are analyzed with a set of parameters $K_s, K_r, c_r, c_d, T_0, \alpha$ sampled using the Monte Carlo method with three values of the Hill coefficient ($n = 1, 2, 3$). Each parameter except K_r is sampled from the uniform distribution $\mathcal{U}(0, 1)$ and K_r is sampled from $\mathcal{U}(K_s, 1)$. The Monte Carlo simulation sampled 10^6 sets of parameters in each case. First row: $(i, j) = (\text{rCo}, \text{sCh})$. Regardless of the value of n, the majority of parameter sets show the failure of mono-culture of rCo and/or sCh (green in left), or the failure of the invasion of rCo into sCh (yellow in left). When rCo can invade sCh, either the equilibrium of exclusion (purple) or that of coexistence (red) is stable. In other words, there is no parameter set wherein both equilibria are unstable (blue). As n gets larger, the frequency of stable exclusion (purple) becomes smaller while that of coexistence (red) becomes larger when rCo invades sCh. Second row: $(i, j) = (\text{sCo}, \text{rCh})$. The frequency of stable exclusion increases as n grows while that of coexistence decreases. Notice that in this case, it is possible that both equilibria are stable (orange). The fixed parameter values are $d_{\max} = 1$, $f_{\max} = 0.5$, $K_d = 0.2$, and $r = 1$.

Appendix S4   Three-strategy situations

If two strategies i, j stably coexist, mutation can provide the third strategy. Notice that the stable coexistence of the two strategies occurs only if $(i, j) \in \{(\mathrm{rCo}, \mathrm{sCh}), (\mathrm{sCo}, \mathrm{rCh})\}$. As the fitness of cooperators is lower than that of cheaters which have the same level of the resistance, these cooperators (i.e., sCo and rCo, respectively) disappear even when they arrive by mutation. In this section, we shall see the evolutionary dynamics where the new cheaters (i.e., rCh or sCh) invade the population where two strategies, rCo and sCh, or sCo and rCh coexist.

At the beginning, two strategies i and j $((i, j) \in \{(\mathrm{rCo}, \mathrm{sCh}), (\mathrm{sCo}, \mathrm{rCh})\})$ coexist and the following equation should hold in Appendix S3:

$$W_i\left(T_{ij}^\dagger\right) = W_j\left(T_{ij}^\dagger\right). \tag{S.49}$$

Then, a new strategy k, a cheater with the same level of resistance as strategy i $((i, k) \in \{(\mathrm{sCo}, \mathrm{sCh}), (\mathrm{rCo}, \mathrm{rCh})\})$, can invade the population once it appears by the mutation. This is because

$$W_k(T) > W_i(T) \tag{S.50}$$

holds regardless the value of T. Hence, the two strategies i and k never coexist. In addition, the coexistence of two strategies j and k cannot be stable as the cheaters cannot change the toxin concentration (see Appendix S3). Therefore, the only stable equilibrium is the mono-culture of one of the three strategies.

Indeed, the only stable equilibrium is the mono-culture of strategy k. At the equilibrium of the mono-culture of strategy k $((T, x_i, x_j, x_k) = (T_{\mathrm{in}}, 0, 0, x_k^*))$, the Jacobian matrix is written as below:

$$J = \begin{pmatrix} -\alpha & -T\partial f/\partial x_i & 0 & 0 \\ 0 & r_i(1 - x_k^*) - \delta_i(T_{\mathrm{in}}) - \alpha & 0 & 0 \\ 0 & 0 & r_j(1 - x_k^*) - \delta_j(T_{\mathrm{in}}) - \alpha & 0 \\ -x_k^*\delta_k'(T_{\mathrm{in}}) & -r_k x_k^* & -r_k x_k^* & -r_k x_k^* \end{pmatrix}. \tag{S.51}$$

Then, the eigenvalues of the Jacobian matrix λ are the diagonal elements of the Jacobian matrix by cofactor expansion:

$$|\lambda I - J| = 0$$
$$\Leftrightarrow \quad \prod_{l=1}^{4}(\lambda - J_{ll}) = 0 \tag{S.52}$$
$$\therefore \quad \lambda = J_{ll}, l = 1, 2, 3, 4. \tag{S.53}$$

J_{11} and J_{44} are negative. In addition, J_{22} is also negative for the following reason:

$$W_i(T_{\mathrm{in}}) < W_k(T_{\mathrm{in}}) = \frac{1}{1 - x_k^*}$$
$$\Leftrightarrow \quad r_i(1 - x_k^*) - \delta_i(T_{\mathrm{in}}) - \alpha < r_k(1 - x_k^*) - \delta_k(T_{\mathrm{in}}) - \alpha = 0. \tag{S.54}$$

To evaluate the sign of J_{33}, it should be noted that

$$W_j(T_{\mathrm{in}}) < W_i(T_{\mathrm{in}}), \tag{S.55}$$

which is satisfied due to the invasion of strategy i to j. Therefore,

$$W_j(T_{\mathrm{in}}) < W_k(T_{\mathrm{in}}), \tag{S.56}$$

which suggests that $J_{33} < 0$ as shown in inequality(S.54). Thus, all eigenvalues are negative and the mono-culture of strategy k is stable. The stability of the mono-culture of the other two strategies is analyzed similarly. However, the mono-culture of strategy i or strategy j is unstable because the sign of one eigenvalue is the same as $W_k(T_i^*) - W_i(T_i^*) > 0$ or $W_k(T_{\mathrm{in}}) - W_j(T_{\mathrm{in}}) > 0$, respectively.

Appendix S5   ODE including mutation rates

In the main text, we ignore mutations as shown in Eq (1). This is because it is difficult to analytically find an equilibrium and to analyze the stability of the equilibrium when mutations are taken into account.

In this section, we show some examples of the dynamics when mutations are introduced into Eq (1). In this case, the dynamics of the density of strategy i are defined by

$$\frac{dx_i}{dt} = \left(1 - \sum_j x_j\right) \sum_j Q_{ji} r_j x_j - \{\delta_i(T) + \alpha\} x_i, \tag{S.57}$$

where Q_{ji} is the probability that strategy j mutates into i. In our scenario, there exist two independent mutations: (i) changing cooperators to cheaters and vice versa, and (ii) changing the level of resistance. The mutation probabilities are denoted as μ_1 and μ_2, respectively. The mutation matrix $Q = (Q_{ji})$ is composed of μ_1 and μ_2 as follows:

$$Q_{ji} = \begin{cases} (1 - \mu_1)(1 - \mu_2) & j = i \\ \mu_1(1 - \mu_2) & \text{mutated only in the detoxification ability} \\ (1 - \mu_1)\mu_2 & \text{mutated only in the resistance level} \\ \mu_1\mu_2 & \text{both mutated} \end{cases} \tag{S.58}$$

At an equilibrium of this case, the density of each strategy i should satisfy:

$$1 - \sum_j x_j^* = \frac{\{\delta_i(T^*) + \alpha\} x_i^*}{\sum_j r_j x_j^* Q_{ji}}, \tag{S.59}$$

where $T^* = \alpha T_{\text{in}} / \left\{\alpha + f\left(x_{\text{sCo}}^* + x_{\text{rCo}}^*\right)\right\}$. Note that $x_i^* > 0$. It is difficult to find the equilibrium state, or to analyze the stability of this equilibrium as one has to calculate a 5×5 Jacobian matrix.

However, computer simulations show the effect of mutation rates on an equilibrium. When both mutation probabilities are large ($\mu_1 = 10^{-1}, \mu_2 = 10^{-1}$), all four strategies coexist at relatively large density (Fig. S.8 top left). This is because all strategies produce the other strategies with relatively large probabilities. In other words, natural selection is weak as mutation probabilities are large.

If we reduce μ_2, the majority of cells are sCo or sCh, while the densities of rCo and rCh are close to (but not equal to) zero (Figure S.8 top right). As μ_2 is small, sensitive strains rarely produce resistant strains, while cheaters produces cooperators of the same level of resistance (and vice versa). In addition, as the toxin concentration is low, sensitive strains are more advantageous than resistant strains (see Appendix S2). For these reasons, most cells are sCo and sCh.

If μ_1 is small but μ_2 is large (Figure S.8 bottom left), cooperators almost go extinct. Cheaters rarely produce cooperators as μ_1 is small, and therefore, cooperators are excluded due to natural selection. On the other hand, both sensitive and resistant cheaters coexist as μ_2 is large.

When both mutation rates are small (Figure S.8 bottom right), one of the four strategies dominates most of the time. At the beginning, sCo grows and and the toxin concentration decreases. Following that, sCh increases and sCo are excluded. Then, rCh invades and sCh are swept out as the toxin concentration is intermediate. Notice that rCo does not increase when sCh is dominant and the toxin concentration is increasing.

In summary, multiple strategies can coexist when both or either of the two mutation probabilities are large. As the mutation probabilities decrease, this coexistence collapses because the strength of natural selection increases. In particular, cooperators almost go extinct if μ_1 is small. When both μ_1 and μ_2 are small, we can see similar evolutionary dynamics as in the state transition in Figure 3.

Appendix S6   Optimum culture conditions at the equilibrium

Defining the objective function as Eq (7), one can consider optimizing the efficiency of detoxification at an equilibrium state. As the equilibria can be classified into three types (the existence of only cheaters, coexistence of cooperators with cheaters, and a mono-culture of cooperators), we can consider the optimization problem in each equilibrium class. In this section, we shall see how to obtain the maximum detoxification efficiency at each equilibrium class by changing the dilution rate α and the toxin concentration flowing into the system T_{in}.

If there exist only sensitive or resistant cheaters in the population, the toxin concentration remains that flowing into the system $T = T_{\text{in}}$, regardless of the density of cheaters. Then, the detoxification efficiency is zero.

When cooperators coexist with cheaters of different resistance level, the toxin concentration is $T_{\text{rCo,sCh}}^{\dagger}$ or $T_{\text{sCo,rCh}}^{\dagger}$. Here, we denote strategies i and j as cooperators and cheaters, respectively. As shown
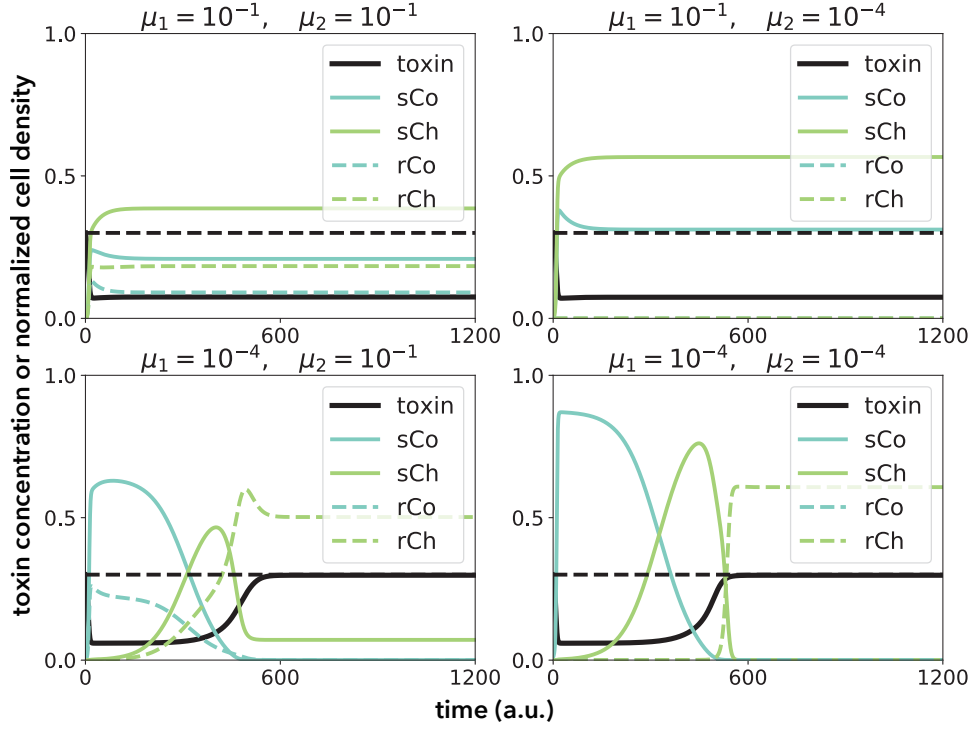
Figure S.8: Evolutionary dynamics including mutation rates in ODE

Examples of evolutionary dynamics with Eq (S.57) with different values of μ_1 and μ_2. In each panel, the dashed black line represents the toxin concentration flowing into the chemostat T_{in}, while the solid black line is the toxin concentration flowing out of the chemostat T. The other lines represent the densities of each strategy (solid dark green: sCo, solid thick lime green: sCh, dashed dark green: rCo, and dashed thick lime green rCh). In each case, the initial condition is $T(0) = T_{\text{in}}$, $x_{\text{sCo}}(0) = 0.01$, and $x_i(0) = 0$ where $i \in \{\text{rCo}, \text{sCh}, \text{rCh}\}$. The other parameter values are $\alpha = 0.1$, $T_{\text{in}} = 0.3$, $f_{\text{max}} = 0.5$, $K_d = 0.2$, $r = 1$, $c_d = 0.15$, $c_r = 0.3$, $d_{\text{max}} = 1$, $K_s = 0.3$, $K_r = 0.5$, and $n = 3$.

17

in Appendix S3, the coexistence equilibrium is analytically calculated. In this case, there exist three constraints: (i) strategy i can invade the population of j and strategy j can persist in a mono-culture ($W_i(T_{\text{in}}) > W_j(T_{\text{in}})$), (ii) there exists a smaller (larger) real root of $F_{ij}(\tau)$, and (iii) the coexistence equilibrium is stable, which is given by inequality (S.46):

$$\underset{\alpha, T_{\text{in}}}{\text{maximize}} \quad \phi(\alpha, T_{\text{in}}) = \alpha\left(T_{\text{in}} - T_{ij}^{*\dagger}\right)$$

$$\text{subject to} \quad W_i(T_{\text{in}}) > W_j(T_{\text{in}}) > 0,$$

$$F_{ij}(\tau) \quad \text{has real roots},$$

$$-\left\{r_i x_i^\dagger + r_j x_j^\dagger\right\} / \left\{x_i^\dagger \delta_i'\left(T_{ij}^\dagger\right)\right\} < -\left\{\alpha T_{\text{in}} f'\left(x_i^\dagger\right)\right\} / \left\{\alpha + f\left(x_i^\dagger\right)\right\}^2.$$

Notice that $T_{ij}^\dagger$ is independent of T_{in}. This leads to

$$\frac{\partial \phi}{\partial T_{\text{in}}} = \alpha > 0. \tag{S.60}$$

Therefore, the maximum efficiency of detoxification at this class of equilibria is given by the maximum value of $T_{\max \text{in}}$ that satisfies the constraints with each value of α. As the constraints are complex, however, it would be easier to numerically find the maximum detoxification efficiency (Fig. 4b and d).

If only cooperators i exist, the toxin concentration at equilibrium is T_i^*. In this case, one can consider two types of constraints: (i) the equilibrium (T_i^*, x_i^*) is stable (unless mutation occurs), which is defined by inequality (S.6), or (ii) strategy i can grow from low density ($r - d_i(T_{\text{in}}) - \alpha > 0$). Notice that constraint (ii) is sufficient for constraint (i). If one uses (i) as a constraint, however, the trivial equilibrium $(T, x) = (T_{\text{in}}, 0)$ can be stable (see Fig. S.1 right), which reduces the detoxification efficiency to zero. In other words, the efficiency at equilibrium depends on the initial condition if constraint (i) is used. This problem is avoided when constraint (ii) is used, because the trivial equilibrium is unstable in this case. For this reason, we used constraint (ii) for the optimization problem with this class of equilibrium.

$$\underset{\alpha, T_{\text{in}}}{\text{maximize}} \quad \phi(\alpha, T_{\text{in}}) = \alpha(T_{\text{in}} - T_i^*)$$

$$\text{subject to} \quad r_i > \delta_i(T_{\text{in}}) + \alpha.$$

Notice that it is impossible to analytically find the maximum detoxification efficiency because one cannot write T_i^* in a closed form as shown in Eq (9a). However, due to the low dimensionality of the objective function, one can numerically find its global optimum (Figure 4a and c). It should be noted that the mono-culture of cooperators can be invaded by cheaters of opposite resistance only with constraint (i) or (ii). Therefore, to achieve maximum detoxification efficiency with this class of equilibrium, it is necessary that cheaters are excluded from the population before changing the parameters (α, T_{in}) so that the detoxification efficiency of the mono-culture of cooperators is maximized.

When cooperators can coexist with cheaters of opposite resistance with certain parameter values (α, T_{in}), this parameter set gives the higher detoxification efficiency when there are only cooperators in the population. This is because the necessary condition for the existence of equilibria where cooperators can coexist with cheaters of opposite resistance is that $T_{ij}^\dagger > T_i^*$ (see Fig. S.4), although the stability of these two types of equilibria is unclear. This implies that the detoxification efficiency of a mono-culture of cooperators is larger with the parameters (α, T_{in}) at which cooperators coexist with cheaters:

$$\phi\left(\alpha, T_{\text{in}}, T_{ij}^\dagger(\alpha, T_{\text{in}})\right) < \phi(\alpha, T_{\text{in}}, T_i^*(\alpha, T_{\text{in}})). \tag{S.61}$$

Therefore, the maximum detoxification efficiency of a mono-culture of cooperators is larger than that of a co-culture of cooperators with cheaters that differ in the level of resistance.

$$\max_{\alpha, T_{\text{in}}} \phi\left(\alpha, T_{\text{in}}, T_{ij}^\dagger(\alpha)\right) < \max_{\alpha, T_{\text{in}}} \phi(\alpha, T_{\text{in}}, T_i^*(\alpha, T_{\text{in}})). \tag{S.62}$$

If cheaters can be excluded from coexistence with cooperators of different resistance by changing (α, T_{in}), the efficiency of detoxification is maximized as follows: (i) change (α, T_{in}) so that the equilibrium where a mono-culture of cooperators is stabilized while the equilibrium of coexistence is destabilized, (ii) maintain the parameter values until cheaters go extinct, and then, (iii) change the parameters (α, T_{in}) so that the detoxification efficiency given by the mono-culture of cooperators is maximized. If coexistence cannot be destabilized, the maximum detoxification efficiency is achieved just by maximizing it for the co-culture of cooperators with cheaters.

Appendix S7 Optimization of cumulative detoxification efficiency

In Appendix S2, we found that cooperators can invade a population of cheaters at some toxin concentration if the level of resistance is different. In addition, Appendix S6 shows that cooperators can exclude cheaters with different resistance level by changing the dilution rate or the toxin concentration flowing into the system, and then one can maximize detoxification efficiency in a mono-culture of cooperators. Once cheaters with the same level of resistance appear by mutation, however, cooperators are swept out and the detoxification efficiency defined by Eq (7) becomes zero as cheaters do not detoxify.

To recover detoxification, it is necessary to introduce cooperators which have a different level of resistance from the resident cheaters and to increase the toxin concentration so as to satisfy inequality (S.10), because cooperators with different resistance are unlikely to appear by mutation as they require double mutations. When the state of the population (existence of four strategies) is not observable, one problem is how often sCo and rCo should be introduced into the population in order to maximize detoxification from the beginning of the inoculation to a certain time. If cooperator inoculation probabilities are too small, the population will consist of cheaters most of the time and detoxification efficiency will be low. If cooperator inoculation probabilities are too large, cooperators are likely to be introduced into a population already containing cooperators unnecessarily, or where the invasion of cooperators fails because the resident cheaters have the same level of resistance. Such unfavorable introduction of cooperators is costly because it decreases detoxification efficiency.

If we assume a small enough mutation probability, we can redefine the time scale to discrete time steps. Then, focusing on the existence of each strategy rather than their densities, one can consider discrete population states (e.g., a mono-culture of one of the four strategies, or a transient state where two strategies coexist and one strategy is invading and sweeping out another). Combining discrete time steps and discrete population states, the model is defined as a Markov chain. We assume that only one mutation or introduction of cooperators occurs in one time step and these events never occur during the transient state. In other words, it is assumed that mutation or the introduction of cooperators occurs only when the population state is an equilibrium state of mono- or co-culture of one or two of the four strategies.

In this section, we first show the details of the special case shown in Box 1 where $\mu_2 = 0$. Note that the Markov chain remains ergodic when $\mu_2 > 0$ in this case (relaxing assumption (i)) because a positive μ_2 adds transient states between sCo-rCo and sCh-rCh, respectively, and the two states where sCo (rCo) excludes sCh (rCh) or they coexist but $(\alpha, T_{\mathrm{in}})$ are not optimized. Then, we will investigate whether relaxing the other three assumptions in Box 1 affect the ergodicity of the Markov chain or not: (ii) mutation and introduction do not occur when the population is in one of the transient states, nor at the same time, (iii) both types of cooperators can mutually exclude each other, and (iv) sCo and rCo can exclude cheaters with different resistance levels by changing α and T_{in}. In the latter part of this section, we show how to maximize the cumulative detoxification efficiency when Markov chains are not ergodic.

When $\mu_2 = 0$ and the above three assumptions in Box 1 are held, there exist at least 14 states as shown in Fig. S.9; the mono-culture of sCo (state 1), the transient state (which is unstable in short-term dynamics Eqs (1) and (5)) when sCh appears in the mono-culture of sCo by mutation (state 2), the mono-culture of sCh (state 3), the transient state when rCo is introduced to the mono-culture of sCh (state 4), the mono-culture of rCo (state 5), the transient state when rCh appears in the mono-culture of rCo by mutation (state 6), the mono-culture of rCh (state 7), the transient state when sCo is introduced to the mono-culture of rCh (state 8), the transient state when rCo is introduced to the mono-culture of sCo (state 9), the transient state when sCo is introduced to the mono-culture of rCo (state 10), the transient state when sCo is unnecessarily introduced to the mono-culture of sCo (state 11), the transient state when sCo is introduced to the mono-culture of sCh but soon excluded (state 12), the transient state when rCo is unnecessarily introduced to the mono-culture of rCo (state 13), and the transient state when rCo is introduced to the mono-culture of rCh but soon excluded (state 14). One can add more transient states between the mono-culture states when transition from one mono-culture state to another mono-culture state takes more units of time than other transitions. In addition, the number of the transition states depends on the time scale of a discrete time step. These points do not change, however, the ergodicity of the Markov chain, and we will continue the scenario with the least number of states. When $m_1, m_2 > 0$ and $m_1 + m_2 + \mu_1 < 1$, the Markov chain shown in Figure S.9 is ergodic, and therefore, an arbitrary initial probability distribution of the states of population $\boldsymbol{\pi}(0; \boldsymbol{m}, \boldsymbol{\mu})$ converges to

a unique stationary distribution $\boldsymbol{\pi}^*(\boldsymbol{m};\boldsymbol{\mu})$, which is obtained below:

$$\sum_{i=1}^{14} \pi_i^* = 1 \tag{S.63a}$$

$$\pi_1^* = (1 - \sum_{j=1}^{2} m_j - \mu_1)\pi_1^* + \pi_8^* + \pi_{10}^* + \pi_{11}^* \tag{S.63b}$$

$$\pi_2^* = \mu_1 \pi_1^* \tag{S.63c}$$

$$\pi_3^* = (1 - \sum_{j=1}^{2} m_j)\pi_3^* + \pi_2^* + \pi_{12}^* \tag{S.63d}$$

$$\pi_4^* = m_2 \pi_3^* \tag{S.63e}$$

$$\pi_5^* = (1 - \sum_{j=1}^{2} m_j - \mu_1)\pi_5^* + \pi_4^* + \pi_9^* + \pi_{13}^* \tag{S.63f}$$

$$\pi_6^* = \mu_1 \pi_5^* \tag{S.63g}$$

$$\pi_7^* = (1 - \sum_{j=1}^{2} m_j)\pi_7^* + \pi_6^* + \pi_{14}^* \tag{S.63h}$$

$$\pi_8^* = m_1 \pi_7^* \tag{S.63i}$$

$$\pi_9^* = m_2 \pi_1^* \tag{S.63j}$$

$$\pi_{10}^* = m_1 \pi_5^* \tag{S.63k}$$

$$\pi_{11}^* = m_1 \pi_1^* \tag{S.63l}$$

$$\pi_{12}^* = m_1 \pi_3^* \tag{S.63m}$$

$$\pi_{13}^* = m_2 \pi_5^* \tag{S.63n}$$

$$\pi_{14}^* = m_2 \pi_7^*. \tag{S.63o}$$

From the above simultaneous linear equations, the elements of the stationary distribution are rewritten as follows:

$$^{\mathrm{t}}\boldsymbol{\pi}^* = \begin{pmatrix} \pi_1^* \\ \mu_1 \pi_1^* \\ \mu_1 m_2^{-1} \pi_1^* \\ \mu_1 \pi_1 \\ (\mu_1 + m_2)(\mu_1 + m_1)^{-1} \pi_1^* \\ \mu_1(\mu_1 + m_2)(\mu_1 + m_1)^{-1} \pi_1^* \\ \mu_1(\mu_1 + m_2)m_1^{-1}(\mu_1 + m_1)^{-1} \pi_1^* \\ \mu_1(\mu_1 + m_2)(\mu_1 + m_1)^{-1} \pi_1^* \\ m_2 \pi_1^* \\ m_1(\mu_1 + m_2)(\mu_1 + m_1)^{-1} \pi_1^* \\ m_1 \pi_1^* \\ \mu_1 m_1 m_2^{-1} \pi_1^* \\ m_2(\mu_1 + m_2)(\mu_1 + m_1)^{-1} \pi_1^* \\ \mu_1 m_2(\mu_1 + m_2)m_1^{-1}(\mu_1 + m_1)^{-1} \pi_1^* \end{pmatrix}, \tag{S.64}$$

where $^{\mathrm{t}}\boldsymbol{\pi}^*$ represents a transpose matrix of $\boldsymbol{\pi}^*$. As the sum of the elements of $\boldsymbol{\pi}^*$ should be one, one can find

$$\pi_1^* = \frac{m_1 m_2 (m_1 + \mu_1)}{A}, \tag{S.65}$$

where

$$A = m_1 m_2 (m_1 + \mu_1)(m_1 + m_2 + 2\mu_1 + 1) + m_1 m_2 (m_2 + \mu_1)(m_1 + m_2 + 2\mu_1 + 1)$$
$$+ m_1 \mu_1 (m_1 + 1)(m_1 + \mu_1) + m_2 \mu_1 (m_2 + 1)(m_2 + \mu_1). \tag{S.66}$$

Now, we analytically obtain the stationary distribution $\boldsymbol{\pi}^*$. The expected cumulative efficiency of detoxification from the beginning of cultivation to time step $s + 1$, $\Phi(s + 1; \boldsymbol{m}, \boldsymbol{\mu})$, is given by:

$$\Phi(s + 1; \boldsymbol{m}, \boldsymbol{\mu}) = \Phi(s; \boldsymbol{m}, \boldsymbol{\mu}) + \sum_{i=1}^{14} C\phi_i \pi_i(s + 1; \boldsymbol{m}, \boldsymbol{\mu}), \tag{S.67}$$
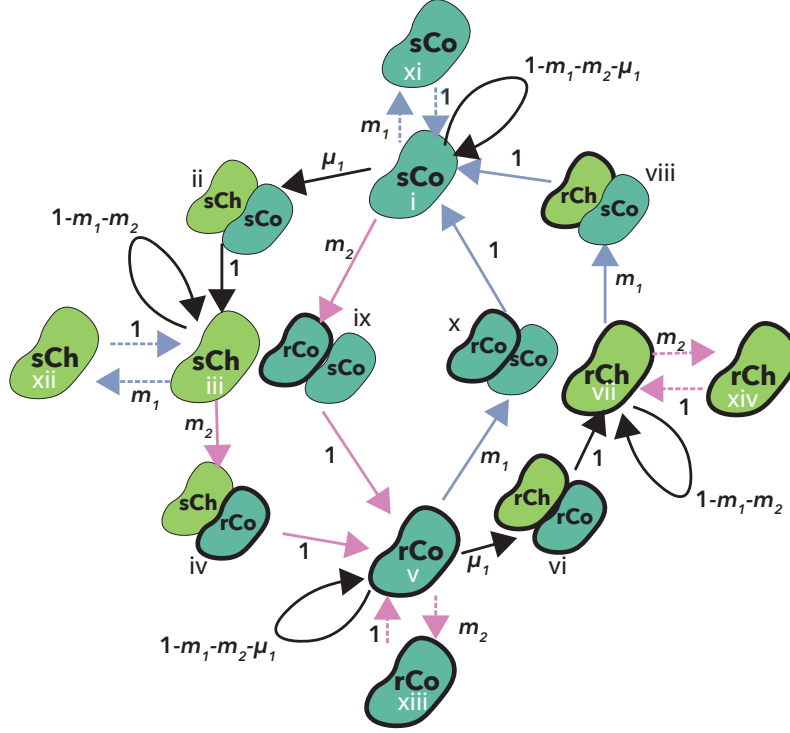
20

Figure S.9: The diagram of state transitions when mutation never occurs in the level of resistance

The diagram of the state transitions when $\mu_2 = 0$. In this example, there exist 14 states; mono-cultures of one of four strategies, transient states caused by the invasion of cheaters (solid black arrows) or by the introduction of cooperators with different resistance levels (solid blue or pink arrows), and the unnecessary introduction of cooperators which currently exist or cannot invade the population (dashed blue or pink arrows). Here, the introduction of sCo is represented by the blue arrows while the pink arrows represents the introduction of rCo. Each value along the arrows represents the corresponding transition probability. The numbers in the roman numerals correspond to state i; $i = 1$ represents the mono-culture of sCo. The optimal introduction rates of the cooperators in this model are shown in the bottom panel of Fig. 5.

where C is a positive constant to change the time scale of ϕ_i into a discrete time step. The second term of Eq (S.67) represents the expected detoxification efficiency at time step $s + 1$. The optimization problem is to find the values of m_1 and m_2 that maximize the expected cumulative efficiency given time step s, which is solved by Dynamic Programming. However, in the limit of $s \to \infty$, $\boldsymbol{\pi}(s, \boldsymbol{m}; \boldsymbol{\mu}) \to \boldsymbol{\pi}^*(\boldsymbol{m}; \boldsymbol{\mu})$ due to the ergodicity of the Markov chain. Once we reach the stationary distribution $\boldsymbol{\pi}^*$, the second term of Eq (S.67) becomes $C\hat{\Phi}(\boldsymbol{m}; \boldsymbol{\mu})$, which is independent on time step s (see Eq (11)).

Suppose that the probability distribution is enough close to $\boldsymbol{\pi}^*$ since time step s^*. At time step $s > s^*$, the cumulative expected detoxification efficiency is approximately:

$$\Phi(s; \boldsymbol{m}, \boldsymbol{\mu}) \approx (s - s^* + 1) C\hat{\Phi}(\boldsymbol{m}; \boldsymbol{\mu}) + \Phi(s^* - 1; \boldsymbol{m}, \boldsymbol{\mu}). \tag{S.68}$$

At large s, the second term of Eq (S.68) is small relative to its first term. The optimization problem for the cumulative expected detoxification efficiency then simplifies to maximizing Eq (11). In other words, for a large time step, the optimal values of m_1 and m_2 converge to those which maximize Eq (11) as shown in Fig. 5B.

Next, we relax the three assumptions in Box 1. If a mutation occurs in the transient states, or if a mutation and the intentional introduction of cooperators occur at the same time (relaxing assumption (ii)), the Markov chain does not lose ergodicity (Figure S.10 top left). Although relaxing this assumption allows the population to hold three strategies at once, there exist at most two strategies in the population at an equilibrium state (see Appendix S3 and Appendix S4).

If cooperators cannot invade a population of cooperators with different resistance level (relaxing

assumption (iii)), (a) resident cooperators remain in the population but detoxification efficiency is not optimized in that state, or (b) resident cooperators also go extinct because they cannot persist at (α, T_{in}) which maximizes detoxification efficiency in a mono-culture of invader cooperators (Fig. S.10 top right). In both cases, however, the Markov chains are still ergodic. In case (a), resident cooperators are excluded by cheaters that have the same level of resistance regardless of the values of (α, T_{in}). In case (b), on the other hand, sensitive or resistant cooperators can be introduced with probability m_1 and m_2 when the population size is 0. Indeed, if all microbes go extinct, the Markov chains are always ergodic because the state can transit to a mono-culture of either type of cooperator from this state. Therefore, the optimal $\boldsymbol{m}$ is found in the same way as in the main text.

If sCo and/or rCo cannot exclude the cheaters that differ in the level of resistance (relaxing assumption (iv)), the Markov chains can be non-ergodic. The Markov chain is not ergodic if the initial condition is a mono-culture of one type of cooperators where the population state never goes back once it has changed. When the initial condition is a mono-culture of sCo and sCo cannot exclude rCh, for example, the Markov chain is not ergodic if the state transition cannot occur from a mono-culture of rCo or the coexistence of rCo with sCh to a mono-culture or sCo (Figure S.10 bottom left and right). Note that although cooperators cannot invade the population where the other type of cooperators and the cheaters that differ in the level of resistance coexist, the manual introduction of cooperators can collapse this coexistence by changing (α, T_{in}). In addition, the values of (α, T_{in}) in a mono-culture of sCo should be the same as in the coexistence of sCo with rCh; otherwise, the initial condition is transient and ergodicity is lost.

Now we begin the analysis of non-ergodic Markov chains. When sensitive (and/or resistant) cooperators cannot exclude cheaters with different resistance level by changing the culture conditions, the Markov chains are not always ergodic. Indeed, if the initial condition is a mono-culture of cooperators that cannot exclude cheaters with different level of resistance, and if there is no state transition from another state to the mono-culture of this type of cooperator, the Markov chain is non-ergodic. In this case, the set of states is divided into two subsets S_1 and S_2. Subset S_1 is composed of the equilibrium states where there exist only cooperators that cannot exclude cheaters with different levels of resistance (the values of two parameters, α and T_{in} are different at each state), and the transient states from these equilibrium states to other equilibrium states in subset S_1. Subset S_2, on the other hand, includes the rest of the states within the state space, and the Markov chain is ergodic within S_2. In such cases, the expected cumulative efficiency of detoxification is recursively represented as below:

$$
\begin{aligned}
\Phi(s+1, \boldsymbol{m}; \boldsymbol{\mu}) \;=\; & \Phi(s, \boldsymbol{m}; \boldsymbol{\mu}) \\
& + \prod_{i=0}^{s+1} \{1 - p_{12}(i, \boldsymbol{m}; \boldsymbol{\mu})\} \sum_{k \in S_1} C\phi_k \pi_k (s+1, \boldsymbol{m}; \boldsymbol{\mu}) \\
& + \left[1 - \prod_{i=0}^{s+1} \{1 - p_{12}(i, \boldsymbol{m}; \boldsymbol{\mu})\}\right] \sum_{k \in S_2} C\phi_k \pi_k (s+1, \boldsymbol{m}; \boldsymbol{\mu}),
\end{aligned} \tag{S.69}
$$

where $p_{12}(i)$ represents the state transition probability from a state in S_1 to another in S_2 at $s = i$. Then, the second and third terms of the above equation represent the expected detoxification efficiency when the state of the population is in subset S_1 or S_2, respectively.

To calculate the state distribution in this case, we separate the probability distribution vector into two vectors, $\boldsymbol{\pi}_i = (\pi_j)$ where $j \in S_i$ for $i = 1, 2$. Then, each vector at time step $s+1$ is calculated as below:

$$
\begin{aligned}
\boldsymbol{\pi}_1 (s+1, \boldsymbol{m}; \boldsymbol{\mu}) \;=\; & \boldsymbol{\pi}_1 (s, \boldsymbol{m}; \boldsymbol{\mu}) P^{11}(\boldsymbol{m}; \boldsymbol{\mu}), \tag{S.70a} \\
\boldsymbol{\pi}_2 (s+1, \boldsymbol{m}; \boldsymbol{\mu}) \;=\; & \boldsymbol{\pi}_1 (s, \boldsymbol{m}; \boldsymbol{\mu}) P^{21}(\boldsymbol{m}; \boldsymbol{\mu}) + \boldsymbol{\pi}_2 (s, \boldsymbol{m}; \boldsymbol{\mu}) P^{22}(\boldsymbol{m}; \boldsymbol{\mu}), \tag{S.70b}
\end{aligned}
$$

where matrix P^{ij} represents the transition probability from a state in S_i to another in S_j. Note that $p_{12}(i, \boldsymbol{m}; \boldsymbol{\mu})$ is given by the sum of all elements of $\boldsymbol{\pi}_1 (i, \boldsymbol{m}; \boldsymbol{\mu}) P^{21}(\boldsymbol{m}; \boldsymbol{\mu})$.

As there is no state transition from a state in S_2 to another in S_1, $\boldsymbol{\pi}_1 (s, \boldsymbol{m}; \boldsymbol{\mu}) \to \boldsymbol{0}$ in the limit of $s \to \infty$. Once $\boldsymbol{\pi}_1 (s, \boldsymbol{m}; \boldsymbol{\mu}) P^{21}(\boldsymbol{m}; \boldsymbol{\mu})$ becomes small enough, the state distribution $\boldsymbol{\pi}_2$ is approximated as:

$$
\boldsymbol{\pi}_2 (s+1, \boldsymbol{m}; \boldsymbol{\mu}) \approx \boldsymbol{\pi}_2 (s, \boldsymbol{m}; \boldsymbol{\mu}) P^{22}(\boldsymbol{m}; \boldsymbol{\mu}). \tag{S.71}
$$

In the limit of $s \to \infty$, therefore, $\boldsymbol{\pi}_2$ converges to a unique stationary distribution $\boldsymbol{\pi}_2^*$ as a Markov chain given by subset S_2 is ergodic. In addition, $\prod_{i=0}^{s} \{1 - p_{12}(i, \boldsymbol{m}; \boldsymbol{\mu})\}$ in Eq(S.69) converges to zero with
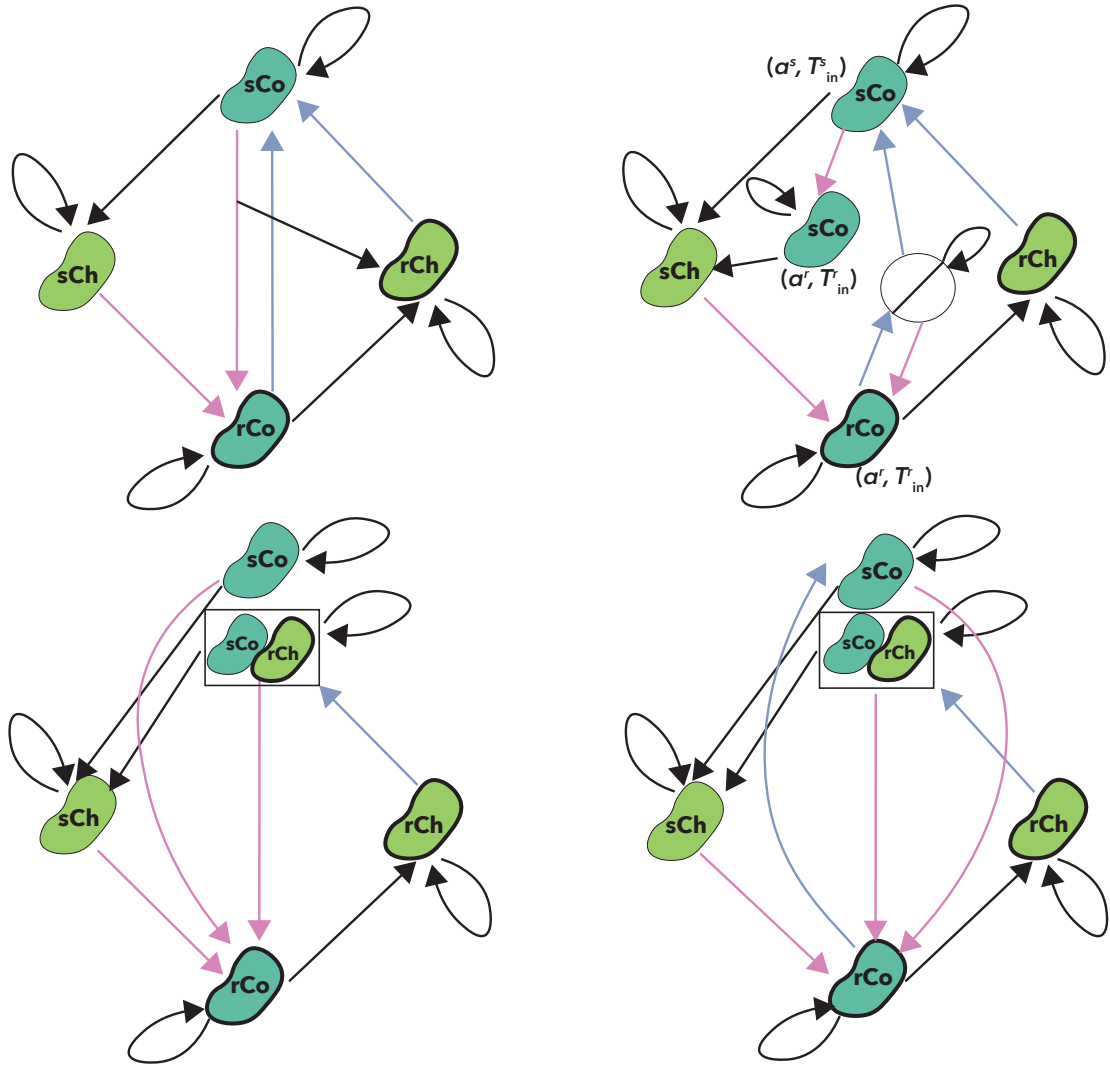
Figure S.10: Diagrams of state transitions

In each panel, the transient states and the states where the manual introduction of cooperators decreases the efficiency of detoxification is ignored. In addition, it is assumed that $\mu_2 = 0$ to reduce the number of nodes. The black lines represent the state transitions caused by mutation while the blue and pink lines represent manual introduction of sCo and rCo, respectively. Top left: when rCo is introduced into a population of sCo, rCo can mutate into rCh, which can add the node to rCh. Top right: if rCo cannot invade sCo, sCo persists in the population. However, the values of $\left(\alpha, T_{\mathrm{in}}\right)$ change and the efficiency of detoxification differs before the invasion. If sCo cannot invade rCo and rCo goes extinct where the efficiency of detoxification is maximized for a population of sCo, the population state transits from rCo to the white, slashed circle where all microbes go extinct. If sCo or rCo are introduced, however, they can grow in the chemostat. Bottom left: if sCo cannot exclude rCh and sCo cannot invade a population of rCo, the Markov chain is not ergodic, when the initial condition is a mono-culture of rCo. Notice, however, that the Markov chain can be regraded as ergodic by ignoring the state of sCo when the initial condition is rCo. Bottom right: Although sCo cannot exclude rCh, the Markov chain can be ergodic when sCo can exclude rCo.
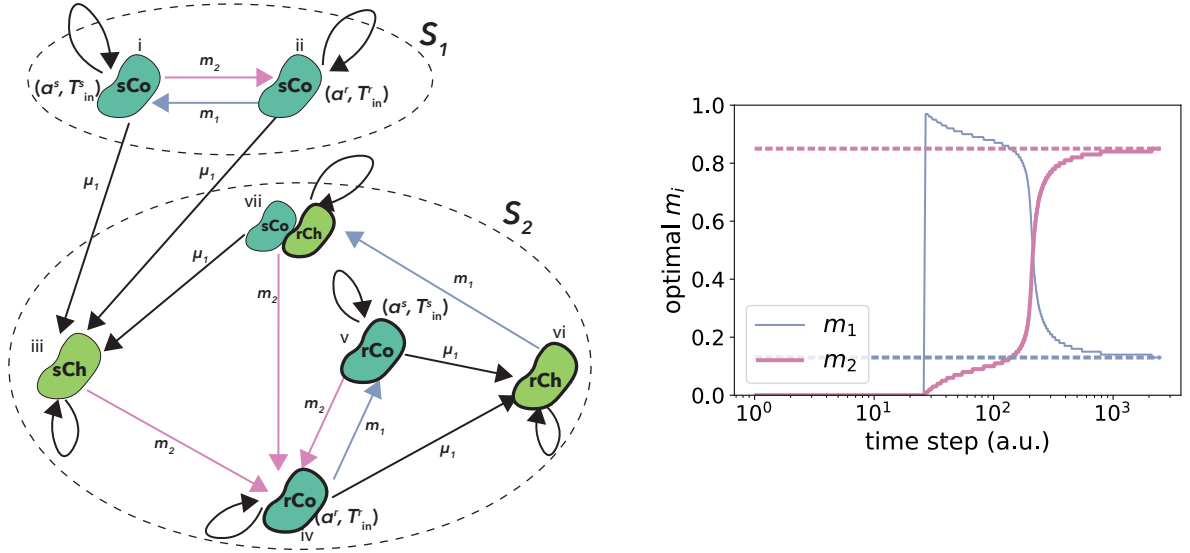
Figure S.11: An example when a Markov chain is not ergodic

Left: the state transition diagram of a non-ergodic Markov chain is shown. In this example, it is assumed that sCo cannot exclude rCh, that $\mu_2 = 0$, and that both types of cooperators cannot invade each other. Therefore, when one type of cooperators is introduced to the population of the other type of cooperators, the resident cooperators remain in the population but the values of α and T_{in} are changed. To reduce the number of states, the transient states are ignored. Notice that there is no state transition from subset S_2 to S_1. In addition, the Markov chain within subset S_2 is ergodic. Each number corresponds to the index of each state. Right: the optimal $\boldsymbol{m}$ (the solid lines) when the simulation finishes at each time step is shown when $\mu_1 = 0.01$ in the left panel. The dashed lines represent $\boldsymbol{m}$ which maximizes $\sum_{k \in S_2} \phi_k \pi_k^* (\boldsymbol{m}; \boldsymbol{\mu})$. The other parameter values are $\boldsymbol{\phi} = \{0.4, 0.2, 0, 0.3, 0.15, 0, 0.2\}$.

large s. Therefore, Eq (S.69) is simplified in the limit of $s \to \infty$ as below:

$$\Phi(s + 1, \boldsymbol{m}; \boldsymbol{\mu}) \approx \Phi(s, \boldsymbol{m}; \boldsymbol{\mu}) + \sum_{k \in S_2} C \phi_k \pi_k^* (\boldsymbol{m}; \boldsymbol{\mu}). \tag{S.72}$$

Thus, the optimal $\boldsymbol{m}$ is approximately obtained by maximizing $\sum_{k \in S_2} \phi_k \pi_k^* (\boldsymbol{m}; \boldsymbol{\mu})$.

Let us show a simple example of a non-ergodic Markov chain. Here, we assume that sCo cannot exclude rCh at any (α, T_{in}). In addition, it is assumed that neither cooperator can invade the population of the other type of cooperator, and that no state exists where all cells go extinct. In this case, the Markov chain is ergodic if the initial state is a mono-culture of sCo. For simplicity, we ignore transient states and assume that $\mu_2 = 0$ as in the main text (Figure S.11 left). At more than $1,000$ time steps, the optimal $\boldsymbol{m}$ converges to the values where $\sum_{k \in S_2} \phi_k \pi_k^* (\boldsymbol{m}; \boldsymbol{\mu})$ is maximized (Figure S.11 right).

References

Murray, J. D. *Mathematical Biology : I . An Introduction*. Springer, New York, 3 edition, 2002. ISBN 0387952233.

Appendix B

Appendices of Chapter 3

The following appendices are also available at `https://doi.org/10.6084/m9.figshare.c.5628917.v1`.

Appendix 1 Details of analysis

Appendix 1.1 Species interaction analysis

Here, we summarize the details of the simulations and parameter values used for species interaction analysis in the main text. In the analysis of species interactions, we used the minimal model ($N = 2$): two species compete for one resource and absorb and are killed by one toxin. We assumed that species 1 grows faster than species 2 but the other parameter values are identical (Table A.1). In each run, the initial condition is either $(r_1(0), t_1(0), s_1(0), s_2(0)) = (150, 100, 10, 0)$ or $(150, 100, 10, 10)$, where species 2 is absent or present, respectively. The initial environmental condition is $\xi = 1$ with a probability of 0.5; otherwise $\xi = -1$. Each simulation continues until at time $\sigma_{end} = 200$ when the distributions of species abundances converge to a quasi-stationary distribution and do not change for a long time. The extinction probability was estimated by running 10^5 simulations for each condition.

In the main text, to investigate how species 2's effect on species 1 changes over the switching rate, we began by analyzing the difference in the extinction probability of species 1 in the presence/absence of species 2 as a proxy for species interactions:

$$\Delta P\left(s_1(\sigma_{end}) = 0\right) \equiv P\left(s_1(\sigma_{end}) = 0; s_2(0) = 0\right) - P\left(s_1(\sigma_{end}) = 0; s_2(0) > 0\right). \tag{A.1}$$

Then, we moved to using the probability of exclusion of the fittest (probability that species 2 excludes species 1) instead:

$$\Delta P\left(s_1(\sigma_{end}) = 0\right) = -\underbrace{P\left(s_1(\sigma_{end}) = 0, s_2(\sigma_{end}) > 0; s_2(0) > 0\right)}_{\text{exclusion of the fittest}}$$
$$+ \left\{ \underbrace{P\left(s_1(\sigma_{end}) = 0; s_2(0) = 0\right)}_{\text{sp 1 goes extinct in mono-culture}} - \underbrace{P\left(s_1(\sigma_{end}) = 0, s_2(\sigma_{end}) = 0; s_2(0) > 0\right)}_{\text{both species go extinct}} \right\}. \tag{A.2}$$

Although Fig.3A and C show that these two measures give similar results, Fig. A.2 confirms this conclusion as the ratio of these two measures is around 1. In other words, the second line on the right-hand-side of Eq (A.2) is ignorable. Intuitively, this is because the environment is very harsh when both species are likely to go extinct (i.e., small resource supplies and high toxin sensitivity); then species 1 is also likely to go extinct in mono-culture under such harsh environment. For this reason, the competitive exclusion probability (the first line on the right-hand-side of Eq (A.2)) changes similarly to difference in extinction probability ($\Delta P\left(s_1\left(\sigma_{end}\right) = 0\right)$) over environmental switching rates and toxin sensitivities (Fig. 3C), although they differ in their signs.

We calculated 95% of highest posterior density intervals (HDPIs) to see the uncertainty of species '1 extinc-

tion probabilities or the probability of exclusion of the fittest. As we were interested in whether species 1 goes extinct (or is excluded by species 2) or not in each simulation run, it is reasonable to assume these probabilities follow beta distributions. As a prior distribution, we assumed the following beta distribution

$$\text{Beta}\,(1,1)\,, \tag{A.3}$$

which is equivalent to a uniform distribution between 0 and 1. After running 10^5 simulations and observing species 1's extinction (or exclusion) X times, the posterior probability distribution of species 1's extinction probability (or the probability of exclusion of the fittest) is given as follows:

$$\text{Beta}\,\left(X + 1, 10^5 - X + 1\right)\,. \tag{A.4}$$

To calculate the 95% HDPIs, we sampled 10,000 samples from the posterior distributions and used pymc3.stats.hpd function. However, the 95% HDPIs are very small and not observable in the main text due to the large number of simulations.

Appendix 1.2   Landscapes of exclusion of the fittest

To understand how the switching rate and toxin sensitivity affect the probability of the exclusion of the fittest, we analyzed three cases in the absence of EFs (Fig. 5), deriving three critical toxin sensitivities where the probability of exclusion of the fittest is maximized. This analysis clarifies what happens at the two extreme switching rates. We see two critical toxin sensitivities (at 0.1 and 0.8 in Figs. 3C and D) at $\nu \to 0$ that correspond to the long time spent with either scarce or abundant resources, while at $\nu \to \infty$, where resources remain at mean abundance, there is a single critical toxin sensitivity (at 0.4 in Figs. 3C and D). As the switching rate increases from one extreme to the other, the form of the probability of exclusion of the fittest in Fig. 3C changes from bi-modal to uni-modal.

We now see that the landscape of competitive exclusion in Fig. 3C contains two "mountain ranges". The first includes two peaks corresponding to the critical toxin sensitivities under scarce and mean resources ($0.1, 0.4$, respectively). The peak at toxin sensitivity 0.1 converges to the peak at 0.4 (Fig. 3D) with increasing environmental switching. The second mountain range has a single peak corresponding to the critical toxin sensitivity under abundant resources (0.8), which vanishes by increasing the switching rate (Fig. 3D). At toxin sensitivities between the critical values under scarce and mean resource supplies ($\delta = 0.2, 0.3$), the probability of exclusion of the fittest changes in a humped shape over the switching rate (species 1's difference in extinction probability changes in a U shape, see Fig. 3B), as it passes over the first mountain range. When the toxin sensitivity is between the critical values under mean and abundant resources (e.g. $\delta = 0.6$), the probability of exclusion of the fittest instead has a "valley" over the switching rate (the difference in extinction probability changes in a humped shape, see Fig. 3B). At toxin sensitivities larger than the abundant resource critical value ($\delta > 0.8$), exclusion of the fittest is very unlikely because both species frequently go extinct (Fig. 5C).

In summary, the transition between the two extreme switching rates results in a highly rugged landscape. This means that the stochastic exclusion of the fittest species can be very high or very low at an intermediate switching rate (either in agreement or in contradiction with the intermediate disturbance hypothesis), depending on species' toxin sensitivities, our proxy for environmental harshness. This makes predicting the outcome at a given switching rate very difficult, as it is dependent on precise parameter values.

Appendix 1.3   Species diversity analysis

In the community diversity analysis, we changed the number of species from $N = 2$ to $N = 10$. Some parameter values were not fixed in this analysis, and we sampled them from the following probability density functions:

$$\mu_{ik} \sim \mathcal{N}\left(1, 0.1^2\right), \tag{A.5a}$$

$$\delta_{jk} \sim \text{Beta}\left(100\bar{\delta}, 100\left(1 - \bar{\delta}\right)\right), \tag{A.5b}$$

$$K_{ik}^r, K_{jk}^t \sim \mathcal{N}\left(100, 10^2\right). \tag{A.5c}$$

Here, each function is uni-modal with a mean of 1.0 for μ_{ik}, $\bar{\delta}$ for δ_{jk}, and 100 for K_{ik}^r, K_{jk}^t. For μ_{ik}, K_{ik}^r, and K_{jk}^t, the mean values are the same as in Table A.1 and they are rarely negative due to the small variances. We set the mean of μ_{ik} as 1 so that μ_{ik} is likely to be larger than δ_{jk} unless $\bar{\delta} = 0.99$: species would easily go extinct when $\mu_{ik} < \delta_{jk}$. The means of K_{ik}^r and K_{jk}^t are chosen so that amounts of resource and/or toxin inflows are larger (smaller) than these values under the abundant (scarce) supply condition. We expected that the growth and/or death rates of species would largely change depending on the environmental conditions with this setting. We sampled δ_{jk} from a beta distribution so that $0 \leq \delta_{jk} \leq 1$. δ_{jk} should be non-negative by definition and not be larger than 1 because a large δ_{jk} is likely to drive species k extinct. Beta distribution satisfies these requirements regardless value of mean $\bar{\delta}$. We systematically vary the value of the mean toxin sensitivity $\bar{\delta}$ to be $\bar{\delta} = 0.1, 0.2, 0.4, 0.6$ or 0.99 in each set of simulations. We used $\bar{\delta} = 0.99$ instead of $\bar{\delta} = 1.0$ because the beta distribution did not generate different values of δ_{jk} with $\bar{\delta} = 1.0$. For each number of species N (2, 4, 6, 8 or 10) and $\bar{\delta}$, 100 sets of parameter values are sampled. With each parameter set, we performed the simulation 100 times until $\sigma_{end} = 200$ at each value of the switching rate ν. Then, we calculated the species richness (the number of surviving species) and beta diversity.

In each run, initial resource amounts, species abundances, and toxin amounts are given by $(r_i(0), t_j(0), s_k(0)) = (150, 150, 10)$ for any i, j, k. As an environmental switching scenario, we chose scenario 1 (Table 1) with $R_i^+ = 200$, $R_i^- = 50$, and $\langle T_k \rangle = 125$ for any i and k. The initial environmental condition is $\xi = 1$ with probability of 0.5; otherwise $\xi = -1$.

Then, at a quasi-stationary distribution (σ_{end}), we evaluated beta diversity and species richness. Beta diversity is calculated as follows:

$$^1D_\beta\left(\sigma_{end}\right) \equiv \frac{^1D_\gamma\left(\sigma_{end}\right)}{^1D_\alpha\left(\sigma_{end}\right)}, \tag{A.6}$$

with alpha and gamma diversities defined as below:

$$^1D_\alpha \left(\sigma_{end} \right) \equiv \exp \left(- \sum_{l=1}^{100} \sum_{k=1}^{N} w_l p_{lk} \left(\sigma_{end} \right) \ln p_{ln} \left(\sigma_{end} \right) \right), \tag{A.7}$$

$$^1D_\gamma \left(\sigma_{end} \right) \equiv \exp \left(- \sum_{k=1}^{N} \bar{p}_k \ln \bar{p}_k \left(\sigma_{end} \right) \right). \tag{A.8}$$

w_l is a weight for community l calculated by size of community l (sum of species abundances in community l relative to the sum of community sizes over l), p_{lk} is the relative abundance of species k in community l (i.e., in community l, $p_{lk} \left(\sigma_{end} \right) = s_k \left(\sigma_{end} \right) / \sum_k s_k \left(\sigma_{end} \right)$), and $\bar{p}_k = \sum_l w_l p_{lk}$ is the mean relative abundance of species k among communities $l = 1, \ldots, 100$. If all species go extinct in community l, it does not affect alpha, beta and gamma diversities as $w_l = 0$. If all species go extinct in all communities, beta diversity becomes $^1D_\beta \left(\sigma_{end} \right) = 1$. Fig. A.3 represents how alpha, beta, and gamma diversities change over the switching rate and mean toxin sensitivities.

Appendix 2   Analysis excluding noise

Appendix 2.1   Deterministic scenarios (absence of DN and EFs)

In this section, we analyse the equilibrium state of a two-species, one-resource, and one-toxin model, in the absence of environmental switching and demographic noise. Although an equilibrium state cannot be analytically obtained, we shall see that there exists at most one stable and feasible equilibrium state. By removing demographic noise and environmental switching from Eq (4) in the main text, the dynamics are governed by the following ordinary differential equations:

$$\dot{r}_1 = \alpha \left(R_1 - r_1 \right) - \sum_{k=1,2} \frac{\mu_{1k}}{Y_{1k}^r} \frac{r_1}{r_1 + K_{1k}^r} s_k, \tag{A.9a}$$

$$\dot{t}_1 = \alpha \left(T_1 - t_1 \right) - \sum_{k=1,2} \frac{\delta_{1k}}{Y_{1k}^t} \frac{t_1}{t_1 + K_{1k}^t} s_k, \tag{A.9b}$$

$$\dot{s}_k = \left(\mu_{1k} \frac{r_1}{r_1 + K_{1k}^r} - \delta_{1k} \frac{t_1}{t_1 + K_{1k}^t} - \alpha \right) s_k, \tag{A.9c}$$

where the dot denotes the time derivative. In spite of the simplicity of the model, we learn from (A.9a) -(A.9c) that, in the absence of consumption by the species, the resources and toxins relax towards the in-flowing values R_1 and T_1 on a timescale of order $1/\alpha$. Furthermore, we infer from (A.9a) -(A.9c) that there are feedback loops between species abundances and resource and toxin concentrations (see Fig. 1A): as s_k increases, both r_1 and t_1 decrease, see (A.9a) and (A.9b), which in turn can result either in a decrease (negative feedback loop) or in an increase of s_k (positive feedback loop), depending on the sign of the parenthesis on the right-hand-side of (A.9c).

When only species k persists in the system, a feasible[1] equilibrium state of Eqs (A.9a) -(A.9c), $(r_1, t_1, s_k) =$

---

[1]Here, feasibility means $r_{1k}^*, t_{1k}^*, s_k^* > 0$

$(r_{1k}^*, t_{1k}^*, s_k^*)$, should satisfy below:

$$\alpha\left(R_1 - r_{1k}^*\right) = \frac{\mu_{1k}}{Y_{1k}^r} \frac{r_{1k}^*}{r_{1k}^* + K_{1k}^r} s_k^*, \tag{A.10a}$$

$$\alpha\left(T_1 - t_{1k}^*\right) = \delta_{1k} Y_{1k}^t \frac{t_{1k}^*}{t_{1k}^* + K_{1k}^t} t_k^*, \tag{A.10b}$$

$$\mu_{1j} \frac{r_{1k}^*}{r_{1k}^* + K_{1k}^r} = \delta_{1k} \frac{t_{1k}^*}{t_{1k}^* + K_{1k}^t} + \alpha. \tag{A.10c}$$

By rearranging the first and second equations, we see that they represent quadratic functions of r_{1k}^* and t_{1k}^*:

$$-r_{1k}^{*2} + \left\{R_1 - K_{1k}^r - \mu_{1k} s_k^*/(\alpha Y_{1k}^r)\right\} r_{1k}^* + K_{1k}^r R_1 = 0, \tag{A.11}$$

$$-t_{1k}^{*2} + \left\{T_1 - K_{1k}^t - \delta_{1k} s_k^*/(\alpha Y_{1k}^t)\right\} t_{1k}^* + K_{1k}^t T_1 = 0. \tag{A.12}$$

Notice that $K_{1k}^r R_1$ and $K_{1k}^t T_1$ are positive. This implies that the above equations always have a unique positive root. In other words, we can obtain unique solutions for r_{1k}^* and t_{1k}^* once we obtain s_k^*.

By substituting the positive roots of Eqs (A.11) and (A.12) into Eq (A.10c), we obtain the following equation whose positive root is s_k^*:

$$f(s) = \frac{1}{2\alpha}\left\{(-\mu_{1k} + 2\alpha + \delta_{1k})s + \sqrt{Q_1(s)} - \sqrt{Q_2(s)} + \alpha\left(-K_{1k}^r Y_{1k}^r + K_{1k}^t Y_{1k}^t - Y_{1k}^r R_1 + Y_{1k}^t T_1\right)\right\} \tag{A.13}$$

where $Q_1(s)$ and $Q_2(s)$ are quadratic functions of s:

$$Q_1(s) = \left\{\mu_{1k}s + \alpha Y_{1k}^r\left(K_{1k}^r - R_1\right)\right\}^2 + 4\alpha Y_{1k}^r K_{1k}^r R_1 > 0 \tag{A.14a}$$

$$Q_2(s) = \left\{\delta_{1k}s + \alpha Y_{1k}^t\left(K_{1k}^t - T_1\right)\right\}^2 + 4\alpha Y_{1k}^t K_{1k}^t T_1 > 0. \tag{A.14b}$$

Notice that $f(s)$ always has a root $s = 0$ because

$$\begin{aligned}
f(0) &= \frac{1}{2\alpha}\left\{\sqrt{Q_1(0)} - \sqrt{Q_2(0)} + \alpha\left(-K_{1k}^r Y_{1k}^r + K_{1k}^t Y_{1k}^t - Y_{1k}^r R_1 + Y_{1k}^t T_1\right)\right\} \\
&= \frac{1}{2\alpha}\left\{\alpha Y_{1k}^r\left(K_{1k}^r + R_1\right) - \alpha Y_{1k}^t\left(K_{1k}^t + T_1\right) + \alpha\left(-K_{1k}^r Y_{1k}^r + K_{1k}^t Y_{1k}^t - Y_{1k}^r T_1 + Y_{1k}^t T_1\right)\right\} \\
&= 0.
\end{aligned} \tag{A.15}$$

Although Newton's method numerically provides a root of $f(s)$, this root depends on the initial value used in Newton's method. In addition, as $f(s)$ has root $s = 0$, Newton's method may provide this root with various initial values, which does not always mean that $f(s)$ does not have positive roots (i.e., s_k^*). In other words, it is recommended to investigate how many positive roots $f(s)$ has before using Newton's method.

To investigate the number of $f(s)$'s positive roots, it is useful to obtain df/ds:

$$\frac{df}{ds} = \frac{1}{2\alpha}\left\{(-\mu_{1k} + 2\alpha + \delta_{1k}) + \frac{dQ_1/ds}{2\sqrt{Q_1(s)}} - \frac{dQ_2/ds}{2\sqrt{Q_2(s)}}\right\}. \tag{A.16}$$

Although it is analytically difficult to obtain the solution(s) of df/ds, we can obtain the maximum number of positive roots of $f(s)$ by analyzing the number of df/ds's sign changes. Notice that dQ_1/ds and dQ_2/ds are linear functions of s:

$$\frac{dQ_1}{ds} = 2\mu_{1k}\left\{\mu_{1k}s + \alpha Y_{1k}^r\left(K_{1k}^r - R_1\right)\right\}, \tag{A.17a}$$

$$\frac{dQ_2}{ds} = 2\delta_{1k}\left\{\delta_{1k}s + \alpha Y_{1k}^t\left(K_{1k}^t - T_1\right)\right\}. \tag{A.17b}$$

In addition, $Q_1(s)$ and $Q_2(s)$ are always positive. Then, the second and third terms of Eq (A.16) change their sign at most once by increasing s. The maximum number of df/ds's sign change is, therefore, two. This implies that the maximum number of positive roots of $f(s)$ is also two. To obtain the exact number of positive roots of $f(s)$, it is necessary to numerically calculate the root(s) of df/ds. Substituting the root(s) into Eq (A.13) and calculating the sign of $f(s)$, it is possible to obtain the exact value of s_k^*.

Once we have obtained a feasible equilibrium state $(r_{1k}^*, s_k^*, t_{1k}^*)$, it is necessary to analyze the stability of this equilibrium state. Although the stability analysis requires the evaluation of the Jacobian matrix at each equilibrium state, we can see that there exists at most one feasible and stable equilibrium state without such a stability analysis. Notice that:

$$\dot{s}_j \lessgtr 0$$

$$\Leftrightarrow \mu_{1k}\frac{r_{1k}^*}{r_{1k}^* + K_{1k}^r} \lessgtr \delta_{1k}\frac{t_{1k}^*}{t_{1k}^* + K_{1k}^t} + \alpha$$

$$\Leftrightarrow f\left(s\right) \gtrless 0. \tag{A.18}$$

These inequalities imply that a stable equilibrium state satisfies the following inequality:

$$\left.\frac{df}{ds}\right|_{s=s_k^*} > 0. \tag{A.19}$$

Although there can exist at most two feasible equilibria, only one of them satisfies the above inequality (Fig. A.1). The number of feasible and stable equilibria is, therefore, one at most.

For the sake of simplicity but without loss of generality, in the main text, we assumed that the maximum growth rate of species 1 is larger than species 2 but the remaining parameter values are identical, resulting in a *per-capita* growth rate of species 1 always larger than species 2. In this setting, in the absence of DN, species 2 always goes extinct after a finite time while species 1 persists if it has a feasible and stable equilibrium state. It is worth noting that this feature also characterizes the two-species, one-resource and one-toxin model in the presence of environmental switching without DN: in this case species 2 always goes extinct in a finite time, and at equilibrium one has either $s_1^* = s_2^* = 0$ (extinction of both species) or $s_1^* > 0, s_2^* = 0$ (survival of species 1, extinction of species 2), see Fig. A.5.

Appendix 2.2   With environmental fluctuations alone

It is also instructive to analyze the long-time dynamics of the two-species, one-resource, one-toxin model in the presence of EFs and without DN. This describes the dynamics of a sufficiently large community in which DN is always negligible and whose time evolution is described by Eqs (A.9a)-(A.9c), but now with $(R_1, T_1) = (R_1(\xi), T_1(\xi))$ switching randomly with rate ν between two states $\xi = \pm 1$, which is applicable to all three scenarios in Table 1. Here, for the sake of simplicity and concreteness, we consider the scenario 1 of Table 1 with $1 \ll R_1^- < \langle R_1 \rangle < R_1^+$, while $T_1 = \langle T_1 \rangle$ does not vary with the environment. In this case, Eqs. (A.9a)-(A.9c) are a system of stochastic differential equations (SDEs). In order to appreciate the effect of EFs on DN, it is useful to study the total population size $n \equiv r_1 + t_1 + s_1 + s_2$ which, according to (A.9a)-(A.9c), obeys

$$\dot{n} = \alpha \left(R_1(\xi) + \langle T_1 \rangle - n \right) - 2 \sum_{k=1,2} \delta_{1,k} \, \frac{t_1 s_k}{t_1 + K_{1,k}^t}, \tag{A.20}$$

where, as in the main text, $Y_{ik}^r = Y_{jk}^t = 1$. Importantly, $\sqrt{n}$ gives the intensity of DN in a total population of size n. Eq (A.20) is, however, difficult to analyze, due to the nonlinear coupling of t_1 with s_k whose dynamics, according to Eqs (A.9b)-(A.9c), also depend on ξ. In order to obtain some insight into the stationary probability density $p(n)$ of the population, we have thus we analyzed $\hat{n} \equiv r_1 - t_1 + s_1 + s_2$ which is related to the total population size, since $\hat{n} = n - 2t_1$. In this simple example $\hat{n}$, according to Eqs. (A.9a)-(A.9c), obeys the following linear SDE

$$\dot{\hat{n}} = \alpha \left(R_1(\xi) - \langle T_1 \rangle - \hat{n} \right). \tag{A.21}$$

As seen above, the system reaches its equilibrium and $(s_1, s_2, t_1) \to (s_1^*, s_2^* = 0, t_1^*)$ after a finite time while the environment varies endlessly .The SDE (A.21) defines a simple piecewise deterministic Markov process (PDMP) (Davis, 1984) whose (marginal) probability density $p_{\nu/\alpha}(\hat{n})$ can be obtained analytically, see Bena (2006) and Horsthemke and Lefever (2006):

$$q_{\nu/\alpha}(\hat{n}) = \mathcal{Z} \left[\left\{ R_1^+ - \langle T_1 \rangle - \hat{n} \right\} \left\{ \hat{n} - R_1^- + \langle T_1 \rangle \right\} \right]^{\frac{\nu}{\alpha} - 1}, \tag{A.22}$$

where $\mathcal{Z}$ is the normalization constant. Hence, in the absence of DN, with environmental switching as the sole source of randomness, the dynamics leads to a population consisting of only individuals of species 1 (species 2 is wiped out) and toxin and resources, with respective abundances s_1, r_1, t_1, in a fluctuating population whose scaled size $\hat{n} = s_1 + r_1 - t_1$ is distributed according to $q_{\nu/\alpha}(\hat{n})$ of finite support $\left[R_1^- - \langle T_1 \rangle, R_1^+ - \langle T_1 \rangle \right]$ (i.e. $R_1^- - \langle T_1 \rangle \le n \le R_1^+ - \langle T_1 \rangle$). While this PDMP probability density ignores DN, it is known to generally provide a useful approximate description of how the quasi-stationary (at finite time $\gtrsim \sigma_{end}$) distribution varies with the environment in a large yet finite system, see, *e.g.*, (Wienand et al., 2017, 2018; West and Mobilia, 2020;

Taitelbaum et al., 2020). Here, Eq. (A.22) readily sheds light on the effect of ν and α on $\hat{n} = n - 2t_1$. It is clear from Eqs. (A.9a)-(A.9c), that t_1 varies with ξ and thus the marginal probability density of population size, $p(n)$, cannot be immediately obtained from (A.22). Yet, we can obtain useful information about how the population varies with ν and α by focusing on the fast and slow switching regimes (see Fig. A.6):

- when $\nu/\alpha \ll 1$ (slow varying environment), $q_{\nu/\alpha}(\hat{n})$ is bimodal and $\hat{n}$ is as likely to fluctuate about $R_1^+ - \langle T_1 \rangle$ or $R_1^- - \langle T_1 \rangle$. Hence, when in this slowly switching regime, the population size fluctuates with the same probability either about $R_1^- - \langle T_1 \rangle + 2t_1^*(\xi = -1)$ or $R_1^+ - \langle T_1 \rangle + 2t_1^*(\xi = +1)$ (see, the first column of Fig. A.6).

- when $\nu/\alpha \gg 1$ (fast changing environment), $q_{\nu/\alpha}(\hat{n})$ is unimodal and $\hat{n}$ fluctuates about its average, i.e. $\hat{n} \approx \langle R_1 \rangle - \langle T_1 \rangle$. Hence, in this fast switching regime, the population size fluctuates about $n \approx \hat{n} + 2\langle t_1 \rangle = \langle R_1 \rangle - \langle T_1 \rangle + 2\langle t_1 \rangle$, where $\langle t_1 \rangle = (t_1^*(\xi = 1) + t_1^*(\xi = -1))/2$ and $t_1^*(\xi = \pm 1)$ are obtained from the equilibria of Eqs. (A.9a)-(A.9c) with $R_1 = R_1(\xi = \pm 1)$ (see, the third column of Fig. A.6).

This picture is confirmed by the simulation results reported in Fig. A.6, where we see that the marginal probability densities $p(n)$ and $q_{\nu/\alpha}(\hat{n})$ have qualitatively the same features: both are bimodal and have two well-separated sharp peaks when $\nu/\alpha \ll 1$, and a single pronounced peak when $\nu/\alpha \gg 1$. At intermediate values of ν/α, the probability densities of n and $\hat{n}$ are much broader with a generally flat profile (exhibiting one or two "bumps"), see the second column of Fig. A.6. This analysis, which can be readily extended to $N > 2$ species and to other scenarios of environmental fluctuations, clearly shows that the population size can greatly vary as the environment changes. In particular, Eq (A.22) shows that when ν is low or α is high, half of the simulation runs lead to communities of "small sizes" where the effect of DN is expected to be significantly larger than in communities obtained in the faster switching regime ($\nu/\alpha \gg 1$) when $R_1^+ \gg R_1^-$.

It is also worth noting from Eqs (A.9a)-(A.9c) and Eq (A.21) that if, say, only the maximum growth rates were subject to environmental switching, i.e. $\mu_{ik} = \mu_{ik}(\xi)$ with all other parameters kept constant, we always obtain a large constant population size (if $R_1 \gg T_1$) and thus no DN-EFs coupling because the distributions of n and $\hat{n}$ are independent on μ_{ik}. On the other hand, if only the maximum death rates were subject to environmental switching, i.e. $\delta_{jk} = \delta_{jk}(\xi)$, the DN-EFs coupling would result from a complicated set of coupled stochastic differential equations obtained from (A.9a)-(A.9c) with $\delta_{jk} \to \delta_{jk}(\xi)$. These analyses clarify the significant difference of our model from others. Some previous studies (e.g., Leigh (1981); Kalyuzhny et al. (2015) and multi-species model of Engen and Lande (1996)) do not include DN-EFs coupling because they assume that EFs affect species' growth rates, but total species abundances or maximum population sizes do not change. Other studies (e.g., Kamenev et al. (2008); Chisholm et al. (2014); Fung et al. (2015) and the single species model of Engen and Lande (1996)) include a form of DN-EF coupling because EFs in their model change both species' growth rates and population sizes. However, these models consist of only one species, and hence do not consider interspecific interactions. Our model includes both DN-EFs and indirect species interactions (resource competition and facilitation via detoxification, see Fig. 1A), and thus we can analyze how DN-EFs

coupling affect species diversity as in Fig. 6. In summary, here we have attempted to make the simplest choice to couple DN and EF in a transparent and biologically-relevant way.

Appendix 3   Alternative environmental switching scenarios

In the main text, environmental switching affects only the resource supply (scenario 1) while the amount of toxin supply is fixed. Here, the results of other environmental switching scenarios are shown: In scenario 2, environmental switching affects only the toxin supply, while both resource and toxin supplies change and correlate negatively in scenario 3 (see Table 1).

In both scenarios 2 and 3, the difference in species 1's extinction probability $\Delta P\left(s_1(\sigma_{end}) = 0\right)$ is very similar to the negative value of the probability of exclusion of the fittest when the sign of $\Delta P\left(s_1(\sigma_{end}) = 0\right)$ is negative (Fig. A.7). This once again confirms that we can use one measure for the other.

In addition, in both scenarios, the probability of exclusion of the fittest are bi-modal across toxin sensitivities at very slow environmental switching $\nu = 10^{-5}$, but uni-modal at very fast environmental switching ($\nu = 10^3$) (Fig. A.7C, D). As explained in the main text, when $\nu \to 0$, there are no switches and the environmental state is randomly allocated to harsh or mild conditions at $t = 0$ with the same probability (the mean of ξ is zero), yielding a bi-modal distribution at low switching rate. In the limit $\nu \to \infty$, there are so many switches that environmental noise averages out, i.e. ξ is replaced by its mean (that is zero). This results in a uni-modal distribution when $\nu \gg 1$.

In scenario 2, however, we do not observe any non-monotonic changes over the switching rate (Fig. A.7C). This is because the critical toxin sensitivity under abundant toxin supply ($\delta = 0.3$) is close to that under mean toxin supply ($\delta = 0.4$). In contrast, we do observe non-monotonic effects of the switching rate in scenario 3 (Fig. A.7D): when $\delta = 0.2$, an intermediate switching rate ($\nu = 10^{-1}$) shows the minimum difference in extinction probability, and the maximum probability of competitive exclusion. Although in scenario 3 the same intermediate switching rate minimizes the probability of competitive exclusion at toxin sensitivity 0.6, the same non-monotonic effect is not observed in the difference in extinction probabilities (Fig. A.7B). Note that the critical toxin sensitivities under the mild environments in scenarios 2 (scarce toxin supply) and 3 (abundant resource supply and scarce toxin supply) are slightly larger than 1.0 (Fig. A.8) and therefore not visible in Fig. A.7D. Table A.2 shows the critical toxin sensitivities in each scenario.

Appendix 4   Effects of resource supply

In this section, we again focus on environmental switching scenario 1 (changing only resource supply). We will see that the amount of resource supplies $R_1(\xi)$ changes the critical toxin sensitivities under scarce, mean, and abundant, resource supplies, affecting the likelihood of the non-monotonic effect of the environmental switching rate.

By increasing the abundant or decreasing the scarce resource supply (Figs. A.9A and D, respectively), the

distance between the critical toxin sensitivities under scarce and mean (or mean and abundant) resource supplies becomes larger (Table A.2). Conversely, decreasing abundant or increasing scarce resource supply (Figs. A.9B and C, respectively), decreases the distance between the critical toxin sensitivities under scarce and mean (or mean and abundant) resource supplies (Table A.2). Once again, changes in competitive exclusion probability (A.9) match changes in species 2's effect on species 1 (Fig. A.10).

Analyzing the probability of exclusion of the fittest instead of the difference in extinction probability is valid only when the difference in extinction probability is negative: if it is positive, the second line in (A.2) cannot be ignored. The sign of the difference can, however, become positive at $\delta = 1.0$. Indeed, when $R_1^+ = 400$ and $\delta = 1.0$, species 2 has a positive effect on species 1 whose strength varies non-monotonically with the rate of environmental switching (A.11A). In this case, we analyze both (i) the probability of exclusion of the fittest (the first line in Eq (A.2), Fig. A.11B) and (ii) the difference in species 1's extinction probability in mono-culture and both species extinction in co-culture (the second line in Eq (A.2), Fig. A.11C). The effects of the environmental switching rate on (i) and (ii) are similar, leading to similar non-monotonic effects of species 2 on species 1. In sum, non-monotonic effects of environmental switching rates on species interactions can be observed whether these interactions are positive or negative. Although the main text explains why non-monotonic effects of environmental switching rates on species interactions happen when interactions are negative, it remains unclear why such non-monotonic changes happen when species interactions are positive.

Appendix 5   Other forms of environmental fluctuations

In this section, we analyze environmental fluctuations other than symmetric switching between two states. Our goal is here to show that our main findings qualitatively still hold and can therefore traced back to the generic interdependence of EFs and DN rather than detail of their coupling.

As in the main text, we assume that the environmental fluctuations change only the resource supply while the toxin supply is constant. First, we investigate asymmetric switching between two resource supply conditions. Second, we increase the number of environmental states and introduce a cyclic change of the resource supply.

Under asymmetrically switching or cyclically fluctuating environments, we find similar patterns of how species interactions change over species' toxin sensitivities and a rate of environmental fluctuations.

Appendix 5.1   Asymmetric switching

In the main text and appendices other than this section, for the sake of simplicity, we assume symmetric switching rates between two states by Eq. (3) (see Taitelbaum et al. (2020)). Here, we relax this assumption and introduce asymmetric switching rates because perfectly symmetric switching environments are very unlikely in nature; in gut microbiota, for example, duration that their host is starving would be longer than that the

host is eating food. We implemented asymmetric switching as follows:

$$\xi = 1 \xrightarrow{\nu_1} \xi = -1 \tag{A.23a}$$

$$\xi = -1 \xrightarrow{\nu_2} \xi = 1. \tag{A.23b}$$

Without loss of generality, we define the two switching rates as follows:

$$\nu_1 = \beta_1 \nu \tag{A.24a}$$

$$\nu_2 = \beta_2 \nu \tag{A.24b}$$

where ν is the basal switching rates. In extreme cases $(\beta_1 \gg \beta_2)$, for asymmetric switching scenarios can correspond to systems with rare disturbances in nature. In this extremely asymmetric case, the sojourn time in the harsh environment exceeds greatly that in the mild environment, which results in a strong effect of DN. In contrast, DN effects are less important when $\beta_2 \gg \beta_1$ and the population experiences more frequently the mild than the harsh environment(Taitelbaum et al., 2020). We recover a symmetric environmental switching when $\beta_1 = \beta_2$.

Fig. A.12 summarizes the two species interactions under resource supply fluctuations with asymmetric environmental switching rates when the initial environmental condition is $\xi(0) = 1$ with probability of 0.5 (otherwise $\xi(0) = -1$). In Figs. A.12 A and C, $\nu_1 > \nu_2$ and therefore the sojourn time of $\xi = -1$ (an harsh environment) is longer than that of $\xi = 1$ (a mild environment). On the other hand, Figs. A.12 B and D shows the cases when the sojourn time the mild environment is longer than that of the harsh environment because $\nu_1 < \nu_2$. In both cases, species 1's difference in extinction probabilities and the probability of exclusion of the fittest show monotonically increasing, monotonically decreasing, or non-monotonic changing with a minimum or maximum value at an intermediate switching rate, although they quantitatively differ from Figs. 3 A and C.

Appendix 5.2 Cyclic changes

Here, we analyze cases when the number of environmental states in terms of resource supply is greater than two, but remains discrete and finite. This simply reflects that natural environments do not always fluctuate between two states. In this subsection, an environmental state is given by $\xi = 1, 2, \ldots, n$ and environments cyclically fluctuate with rate ν as follows:

$$\xi \xrightarrow{\nu} \begin{cases} \xi + 1 & \text{if } \xi = 1, \ldots, n-1 \\ 1 & \text{otherwise.} \end{cases} \tag{A.25}$$

This is a natural extension of Eq (3) by increasing the number of environmental conditions: $n = 2$ recovers a symmetrically switching environment between two conditions ($\xi = 1 \to 2 \to 1 \to \ldots$ although we use the notation $\xi = \pm 1 \to \mp 1 \to \pm 1 \to \ldots$ in the main text).

Fig. A.13 shows how species interactions between two species change when $n = 4$ and the resource supply fluctuates such that $R_1 (\xi = 1) > R_1 (\xi = 2) > R_1 (\xi = 3) > R_1 (\xi = 4)$. In this analysis, an initial environmental condition $\xi(0)$ is one of four conditions (probability of 0.25 for each). As in Figs. 3 A and C, the rate of cyclical environmental change affects species 1's extinction probability and the probability that species 2 excludes species 1. We frequently observe that the probability of exclusion of the fittest non-monotonically changes (toxin sensitivity: $0.1 - 0.6$).

In this work, the fluctuating environment has been modeled as randomly switching between a finite number of environmental states ξ. This choice is particularly convenient as it allows us to deal with bounded noise, and hence R_i and T_j to always remain positive, and are straightforward to simulate using the standard Gillespie algorithm. The case of environmental noise varying continuously in time is also of great interest, as it allows R_i and T_j to take any values in a domain. For instance, the environmental noise can be an Ornstein-Uhlenbeck process (ξ_{OU}), e.g. by letting $R_i = R_i(\xi_{OU}) = \bar{R}_i(1 + k\xi_{OU})$, where $\bar{R}_i$ is constant, ξ varies in time, and $k > 0$, see, e.g., Assaf et al. (2013). This poses a number of challenges since, ξ_{OU} being unbounded, R_i can take negative (unphysical) values. Furthermore, there are no general methods to simulate exactly birth-death processes subject to continuous external noise, see, e.g., Berríos-Caro and Galla (2020). Here, while the assumption of discrete environmental noise is a simplification of many real situations, we think that our main findings are generic and shall hold also under continuous environmental noise. In fact, since our results stem chiefly from the coupling of EFs and DN, a feature shared by discrete and continuous noise, they are expected to qualitatively hold also in the case of continuous external noise.

Appendix 6   Diversity in communities of increasing species number

In the main text, we show the distributions of beta diversity and species richness in communities when the initial number of species N is two or ten. This section shows the results of intermediates values of $N = 4, 6, 8$. The number of initial species does not change how beta diversity changes over the environmental switching rate, except when the mean toxin sensitivity $\bar{\delta} = 0.4$ (Fig. A.18). These results indicate that beta diversity and the probability of competitive exclusion change similarly over the switching rate when the initial number of species is larger than two.

The initial number of species N in a community affects the maximum values of species richness (Fig. A.19). However, how the switching rate and the mean toxin sensitivity affect the distribution of species richness was consistent for different values of N. In particular, increasing the mean toxin sensitivity decreases species richness. The effects of the environmental switching rate also depend on the mean toxin sensitivity: at mean toxin sensitivity 1.0, species in all cases are more likely to go extinct as the switching rate increases, while the likelihood of all species going extinct consistently shows a humped shape at toxin sensitivity 0.4 or 0.6.

Appendix 7 Quantified the similarity between the exclusion of fittest and beta diversity

In this section, we first quantify the similarity between the exclusion of the fittest and beta diversity in two-species communities. Then, we investigate how many species pairs we should analyze to predict the patterns of beta diversity in ten-species communities.

In the main text, we show that the probability of exclusion of the fittest and beta diversity exhibit similar patterns (see columns A and B in Fig. 6). We quantified the similarities with Pearson's correlation coefficients. Fig. A.15 shows the distributions of the correlation coefficients of 100 two-species communities at each mean toxin sensitivity (see also Appendix 1.3). Except for the case that mean toxin sensitivity is 0.4 (where the probability of exclusion of the fittest and beta diversity do not match), the correlation coefficients are large positive (> 0.6). Therefore, a large correlation coefficients indicates the similarity between the probability of the exclusion of the fittest and beta diversity.

We continued the analysis in ten-species communities. In these cases, we have 45 pairs of species in each community and we calculated the probabilities of the exclusion of the fittest by running 100 replicates in each species pair of thirty ten-species communities (six communities at five mean toxin sensitivities). Fig. A.16 shows that the exclusion of the fittest in some pairs match the patterns of beta diversity of whole communities but other pairs do not. Then, we investigated the number of species pairs m that is necessary to predict a pattern of beta diversity over the environmental switching rate (i.e., a large correlation between probability of exclusion of the fittest and beta diversity). When $m \geq 2$, we calculated Pearson's correlation coefficient between beta diversity and mean probability of exclusion of the fittest within m pairs. However, we have many possible choices of m pairs when $1 < m < 45$. In these cases, we randomly chose 300 sets of m pairs in each of community and thus we obtained 1800 correlation coefficients at each mean toxin sensitivities. When $m = 1$ or 45, we analyzed all possible choices of m species pairs and we obtained 45 or 1 correlation coefficient(s) in each community, respectively. Fig. A.17 indicates that the correlation coefficients can be large even when $m = 1$ or 2, but larger m increases correlation coefficients. We suggest that $m = 5$ is the best because the $m = 5$ and $m = 9$ show little difference in the distributions of the correlation coefficients and we are very likely to obtain a large correlation coefficient with $m = 5$. In conclusion, we do not have to analyze the exclusion of the fittest for all species pairs to predict how the environmental switching rate affect beta diversity.

References

Assaf, M., Mobilia, M., and Roberts, E. Cooperation dilemma in finite populations under fluctuating environments. Physical Review Letters, 111(23):1–7, 2013. ISSN 00319007. doi: 10.1103/PhysRevLett.111.238101.

Bena, I. Dichotomous Markov noise: Exact results for out-of-equilibrium systems. A review. International Journal of Modern Physics B, 20(20):2825–2888, 6 2006. ISSN 02179792. doi: 10.1142/S0217979206034881. URL http://arxiv.org/abs/cond-mat/0606116.

Berríos-Caro, E. and Galla, T. Beyond the adiabatic limit in systems with fast environments: a τ-leaping algorithm. arXiv, pages 1–22, 2020. URL http://arxiv.org/abs/2011.10748.

Chisholm, R. A., Condit, R., Rahman, K. A., Baker, P. J., Bunyavejchewin, S., Chen, Y.-Y., Chuyong, G., Dattaraja, H. S., Davies, S., Ewango, C. E. N., Gunatilleke, C. V. S., Nimal Gunatilleke, I. A. U., Hubbell, S., Kenfack, D., Kiratiprayoon, S., Lin, Y., Makana, J.-R., Pongpattananurak, N., Pulla, S., Punchi-Manage, R., Sukumar, R., Su, S.-H., Sun, I.-F., Suresh, H. S., Tan, S., Thomas, D., and Yap, S. Temporal variability of forest communities: empirical estimates of population change in 4000 tree species. Ecology Letters, 17(7): 855–865, 7 2014. ISSN 1461023X. doi: 10.1111/ele.12296. URL http://doi.wiley.com/10.1111/ele.12296.

Davis, M. H. A. Piecewise-Deterministic Markov Processes: A General Class of Non-Diffusion Stochastic Models. Journal of the Royal Statistical Society: Series B (Methodological), 46(3):353–376, 1984. doi: 10.1111/j.2517-6161.1984.tb01308.x.

Engen, S. and Lande, R. Population Dynamic Models Generating Species Abundance Distributions of the Gamma Type. Journal of Theoretical Biology, 178(3):325–331, 2 1996. ISSN 00225193. doi: 10.1006/jtbi. 1996.0028. URL https://linkinghub.elsevier.com/retrieve/pii/S0022519396900284.

Fung, T., Villain, L., and Chisholm, R. A. Analytical formulae for computing dominance from species-abundance distributions. Journal of Theoretical Biology, 386:147–158, 2015. ISSN 10958541. doi: 10.1016/j.jtbi.2015. 09.011. URL http://dx.doi.org/10.1016/j.jtbi.2015.09.011.

Horsthemke, W. and Lefever, R. Noise-Induced Transitions, volume 15 of Springer Series in Synergetics. Springer Berlin Heidelberg, 2nd edition, 4 2006. ISBN 978-3-540-11359-1. doi: 10.1007/3-540-36852-3. URL http: //link.springer.com/10.1007/3-540-36852-3.

Kalyuzhny, M., Kadmon, R., and Shnerb, N. M. A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities. Ecology Letters, 18(6):572–580, 6 2015. ISSN 1461023X. doi: 10.1111/ele.12439. URL http://doi.wiley.com/10.1111/ele.12439.

Kamenev, A., Meerson, B., and Shklovskii, B. How Colored Environmental Noise Affects Population Extinction. Physical Review Letters, 101(26):268103, 12 2008. ISSN 0031-9007. doi: 10.1103/PhysRevLett.101.268103. URL https://link.aps.org/doi/10.1103/PhysRevLett.101.268103.

Leigh, E. G. The average lifetime of a population in a varying environment. Journal of Theoretical Biology, 90(2):213–239, 5 1981. ISSN 00225193. doi: 10.1016/0022-5193(81)90044-8. URL https://linkinghub. elsevier.com/retrieve/pii/0022519381900448.

Taitelbaum, A., West, R., Assaf, M., and Mobilia, M. Population Dynamics in a Changing Environment: Random versus Periodic Switching. Physical Review Letters, 125(4):048105, 7 2020. ISSN 0031-9007. doi: 10.1103/PhysRevLett.125.048105. URL https://link.aps.org/doi/10.1103/PhysRevLett.125.048105.

West, R. and Mobilia, M. Fixation properties of rock-paper-scissors games in fluctuating populations. Journal of Theoretical Biology, 491:110135, 4 2020. ISSN 00225193. doi: 10.1016/j.jtbi.2019.110135. URL http://dx.doi.org/10.1016/j.jtbi.2019.110135.

Wienand, K., Frey, E., and Mobilia, M. Evolution of a Fluctuating Population in a Randomly Switching Environment. Physical Review Letters, 119(15):158301, 10 2017. ISSN 0031-9007. doi: 10.1103/PhysRevLett.119.158301. URL https://link.aps.org/doi/10.1103/PhysRevLett.119.158301.

Wienand, K., Frey, E., and Mobilia, M. Eco-evolutionary dynamics of a population with randomly switching carrying capacity. Journal of the Royal Society Interface, 15(145):20180343, 8 2018. ISSN 17425662. doi: 10.1098/rsif.2018.0343. URL https://royalsocietypublishing.org/doi/10.1098/rsif.2018.0343.

Figure A.1: Feasible and stable equilibrium state

An example of $f(s)$ in Eq (A.13) and equilibrium states. In this example, $f(s)$ has three roots: $s = 0$, and two feasible equilibria. The blue arrows indicate that s decreases or increases when $f(s)$ is positive or negative, respectively. As df/ds is negative at the left equilibrium state, this equilibrium state is unstable. On the other hand, the right feasible equilibrium has a positive df/ds and thus this equilibrium can be stable. Note that the equilibrium state corresponding to $s = 0$ can be also stable. Parameter values are $\delta_{1k} = 1.2$, $R_1 = 200$, $T_1 = 125$ and as in Table A.1 otherwise.
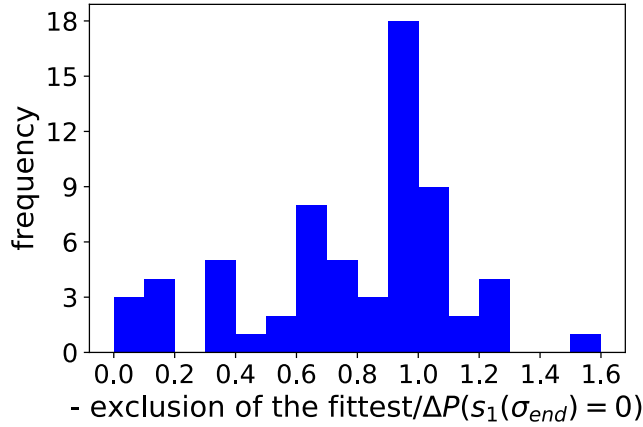
Figure A.2: Probability of exclusion of the fittest relative to the difference in extinction probability of species 1

A histogram of the ratio of the probability of exclusion of the fittest and the difference in species 1's extinction probabilities alone versus in the presence of species 2, showing the results of 81 sets of the environmental switching rate ν and the toxin sensitivity δ (9 values of $\nu = 10^{-5}, \ldots, 10^3$ and 9 values of $\delta = 0.1, \ldots, 0.9$). For each set of the parameter values, 10^5 simulations were run to calculate the competitive exclusion probability and the difference in species 1's extinction probabilities in the presence/absence of species 2. In many of these 81 parameter sets, this ratio is close to 1, indicating that both measures yield similar results. As in the manuscript we focus on conditions leading to competition between the two species, we ignore toxin sensitivity $\delta = 1.0$ where species 2's effect on species 1 can be positive.
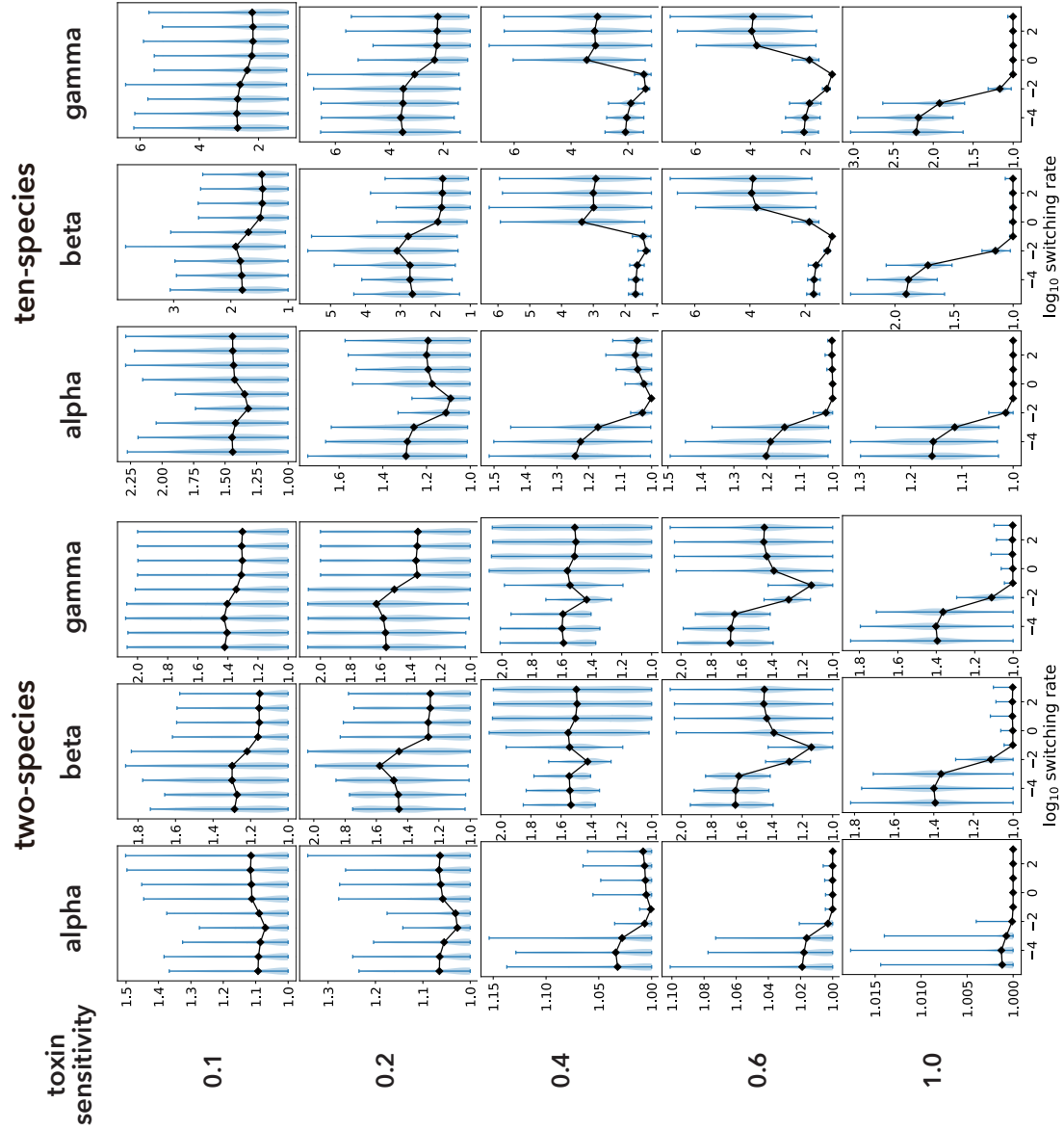
Figure A.3: Changes of alpha, beta, and gamma diversities

Alpha, beta, and gamma diversities over the switching rate and mean toxin sensitivity in two- and ten-species communities at the end of simulations. As alpha diversity is always closed to one, beta diversity and gamma diversity show similar trend. The black lines and blue areas represent the mean values and the probability distributions of the diversities calculated from 10'000 simulations. See Appendix 1.3 for more detail.
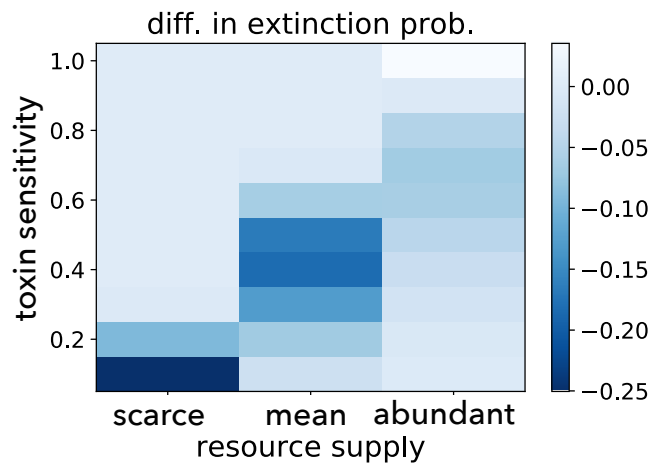
Figure A.4: Species 2's effect on species 1 in the absence of environmental switching

Species 2's effect on species 1 when the resource supply is fixed to be scarce (R_1^-), mean $(\langle R_1 \rangle)$, or abundant (R_1^+). The toxin sensitivities that minimize species 2's effect on species 1 correspond to the peak sensitivities in Fig. 5.
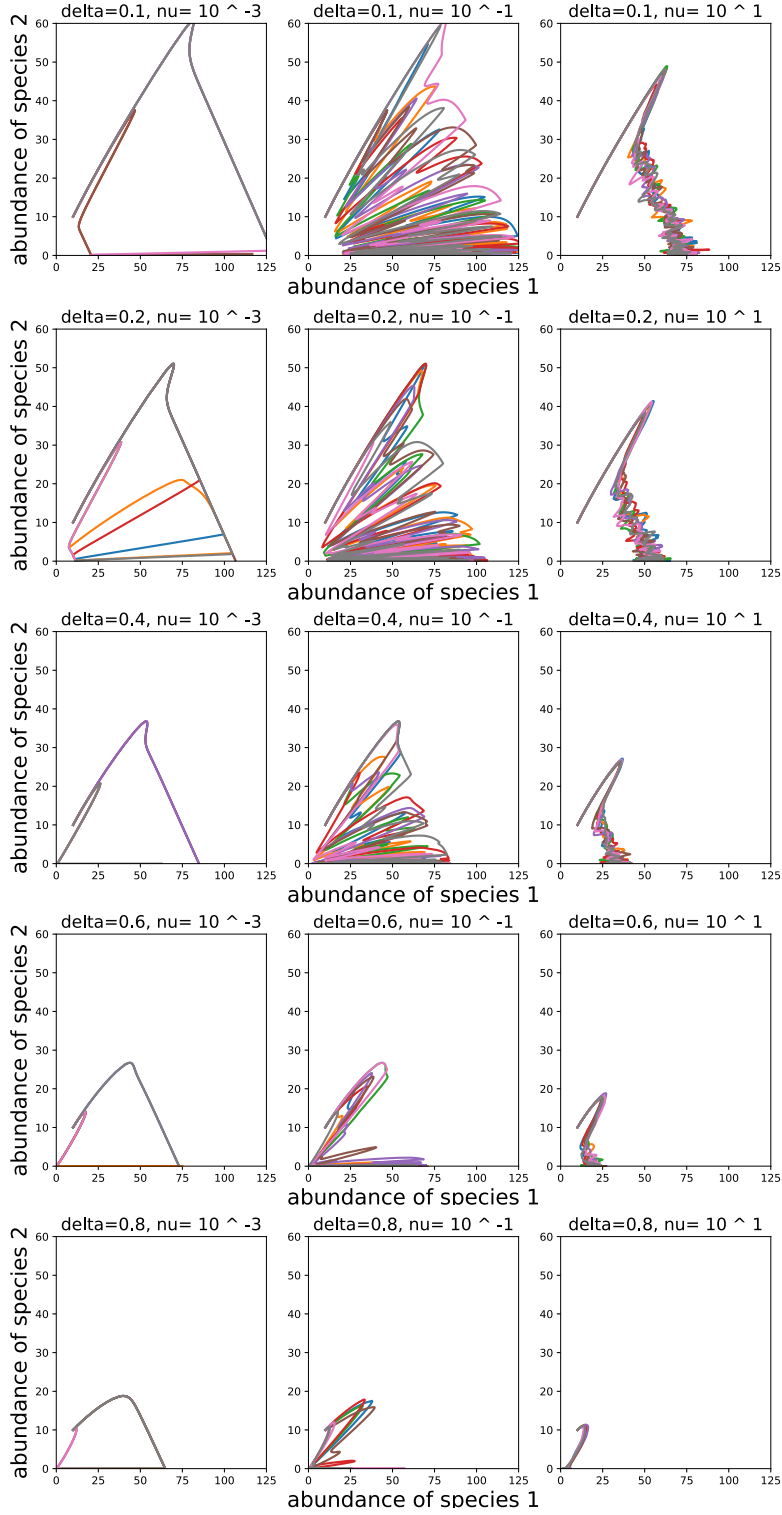
Figure A.5: Examples of the dynamics with only the environmental fluctuations

State transition of two species abundances (s_1, s_2) in the absence of DN but the presence of the EFs are shown. Here, EFs switch resource supply (scenario 1 in Table 1) and the initial population abundances are $s_1(0) = s_2(0) = 10$. In this setting, species 2 always goes extinct ($\lim_{\sigma \to \infty} s_2(\sigma) = 0$) regardless of the values of δ and ν. On the other hand, species 1 survives ($\lim_{\sigma \to \infty} s_1(\sigma) > 0$) if the environment is not too harsh. In each panel, different colors represent different samples of the dynamics with EFs alone. The values of ν and δ are shown on the top of each panel and the rest parameter values are shown in Table A.1.

19
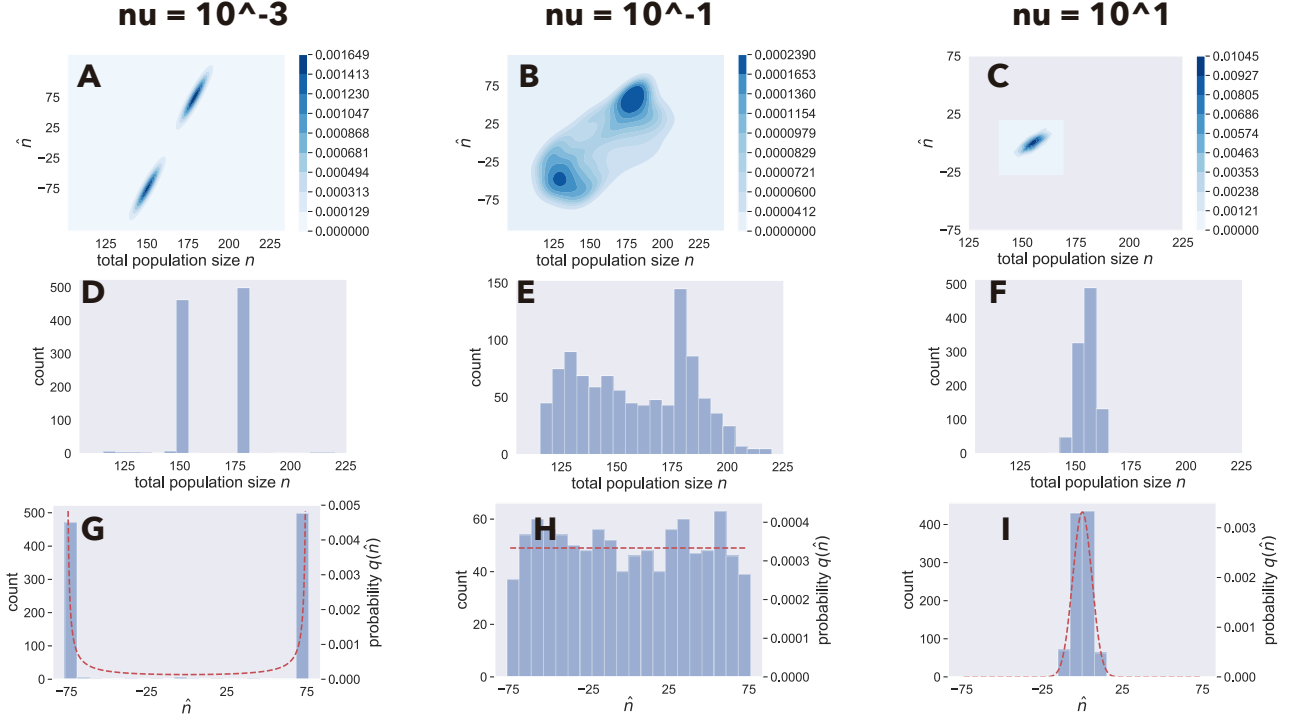
Figure A.6: Distributions of total population sizes

Probability distributions of total population size $n = r_1 + t_1 + s_1 + s_2$ and the auxiliary quantity $\hat{n} = r_1 - t_1 + s_1 + s_2$ obtained from 1,000 runs of Eqs (A.9a)-(A.9c) with environmental switching of scenario 1 in Table 1 at the slow $\nu = 10^{-3}$ (first column), intermediate $\nu = 10^{-1}$ (second column), or fast $\nu = 10^1$ (third column) switching rates. We collected the simulation data at time $\sigma_{end} = 200$. At the beginning of each simulation, $\xi = 1$ with 50 percents; otherwise $\xi = -1$. (A-C): the contour plots show the joint probability distributions of n and $\hat{n}$: large n corresponds to large $\hat{n}$. (D-F): the histograms show the distributions of the total population size n. (G-I): the histograms of $\hat{n}$ (left y-axis) and its theoretical probability distributions $q_{\nu/\alpha}(\hat{n})$ (right y-axis), which is given by Eq(A.22), are shown in blue bars and red dashed lines, respectively. We used $\delta_{1,1} = \delta_{1,2} = 0.2$ and all other parameter values are shown in Table A.1, and thus $\nu/\alpha = 0.01$ corresponds to the slow switching rate (first column), $\nu/\alpha = 1$ corresponds to the intermediate switching rate (second column), and $\nu/\alpha = 100$ corresponds to the fast switching rate (third column), respectively, see Appendix 2.2. Exceptionally, we used $\nu/\alpha = 80$ instead of $\nu/\alpha = 100$ to show $q_{\nu/\alpha}(\hat{n})$ in the third column because $\nu/\alpha = 100$ causes overflow during the calculation of $q_{\nu/\alpha}(\hat{n})$, but this modification does not change the qualitative feature of $q_{\nu/\alpha}(\hat{n})$.
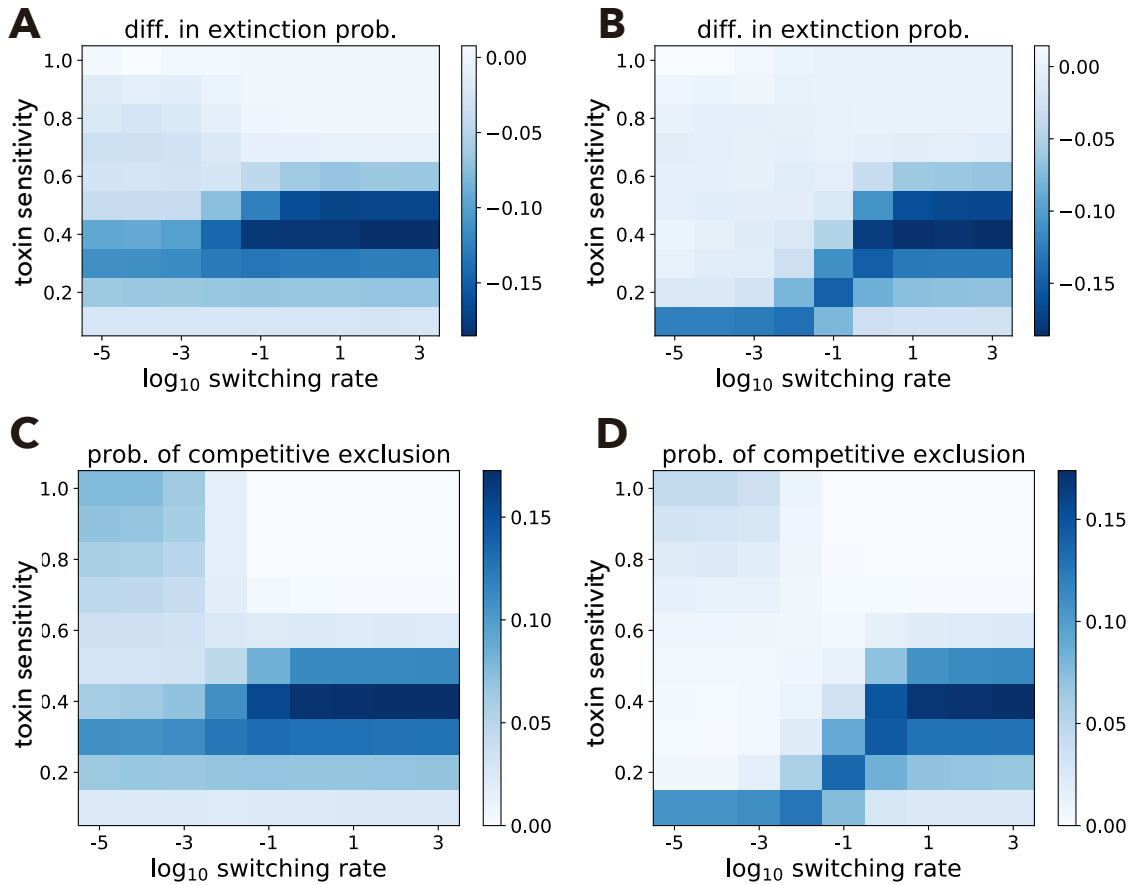
Figure A.7: Effects of the environmental switching rate in alternative scenarios

Examples of the effect of switching rate in alternative scenarios. In the left column (A and C), toxin supply is switching (scenario 2), while both resource and toxin supplies switch and are negatively correlated (scenario 3) in the right column (B, D). A and B: difference between extinction probabilities in absence and presence of species 2. C and D: competitive exclusion probability. Parameter values: $R_1^+ = 200$, $R_1^- = 50$, $T_1^+ = 200$, and $T_1^- = 50$.
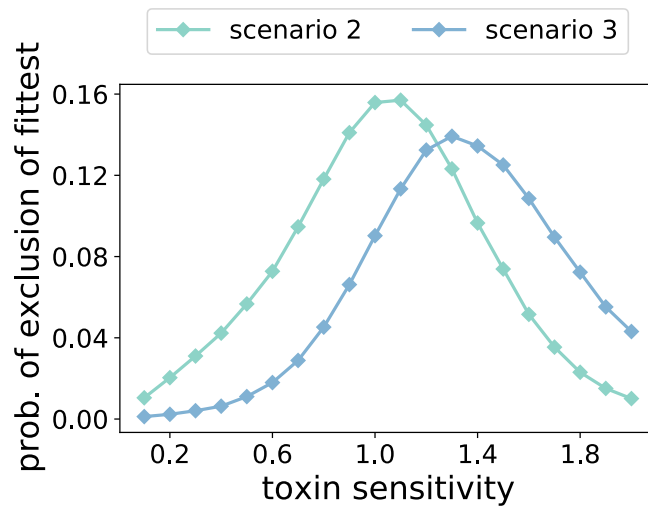
Figure A.8: Critical toxin sensitivities under mild environments

The critical toxin sensitivities (i.e., toxin sensitivity that maximizes the probability of exclusion of the fittest in the absence of environmental switching) under the mild environments (scenario 2 :scarce toxin supply $T_1^- = 50$, and scenario 3: abundant resource supply $R_1^+ = 200$ and scarce toxin supply) are > 1.
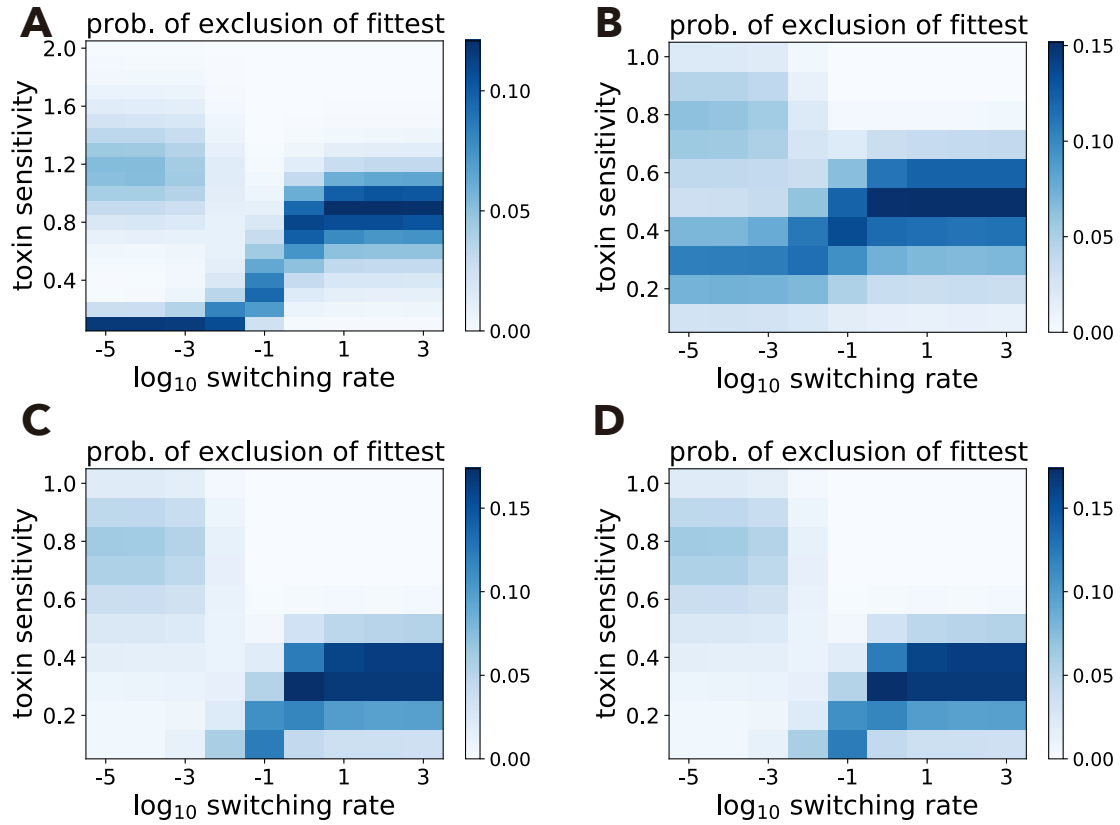
Figure A.9: Effects of resource supplies on exclusion of the fittest

Top: abundant resource supply becomes twice (A) or half (B) of $R_1^+ = 200$, i.e. $R_1^+ = 400$ in panel (A) and $R_1^+ = 100$ in panel (B). Bottom: scarce resource supply becomes twice (C) or half (D) of $R_1^- = 50$, i.e. $R_1^- = 100$ in panel (C) and $R_1^- = 25$ in panel (D).
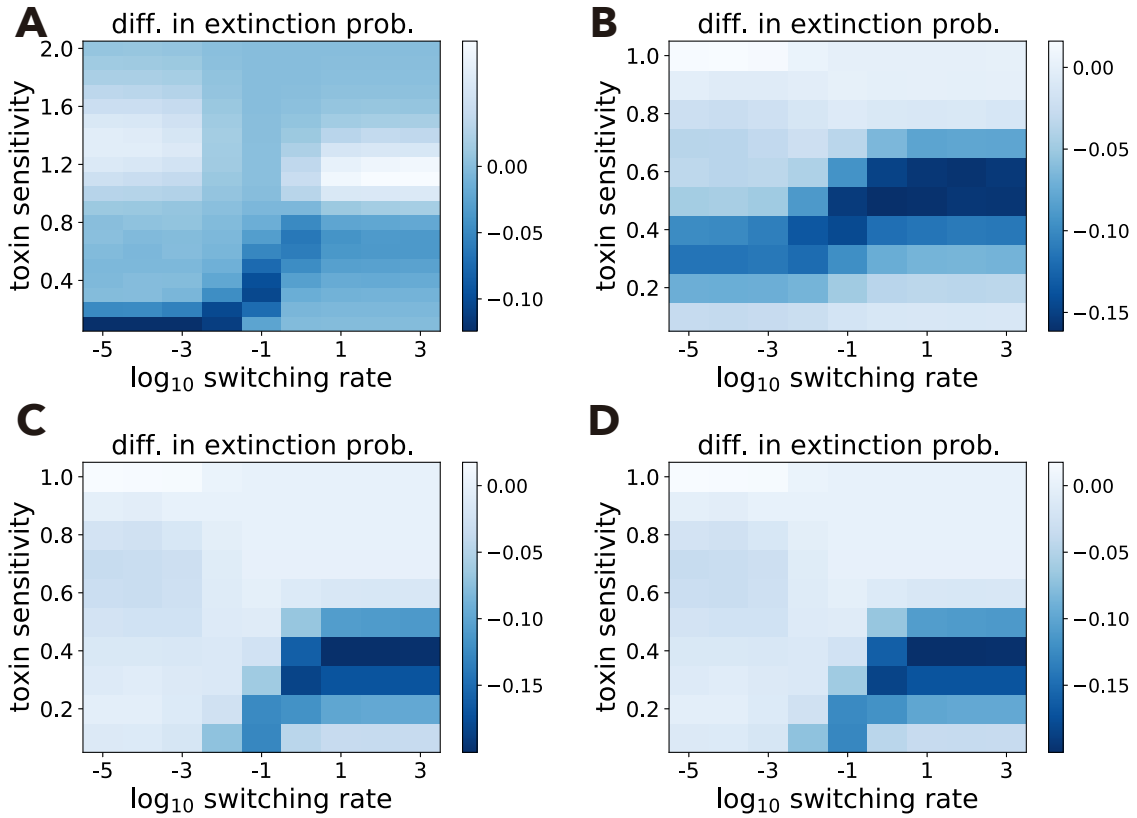
Figure A.10: Effects of resource supplies on difference in extinction probability

Similar to Fig. A.9 but showing species 2's effect on species 1's extinction probability. Top: abundant resource supply becomes twice (A) or half (B) of $R_1^+ = 200$, i.e. $R_1^+ = 400$ in panel (A) and $R_1^+ = 100$ in panel (B). Bottom: the scarce resource supply becomes twice (C) or half (D) of $R_1^- = 50$, i.e. $R_1^- = 100$ in panel (C) and $R_1^- = 25$ in panel (D). We plotted toxin sensitivity from 0.1 to 2.0 in panel A to see non-monotonic positive species interactions (see also Fig.A.11)
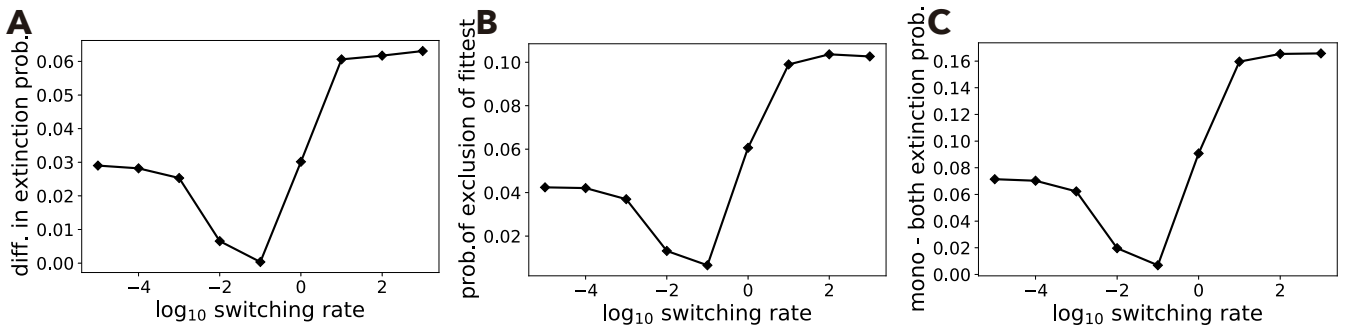

Figure A.11: Positive interaction strength varies non-monotonically with switching rate

A: Difference in species 1's extinction probability with positive sign at $\delta = 1.0$ and $R_1^+ = 400$, showing a non-monotonic effect of the environmental switching rate. B: Probability of exclusion of the fittest. C: Difference between the probability of species 1 going extinct alone and both species going extinct.
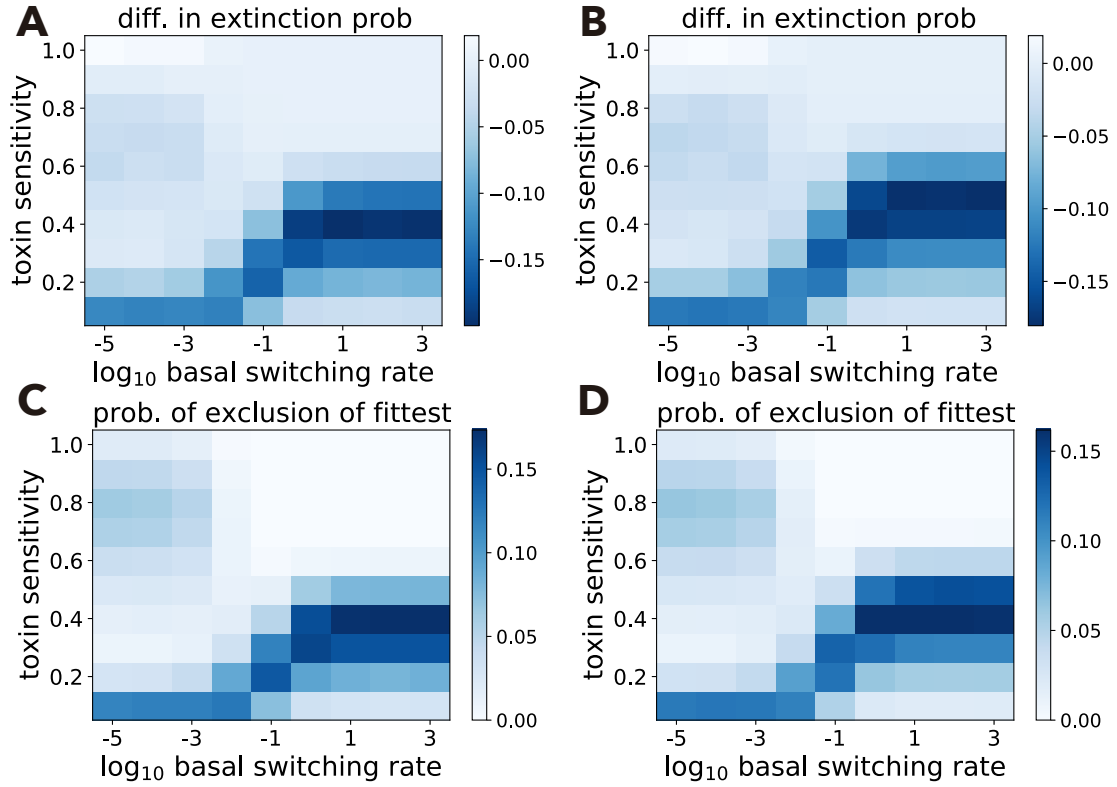
Figure A.12: Extinction probabilities under asymmetrically switching environments

Difference of species 1's extinction probability and the probability of exclusion of the fittest under asymmetric switching rates over the baseline of the switching rate ν and toxin sensitivity. Here, environmental fluctuations change the amounts of resource supplies. A and B: Difference of species 1's extinction probabilities in mono-culture minus co-culture wit species 2 when (A) the harsh environment continues longer than the mild environment ($\beta_1 = 1.2$ and $\beta_2 = 1$), or when (B) the mild environment continues longer than the harsh environment ($\beta_1 = 1$ and $\beta_2 = 1.2$), respectively. C and D: Probabilities that species 2 excludes species 1 when the environmental switching rates are identical to panels A and B, respectively. Other parameter values are shown in Table A.1.
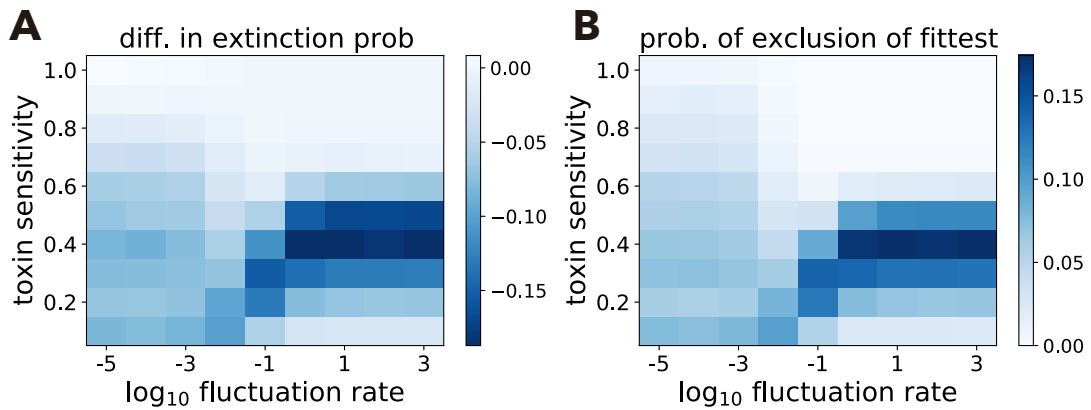

Figure A.13: Extinction probabilities under cyclic environmental changes

Difference in species 1's extinction probability and the probability of exclusion of the fittest under under cyclically fluctuating resource supplies among four states: $R_1\ (\xi = 1) = 200$, $R_1\ (\xi = 2) = 150$, $R_1\ (\xi = 3) = 100$, and $R_1\ (\xi = 4) = 50$. A: Difference of species 1's extinction probabilities in mono-culture minus co-culture with species 2. B: Probability that species 1 goes extinct but species 2 survives. Other parameter values are shown in Table A.1.
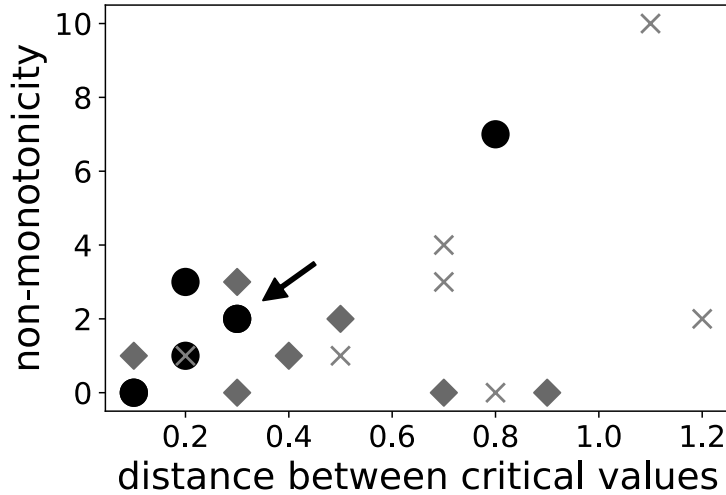
Figure A.14: Non-monotonicity and critical toxin sensitivities

The number of times we observe non-monotonic changes of species 1's difference in extinction probability across the explored parameter range varies with the distance between the two critical toxin sensitivities, depending on where distances are measured (between harsh and mean environments: dots, between mean and mild: environment diamonds, and between harsh and mild environments: crosses). These three distances were measured in each of the following seven scenarios: three different scenarios of environmental switching (Table 1 and Appendix 3) and four environmental switching scenario 1s with changing amounts of resource supplies (Appendix 4). The correlation is only significantly positive for the distance between scarce resource or abundant toxin supplies (i.e., harsh environments) and mean resource/toxin supplies (Spearman's $\rho = 0.77$, P-value: 0.043). The dot indicated by the arrow corresponds to the scenario analyzed in the main text.
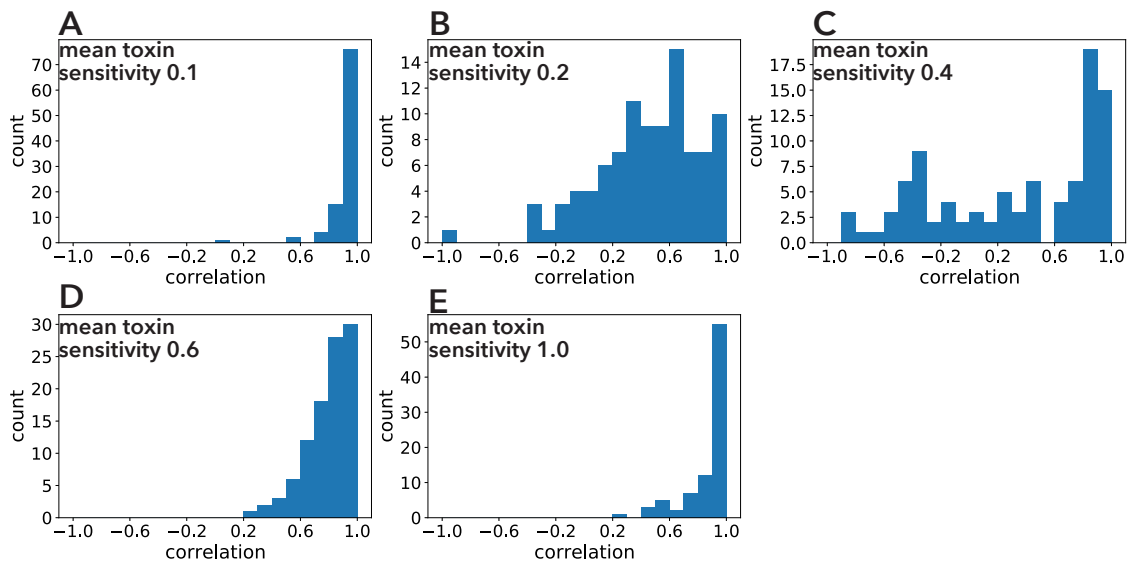

Figure A.15: Correlations between exclusion of the fittest and beta diversity in two-species communities

The Pearson correlation coefficients between probability of exclusion of the fittest (first column of Fig. 6) and beta diversity (second column of Fig. 6) in two-species communities are shown. Each panel differs in mean toxin sensitivity (A: 0.1, B: 0.2, C: 0.4, D: 0.6, and E: 1.0) and we analyzed 100 two-species communities in each case.
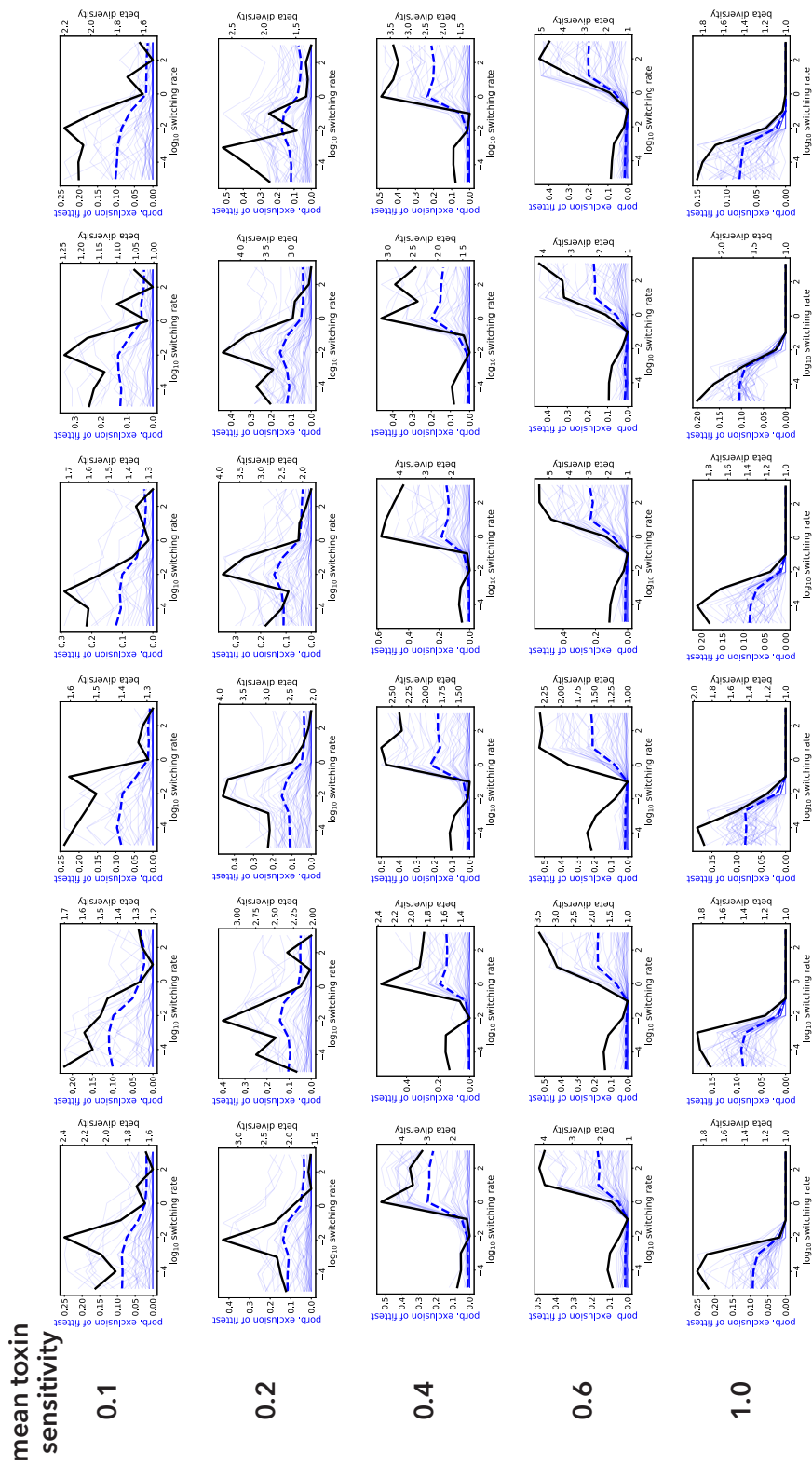
Figure A.16: Beta diversity and exclusion of the fittest in ten-species communities

Each panel represents the beta diversity of a ten-species community (black lines), probabilities of exclusion of the fittest in 45 species pairs (solid blue lines) and the mean probability of exclusion of the fittest over the 45 pairs (dashed blue lines). Each row corresponds to mean toxin sensitivity $\bar{\delta}$. One can see six examples of ten-species communities at each mean toxin sensitivity.

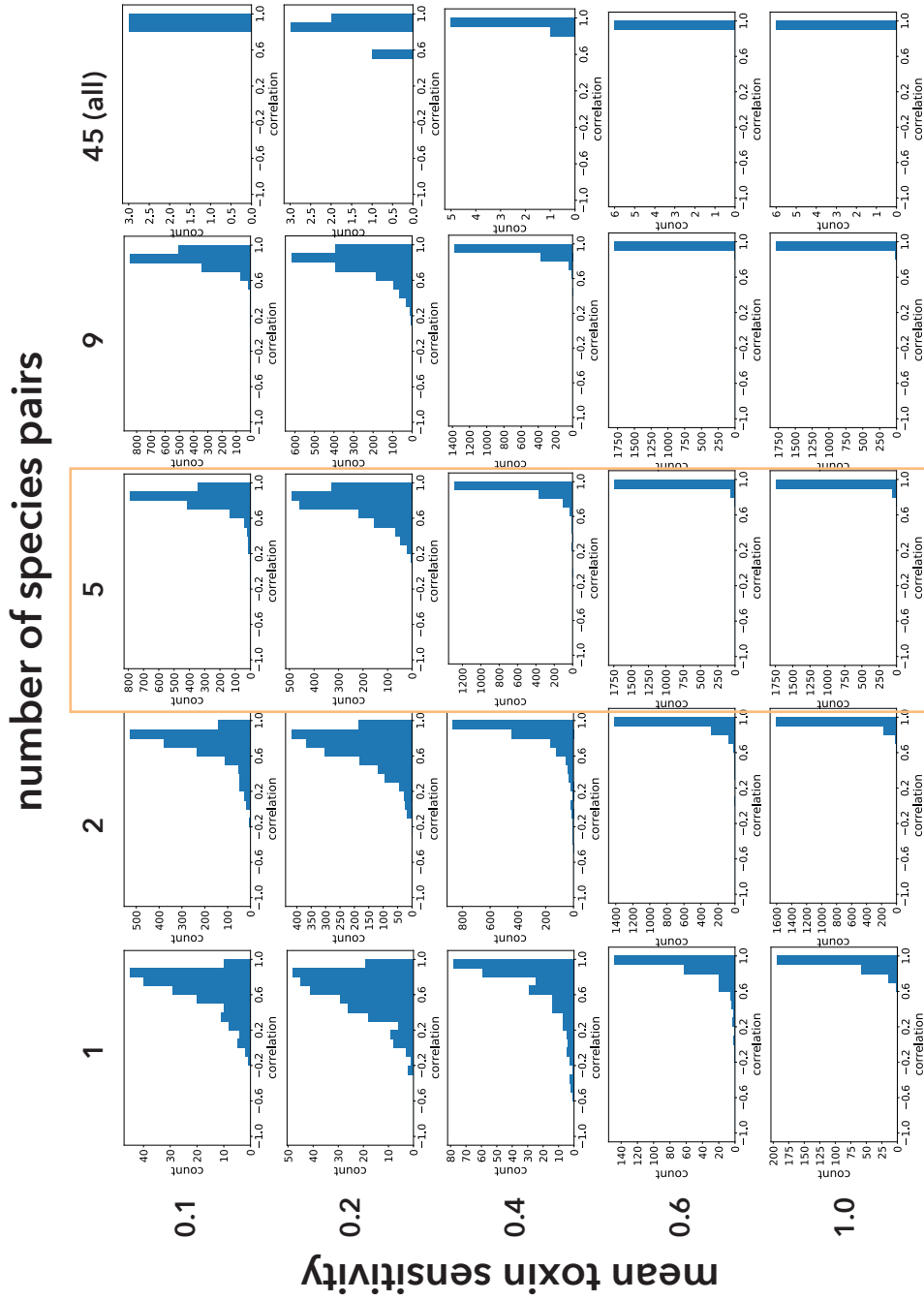Figure A.17: Correlations between exclusion of the fittest and beta diversity in ten-species communities

The Pearson correlation coefficients between mean probability of exclusion of the fittest and beta diversity (Fig. A.16) in ten-species communities are shown. Each row represents the different value of mean toxin sensitivity $\bar{\delta}$ while each column represents the different number of sampled species pairs m.
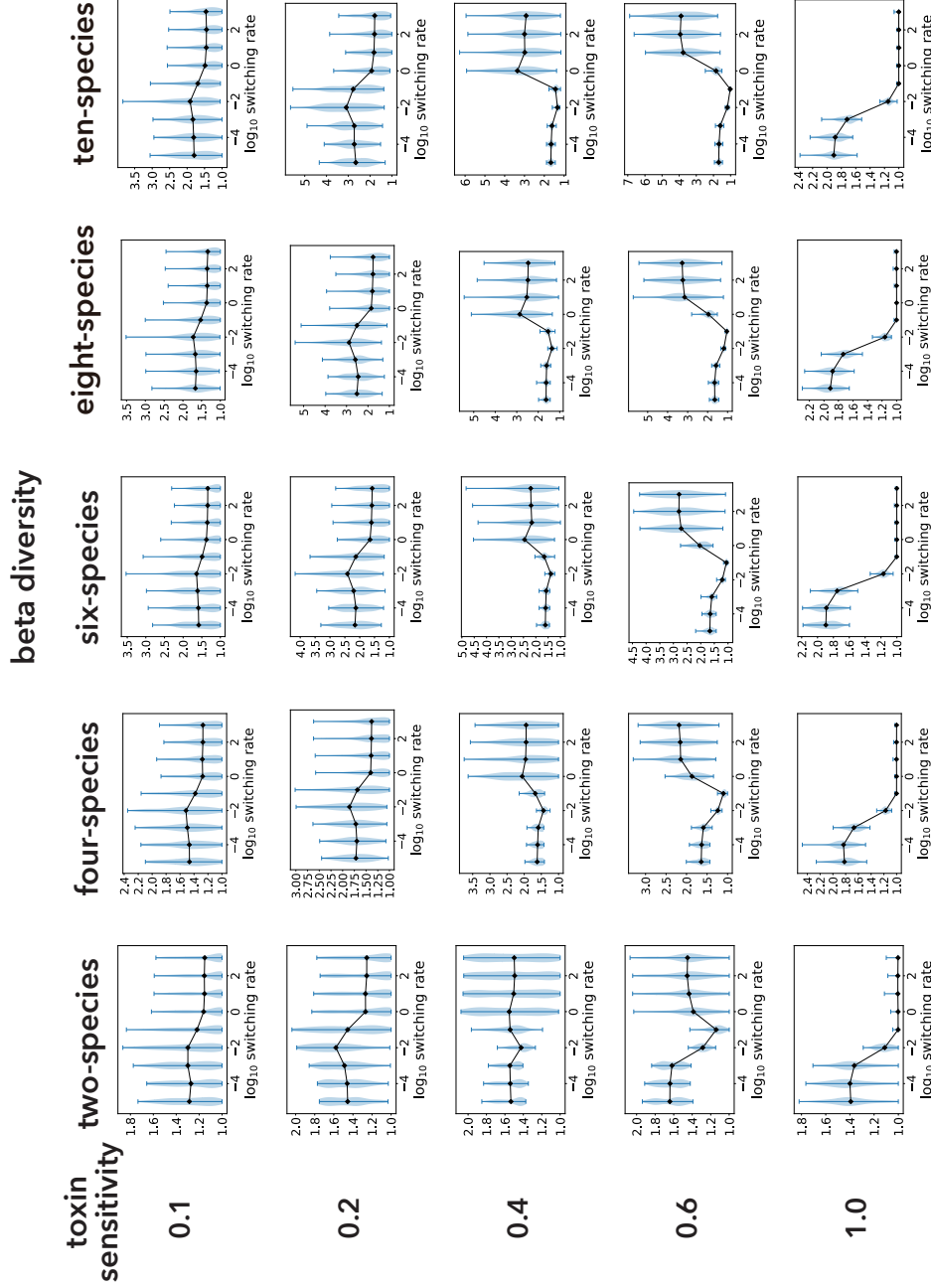
Figure A.18: Beta diversity with increasing initial number of species in a community

Beta diversities with increasing initial numbers of species N and mean toxin sensitivities $\bar{\delta}$. The black lines show the means and blue areas represent the probability distributions calculated by 10'000 simulations (100 beta diversity measurements using different parameter sets, each from 100 replicate runs).
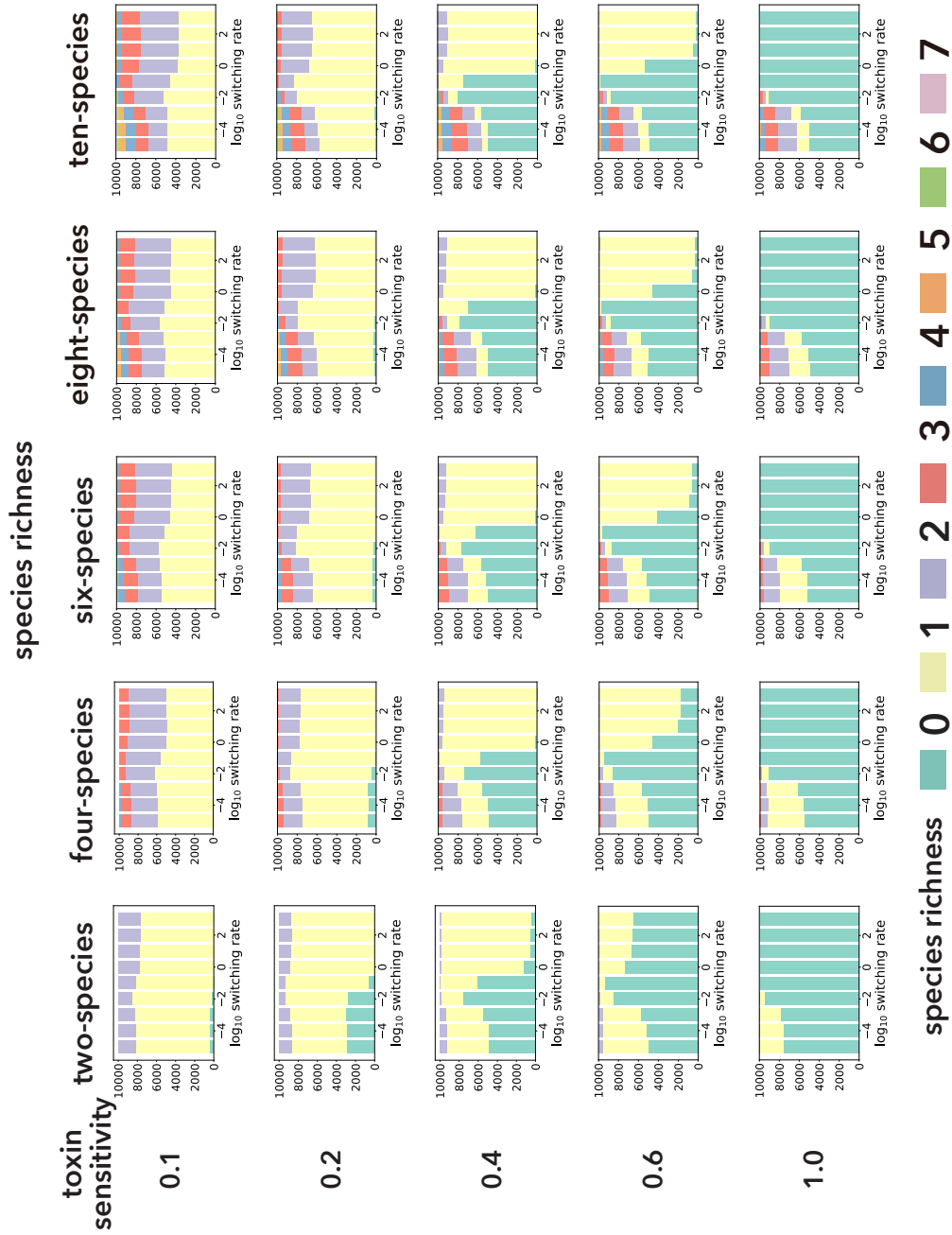
Figure A.19: Species richness with increasing initial number of species in a community

Species richness with increasing initial numbers of species N and mean toxin sensitivities $\bar{\delta}$. Each bar plot represents results of 10'000 simulations.
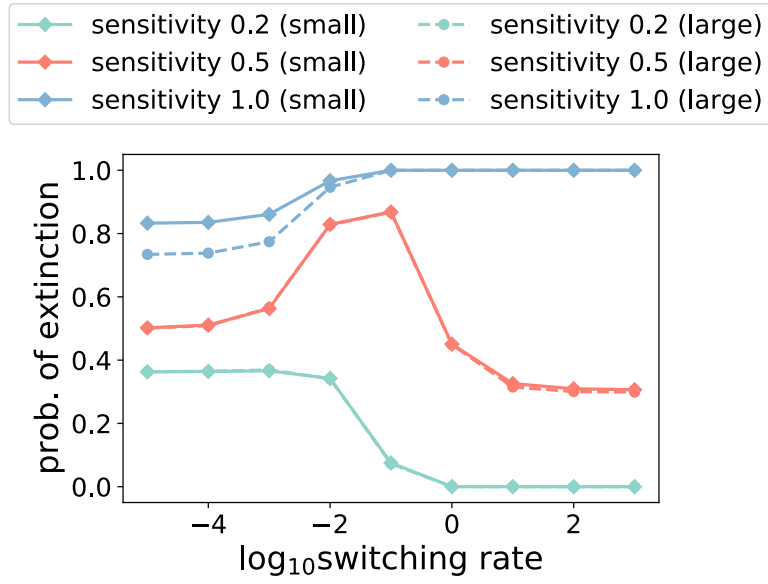
Figure A.20: Initial Population size's effect on extinction probabilities

Extinction probability of species 1 in mono-culture when the initial species abundance is default (small: $s_1(0) = 10$) or larger (large: $s_1(0) = 20$). When the switching rate is small and the toxin sensitivity is 1.0, the extinction probability is lower with the larger initial species abundance. In the rest cases, the extinction probabilities are not affected by the initial species abundance. The parameter values are as shown by Table A.1.
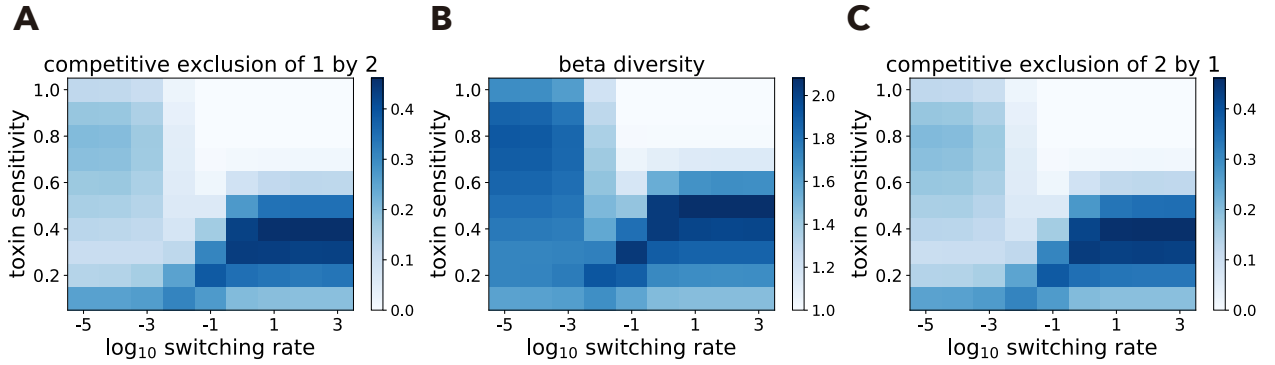
Figure A.21: Exclusion probabilities and beta diversity in neutral cases

Two species dynamics under neutral cases (i.e., two species differ only in their labels). In this case, the probabilities that species 2 outcompete species 1 (A) are identical to those that species 1 outcompete species 2 (C). Without loss of generality, we can call exclusion of either species as exclusion of the fittest in the neutral scenarios. As in Fig. 6, there are similarities between how the exclusion of the fittest (A or C) and beta diversity (B) changes over the switching rate at each toxin sensitivity. The parameter values are as shown by Table A.1.
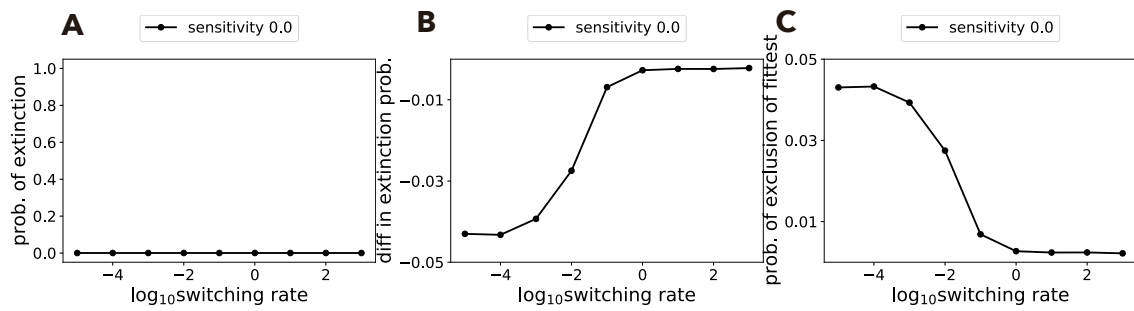
Figure A.22: Extinction probabilities with zero toxin sensitivity

When species's toxin sensitivities are zero, the dynamics are equivalent with the case without toxins because species die only due to dilution. A: Species 1 does not go extinct in mono-culture. B: Difference of species 1's extinction probabilities in mono-culture minus co-culture. C: Probability that species 2 outcompetes species 1 in co-culture. Panels B and C show qualitatively similar results with the case of toxin sensitivity 0.1, see Figs. 3A and C because species 2 can outcompete species 1 under the harsh environment with low switching rates, due to DN.

Table A.1: List of fixed parameters in the interaction analysis

Symbol	Value	Description
α	0.1	dilution rate of the chemostat
$R_1^{\pm}$	$R_1^+ = 200, R_1^- = 50$	abundant or scarce resource supply concentration
$T_1^{\pm}$	$T_1^+ = 200, T_1^- = 50$	abundant or scarce toxin supply concentration
Y_{1k}^r	$Y_{1k}^r = 1$ for $k = 1, 2$	species k's biomass yields of resource
Y_{1k}^t	$Y_{1k}^t = 1$ for $k = 1, 2$	species k's biomass yields of toxin
μ_{11}	1.0	maximum growth rate of species 1 on resource 1
μ_{12}	0.91	maximum growth rate of species 2 on resource 1
δ_{1k}	$[0.1, \ldots, 1.0]$ and $\delta = \delta_{11} = \delta_{12}$	sensitivity of species k to toxin 1.
K_{1k}^r	100	amount of resource 1 that gives half-max growth rate of species k
K_{1k}^t	100	amount of toxin 1 that gives half-max death rate of species k

Table A.2: Summary of critical toxin sensitivities and number of times non-monotonicity is observed

Switching scenario		1	2	3		1		
Amounts of supplies		base line *			$\uparrow R_1^+$	$\downarrow R_1^+$	$\uparrow R_1^-$	$\downarrow R_1^-$
Critical toxin	harsh	0.1	0.3	0.1	0.1	0.1	0.3	0.1
sensitivities	mean	0.4	0.4	0.4	0.9	0.2	0.5	0.3
	mild	0.8	1.1	1.3	1.2	0.3	0.8	0.8
number of	mean and harsh	2	0	2	7	0	1	3
non-monotonic changes	mean and mild	1	0	0	3	0	0	2
observed between	mild and harsh	3	0	2	10	0	1	4 †

* For exact parameter values, see Table A.1.
† At the critical toxin sensitivity corresponding to the mean environment ($\delta = 0.3$), the species interaction non-monotonically changes over the switching rate: the frequency of non-monotonicity is $3 + 2 - 1 = 4$.

Appendix C

Appendices of Chapter 4

C.1 Details of the mathematical model

C.1.1 Details of the consumer-resource model

In the analyses of computational costs of the algorithms, the dynamics of the system follow Eqs (4.2a) and (4.2b), with the assumption that each species interacts with three of $M = 10$ compounds. The two of the three compounds are resources that are either substitutable or complementary and these dynamics follow the models in León and Tumpson (1975). The rest compound is either metabolite that the focal species secretes or toxin that the focal species absorbs and degrades. For example, when compounds 1 and 2 are resources and compound 3 is toxin or metabolite for species k, the impact vector and the growth function is given by

$$\text{substitutable resource and toxin} \begin{cases} g_k\left(\vec{C}\right) &= \sum_{j=1,2} \gamma_{jk} \frac{\beta_{jk} C_j}{C_j + K_{jk}} - \gamma_{3j} \frac{\beta_{3j} C_3}{C_3 + K_{3k}} - c_{3k} \\ f_{jk}\left(\vec{C}\right) &= -\frac{\beta_{jk} C_j}{C_j + K_{jk}}, \quad j = 1, 2, 3 \end{cases}$$

(C.1)

$$\text{substitutable resource and metabolite} \begin{cases} g_k\left(\vec{C}\right) &= \sum_{j=1,2} \gamma_{jk} \frac{\beta_{jk} C_j}{C_j + K_{jk}} - c_{3k} \\ f_{jk}\left(\vec{C}\right) &= \begin{cases} -\frac{\beta_{jk} C_j}{C_j + K_{jk}}, & j = 1, 2 \\ \max\{0, g_k/\gamma_{jk}\} & j = 3 \end{cases} \end{cases}$$

(C.2)

$$\text{complementary resource and toxin} \begin{cases} g_k\left(\vec{C}\right) &= \min_{j=1,2} \gamma_{jk} \frac{\beta_{jk} C_j}{C_j + K_{jk}} - \gamma_{3k} \frac{\beta_{3k} C_3}{C_3 + K_{3k}} - c_{3k} \\ f_{jk}\left(\vec{C}\right) &= \begin{cases} -\gamma_{jk}^{-1} \min_i \gamma_{jk} \frac{\beta_{jk} C_j}{C_j + K_{jk}}, & j = 1, 2 \\ -\frac{\beta_{jk} C_j}{C_j + K_{jk}} & j = 3 \end{cases} \end{cases}$$

(C.3)

$$\text{complementary resource and metabolite} \begin{cases} g_k\left(\vec{C}\right) &= \min_{j=1,2} \gamma_{jk} \frac{\beta_{jk} C_j}{C_j + K_{jk}} - c_{3k} \\ f_{jk}\left(\vec{C}\right) &= \begin{cases} -\gamma_{jk}^{-1} \min_i \gamma_{jk} \frac{\beta_{jk} C_j}{C_j + K_{jk}}, & j = 1, 2 \\ \max\{0, g_k/\gamma_{jk}\} & j = 3 \end{cases} \end{cases}$$

(C.4)

where β_{jk} represents species k's maximum consumption or absorption rate of compound j, K_{jk} is the concentration of compound j that gives the half-max value of consumption, absorption, or production of the compound, γ_{jk} is species k's yield of compound j, and c_{3k} is the cost of degradation of toxins or production of metabolites. Each species randomly chose three compounds to interact with and the interaction types to apply. These assumptions enable us to introduce many species-compounds interactions that would exist in nature with a minimum setup. The parameter values are sampled from the following probability distributions:

$$\beta_{jk} \sim \begin{cases} \mathcal{U}(0,1) & \text{if } j \text{ is resource} \\ \mathcal{U}(-1,0) & \text{if } j \text{ is toxin or metabolite} \end{cases} \tag{C.5}$$

$$\gamma_{jk} \sim \begin{cases} \mathcal{N}(1,0.25^2) & \text{if } j \text{ is resource} \\ \mathcal{N}(-1,0.25^2) & \text{if } j \text{ is toxin or metabolite} \end{cases} \tag{C.6}$$

$$c_{jk} \sim \text{Exp}(10) \quad \text{if } j \text{ is toxin or metabolite} \tag{C.7}$$

where $\mathcal{U}(a,b)$ represents the uniform distribution from a to b, $\mathcal{N}(\mu,\sigma^2)$ represents the normal distribution whose mean and standard deviation are μ and σ, respectively, and $\text{Exp}(\lambda)$ represents the exponential distribution with rate parameter λ giving the mean of the exponential distribution $1/\lambda$.

To evaluate the community functions, I assumed that 30% of compounds in the system are aimed to produce while the reset of them to degrade. Following this assumption, the weight vectors are sampled from the following probability distribution:

$$w_i \sim \begin{cases} \text{Exp}(1) & \text{with } 70\% \\ -\text{Exp}(1) & \text{otherwise.} \end{cases} \tag{C.8}$$

Without loss of generality, $\vec{w}$ is normalized so that $|\vec{w}| = 1$. For given number of N, 100 these parameter sets were sampled to calculate the algorithms' computational costs.

C.1.2   Stability analysis

This subsection analyzes the local stability in the multi-stage chemostat system defined by Eqs (4.2a) and (4.2b). Here, we have M ordinary differential equations (ODEs) for compound concentrations and one ODE for an allocated species abundances in each stage: in total, σ-stages system has $\sigma(M+1)$ ODEs. Without loss of generality, I label species i for one that is allocated to stage i in this subsection. When the same species is allocated to multiple stages (stages i and k), we consider that we have two species that have identical parameter values ($g_i\left(\vec{C}\right) = g_k\left(\vec{C}\right)$ and $f_{ji}\left(\vec{C}\right) = f_{jk}\left(\vec{C}\right)$ for $j = 1, \ldots, M$).

The Jacobian matrix of σ-stage chemostat is written as follows:

$$J = \begin{pmatrix} \mathcal{J}_1 & 0 & \ldots & 0 \\ D & \mathcal{J}_2 & 0 & \vdots \\ 0 & \ddots & \ddots & 0 \\ 0 & \ldots & D & \mathcal{J}_\sigma \end{pmatrix} \tag{C.9}$$

where $\mathcal{J}_i$ represents $(M+1) \times (M+1)$ Jacobian matrix corresponding to stage i

$$\mathcal{J}_i = \left(\begin{array}{c|c} \overbrace{-\alpha\delta_{jl} + \dfrac{\partial f_{li}}{\partial C_{ij}}}^{M \times M} & \overbrace{f_{ji}\left(\vec{C}_i\right)}^{M \times 1} \\ \hline \underbrace{\dfrac{\partial g_i}{\partial C_{ij}}S_i}_{1 \times M} & \underbrace{\left\{ g_i\left(\vec{C}_i\right) - \alpha \right\}}_{1 \times 1} \end{array} \right), \quad (j, l = 1, \ldots, M) \tag{C.10}$$

with Kronecker delta δ_{jl}, and D is a $(M+1) \times (M+1)$ diagonal matrix:

$$D = \text{diag}\left(\underbrace{\alpha, \alpha, \ldots, \alpha}_{M}, 0 \right). \tag{C.11}$$

As J is a triangular matrix, its characteristic equation is

$$\det\left(\lambda I - J\right) = \prod_{i=1}^{\sigma} \det\left(\lambda I - \mathcal{J}_i\right) = 0$$

$$\Leftrightarrow \exists i = 1, \ldots, \sigma, \det\left(\lambda I - \mathcal{J}_i\right) = 0 \tag{C.12}$$

where λ is an eigenvalue and I is an identical matrix. The eigenvalues of J are, therefore, composed of those of $\mathcal{J}_i$ for $i = 1, \ldots, \sigma$. This implies that the necessary and sufficient condition for a local stable equilibrium in σ-stages system (i.e., $\max \text{Re}\lambda < 0$) is that each stage $i = 1, \ldots, \sigma$ is locally stable.

An equilibrium state in stage i $(\vec{C}_i^*, S_i^*)$ depends on an equilibrium state in stage $i-1$ because Eq (4.2a) contains C_{i-1j}: the dynamics in stage i equilibriate after those in stage $i-1$ do. If we assume that the dynamics in stage i quickly equilibrate, we can separately simulate the dynamics in each stage rather than simulating the all stages' dynamics at one time. For this reason, I assume such quick stabilization in each stage: inflow compound j' concentration C_{i-1j} in Eq (4.2a) is given by Eq (4.3) and the initial compound concentrations are $\vec{C}_i(0) = \vec{C}_{i-1}^*$. Of course, we can relax this assumption and allocate species to stage i before the dynamics in stage $i-1$ equilibrate, see e.g., Appendix C.4.

C.2   Algorithm 2: the stochastic algorithm

In Chapter 4, I introduced the deterministic algorithm that always finds the best allocations of species using prune and search. In this section, I improve this algorithm by introducing the probability of which species should be allocated to the focal chemostat. This algorithm has smaller computational costs than the deterministic algorithm and is able to find the allocations whose community functions are or are close to the maxima.

Recall that the deterministic algorithm analyzes the allocations where all allocated species survive. If the inoculation sizes of species are assumed to be very small ($S_k(0) \ll 1$), the deterministic algorithm analyzes the allocations where all allocated species have initial growth rates larger than the dilution rate ($g_k\left(\vec{C}_i(0)\right) > \alpha$). Therefore, it is reasonable to ignore the allocations of species whose initial growth rates are smaller than the dilution rate.

In addition, we can consider how much each species affect the chemical compounds' concentrations when they are inoculated. Here, the initial impact of species k at stage i is defined as follows:

$$y_k\left(\vec{C}_{i-1}^*\right) \equiv \begin{cases} -\sum_j w_j C_{i-1\,j}^* \mathrm{sign}\left(f_{jk}\right) & \text{if } g_k\left(C_{i-1}^*\right) > \alpha \\ -\infty & \text{otherwise,} \end{cases} \tag{C.13}$$

because $\vec{C}_i(0) = \vec{C}_{i-1}^*$ is assumed. y_k is large if species k can degrade or produce the compounds that strongly affects the target community function. In addition, the effect of degrading compound j is larger if there are more compound j in stage i.

Using this idea, I implemented the stochastic algorithm (Algorithm 2), where we sample species L time to allocate at stage $i = 2, \ldots, \sigma_m$. This algorithm tests the allocation of all species to stage 1; if we have only a few allocations in the first chemostat, the number of allocations analyzed in $i = 2, \ldots, \sigma_m$ can be very small. At stage $i \geq 2$, given the allocation to the previous stages ($A_1, \ldots, A_{i-1}$), the initial compounds' concentrations in stage i with this allocation is $\vec{C}_{i-1}^*$. Then, we can define the probability that species k is allocated to stage i using the following softmax function:

$$P_k(\vec{y}; K) = \begin{cases} 0 & \text{if } y_k = -\infty \text{ for all } k = 1, \ldots, N \\ \frac{\exp(y_k/K)}{\sum_{k=1}^N \exp(y_k/K)} & \text{otherwise} \end{cases} \tag{C.14}$$

where $\vec{y} = (y_1, \ldots, y_N)$ and K determines the shape of the softmax function. In the limit of $K \to 0$, species k that has the largest value of y_k is always allocated (i.e., the greedy algorithm). On the other hand, with large value of K, all species with positive initial impacts on the community function have equal probabilities to be allocated. The value of K was tuned by cross-validation and I used $K = 0.5$ as default (Fig. C.1). Species k whose initial growth rate is lower than the dilution rate cannot be allocated. If all species cannot be allocated, the algorithm stops.

Fig. C.2 summarizes the comparison of the deterministic algorithm with the stochastic algorithm. The left panel of Fig. C.2 suggests that the computational costs of the stochastic

Table C.1 – Spearman Correlation analysis

	L	Rel. optima	Rel. elapsed time
L	1.00 / 0.00	0.09 / 0.04	0.18 / 4.15×10^{-5}
Rel. optima		1.00/0.00	0.47 / 2.35×10^{-29}
Rel. elapsed time	.		1.00 / 0.00

ρ / p represent Spearman's ρ and p-value. The parameter values are $N = 8$, $\sigma_m = 16$, and $K = 0.5$.
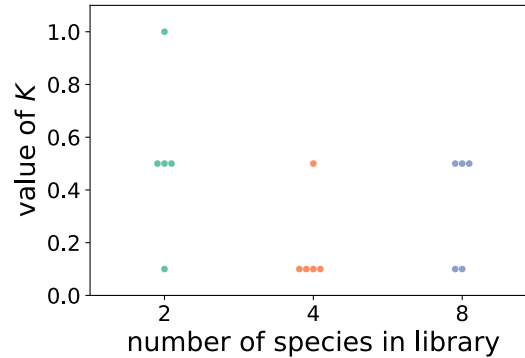
Figure C.1 – Tuned value of K

Optimal values of K in 5-fold cross validation. I used the same species parameter values in Fig.4.2.

algorithm is smaller the deterministic algorithm, especially, when we have many ($N \geq 4$) species in the library. In addition, the stochastic algorithm suggests allocations that have the highest community functions or allocations whose community function values are close to the highest values (i.e., quasi-best allocations) in many cases. The right panel of Fig. C.2 shows the ratio of the maximum community functions suggested by the stochastic algorithm to the true maximum community functions (i.e., relative optima). If relative optima are ones, the stochastic algorithm suggested the best allocations that maximize the community functions. In contrast, relative optima smaller than ones imply that the stochastic algorithm does not suggest the best microbial allocation; however, if the relative optima are close to ones (e.g., 0.95 or larger) the suggested allocations are close to the best ones. If one can accept such quasi-best allocations and wants to decrease the computational cost, the stochastic algorithm would be useful.

The stochastic algorithm can control its computational cost and relative optima by modifying the number of allocations to analyze. In Algorithm 2, I control the computational costs by changing L, which indicates the number of sampled species to allocate to stage $\sigma + 1$ per allocation using σ stage. Table C.1 summarizes how L, relative optima, and elapsed time relative to the deterministic algorithm relate to one another: the relative elapsed time is weakly correlated with L and moderately correlated with relative optima, respectively. Fig. C.3 shows that when $N = 8$, the stochastic algorithm with $L \geq 4$ suggests more than 95% of simulations show relative optima larger than 0.95 and the computational costs are about 7 times smaller than the deterministic algorithm. Setting $L \geq N/2$ as default would be reasonable.

Algorithm 2: Stochastic search algorithm

---

1 Function Main($\vec{C}_0$, σ_m, N, α, $\vec{w}$, K):

2   Initialization: Let $\sigma = 0$ and $\mathcal{B}(\sigma) = \{\}$, which represents a list of allocations and outflow compound concentrations;

3 A_σ^* denotes the allocation whose community function is the largest in the allocations using σ stages;

4   **while** $\sigma \leq \sigma_m$ **do**

5 $\mathcal{B}(\sigma + 1) =$ ADD_Species($\mathcal{B}(\sigma), \sigma, K$);

6   Find $A_{\sigma+1}^*$ from $\mathcal{B}(\sigma + 1)$;

7 $\sigma = \sigma + 1$;

8   **end**

9   **return** $\max_\sigma \phi(A_\sigma^*)$;

10 Function ADD_Species($\mathcal{B}(\sigma), \sigma, K$):

11   Initialize $\tilde{\mathcal{B}} = \{\}$;

12   Set number of samples par allocation L;

13   **if** $\sigma = 0$ **then**

14   In stage 1, try allocation of all species to sample many allocations for the later stages ;

15   **for** $k = 1, \ldots, N$ **do**

16   **if** $g_k(\vec{C}_0) \alpha$ *and species k persist* **then**

17   Add $\mathcal{A} = (k)$, its outflow compounds' concentrations, and its community function to $\tilde{\mathcal{B}}$;

18   **end**

19   **end**

20   **else**

21   **while** $\mathcal{B}$ *is not empty* **do**

22 $\tilde{\mathcal{A}}, \vec{C}_\sigma^* = \text{pop}(\mathcal{B})$.;

23   Note that $\tilde{\mathcal{A}} = \left(\tilde{A}_1, \ldots, \tilde{A}_\sigma\right)$;

24   Calculate species' initial impacts on the community function $y_k\left(\vec{C}_\sigma^*\right)$ and allocation probabilities $P_k(\vec{y}; K)$;

25   **if** $\vec{P} \neq \vec{0}$ **then**

26   **for** $l = 1, \ldots, L$ **do**

27   Choose one species to allocate to stage $\sigma + 1$ using the allocation probabilities;

28   Let denote chosen specie k_l;

29   **if** *Species k_l persist* **then**

30   Add new allocation $\left(\tilde{\mathcal{A}}, k_l\right)$, its outflow compounds' concentrations, and its community function to $\tilde{\mathcal{B}}$;

31   **end**

32   **end**

33   **end**

34   **end**

35   **end**

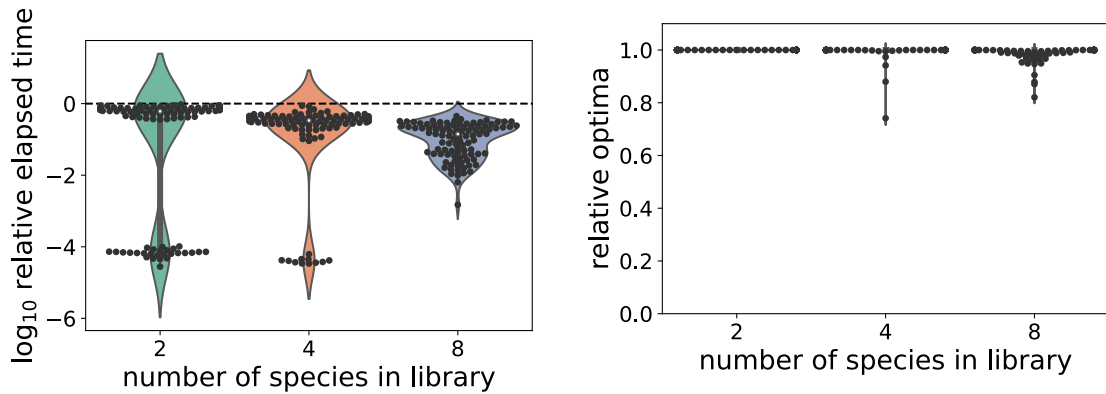**36**   **return** return $\tilde{\mathcal{B}}$;

Figure C.2 – Comparing stochastic algorithm with deterministic one

Left: Elapsed time of the stochastic algorithm relative to those of the deterministic algorithm. The dashed line represents a threshold where the elapsed time in the stochastic algorithm is the same as the deterministic one. The colored areas shows the kernel estimation of the distributions of relative elapsed time while black dots show the actual relative elapsed times. The white dots represent the median values. In this analysis, the number of sampling species per allocation in the stochastic algorithm is the half number of species in a library (i.e., $L = N/2$). Right: Maximum community function values suggested by the stochastic algorithm relative to those by the deterministic algorithm (i.e., relative optima) when $\sigma_m = 16$. If the value is one, the stochastic algorithm found an optimal allocation. If the value is lower than one, on the other hand, the stochastic algorithm fails in finding an optimal allocation. The parameter K for the softmax function is 0.5
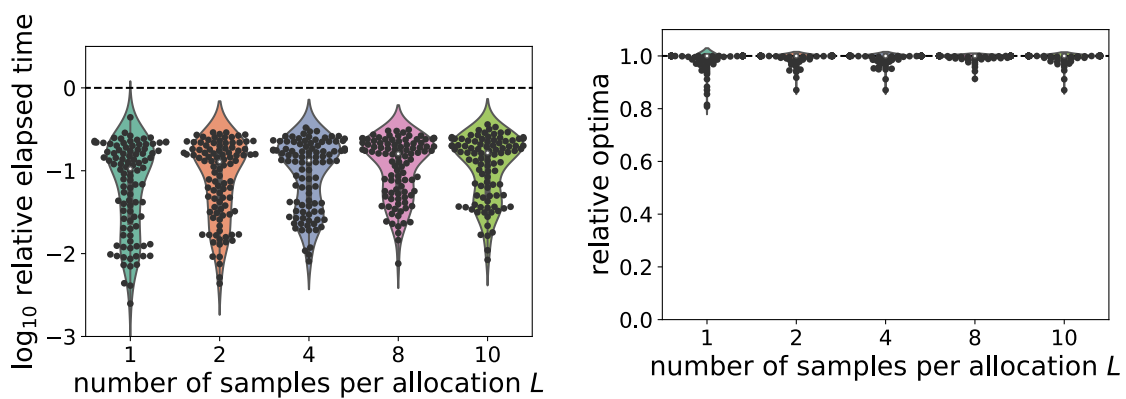

Figure C.3 – The trade-off between relative optima and computational costs

Relative elapsed time (left) and relative optima (right) with $N = 8, \sigma_m = 16$, and $K = 0.5$. The explanations for the symbols, lines and colored areas are the same in Fig. C.2 but this figure shows the effects of L.

C.3   Variation in the network structure of chemostats

In Chapter 4, I assume the chain network structures of the chemostats. However, we can consider various acyclic network structures of chemostats. In this section, I introduce branches in to the networks of the multi-stage chemostats, and analyze what is the typical network structure that maximizes the microbial community functions. The deterministic algorithm and the stochastic one are applicable even when the network has branches: the dynamics in the downstream would not affect those in the upstream even with branches. However, we cannot use these algorithms when the network has loops (e.g., a network with paths from stages i to j and from j to i), because the "downstream" communities can affect the dynamics in the "upstream" ones.

For considering the variation of chemostats' networks, I first redefine the object function. Suppose we use σ chemostats and two other tanks, called source ($i = 0$) and sink ($i = \sigma + 1$), respectively. The source provides the medium to the multi-stage chemostats while the sink collects the outflow medium from the system. The network structure of the whole system is defined by a $\sigma + 2 \times \sigma + 2$ adjacency matrix. In this section, I use α as the adjacency matrix because $\alpha_{ij} > 0$ indicates that there is a flow from stage j to stage i while stages i and j are not connected when $\alpha_{ij} = 0$. Note that $\alpha_{ii} = 0$ for $i = 0, \ldots, \sigma + 1$ because this network does not have any loop. The adjacency matrix in our system should satisfy the following equation:

$$\sum_{i=0}^{\sigma+1} \alpha_{ij} = \sum_{j=0}^{\sigma+1} \alpha_{ij}. \tag{C.15}$$

Otherwise, some chemostats flow all media out at the end. If each chemostat has unique inflow and outflow, we recover the chain network, which is analyzed in Chapter 4. In the following simulations, I assumed $\sum_{j=0} \alpha_{ij} = 0.05$ and $\alpha_{ij} = \alpha_{kj}$ for any pairs of (i, k) that have outflow from stage j.

The dynamics in stage i given by eqs (4.2a) and (4.2b) are re-written as follows by introducing branches in the chemostats' network:

$$\frac{dC_{ij}}{dt} = \sum_{l=0}^{\sigma+1} \alpha_{il} C_{il} - \sum_{l=0}^{\sigma+1} \alpha_{li} C_{li} + f_{jk}\left(\vec{C}_i\right) S_k \tag{C.16a}$$

$$\frac{dS_k}{dt} = \left\{ g_k\left(\vec{C}_i\right) - \sum_{l=0}^{\sigma+1} \alpha_{li} \right\} S_k. \tag{C.16b}$$

where species k is allocated to stage i Because the system may have multiple outflows to the sink, we also redefine the object function as follows:

$$\max \phi\left(\mathcal{A}; \alpha, \vec{w}\right) \equiv \alpha^{\dagger} \sum_{j=1}^{M} w_j \left(C_{0j} - C_{\sigma+1j}\right), \tag{C.17}$$

where $\alpha^\dagger \equiv \sum_{i=0}^{\sigma+1} \alpha_{i0} = \sum_{j=0}^{\sigma+1} \alpha_{\sigma+1j}$ (the sum of inflow from the source or the sum of outflow to the sink), compounds' concentrations in the sink are given by the equilibrated compounds' concentrations at each stage weighted by the dilution rate from each stage to the sink:

$$C_{\sigma+1j} = \sum_i \frac{\alpha_{\sigma+1i} C_{ij}^*}{\sum_i \alpha_{\sigma+1i}}. \tag{C.18}$$

As mentioned before, I do not consider the network structures with loops in this appendix. Without loss of generality, we can assume that there are flows from the source to stage 1 and from stage σ to the source, and that there is no flow from the source to the sink because it does not increase community functions. Then, the network without any loops is represented by a lower triangular matrix α. In this case, the stability of the system is analyzed as discussed in Appendix C.1.2.

When we maximize the microbial community functions allowing branches but without loops, we need to consider the number of possible network structures. If there are many network structures to analyze, the computational costs are too huge to analyze all network structures. Indeed, according to the assumption on the adjacency matrix above, we have

$$F(\sigma) = \frac{(\sigma+1)(\sigma+2)}{2} - 3 \tag{C.19}$$

free elements in the adjacency matrix: for each element, we can choose whether we have a flow between two chemostats (the element > 0) or not. Of course, due to the constrain from Eq (C.15), not all $2^{F(\sigma)}$ network structures are applicable to our system. However, Fig. C.4 suggests the number of the applicable network structures exponentially increases over the number of chemostats. This result implies that analyzing all acyclic networks with branches is not feasible even when σ is small.

To find patterns of how chemostats' network structures affect the community functions, I analyzed all six applicable network structures with $\sigma_m = 2$: two of them do not have branches but use one or two chemostats as discussed in the main text (hereafter, called chain networks with $\sigma = 1$ and 2, respectively), and the rest four have branches using 2 chemostats. I found that about 90% of the simulations suggest that the chain network either with $\sigma = 1$ or $\sigma = 2$ maximizes the community functions (Fig. C.5A). On the other hand, the rest simulations maximize the community functions in network with branches, but the increase of the community function by branches is less than 10% with three exceptions (Fig. C.5B). In two of the three exceptions (red dots in Fig. C.5B), we cannot allocate any species in the chain networks because the dilution rates is too high. However, we can allocate these species to the multi-stage chemostats with branches because the branches decrease the dilution rates in the chemostats. I continued the analysis by setting $\sigma_m = 3$. In this case, there are 60 network structures applicable to our system and analyzing all of them is not feasible. Instead, I randomly sampled 10 network structures with branches and asked whether branches improved the community functions or not. I obtained qualitatively similar results to the cases of $\sigma_m = 2$ (Fig. C.5C and D). 80% of

simulations maximize the community functions in the chain networks while the others maximizes the community functions when the network has three stages and some branches. The increase of the community functions by the branches is also around 10% in this case (except for the three cases). From these results, I conclude that analyzing only chain networks is reasonable because (i) such networks tend to maximize the community functions, (ii) improvement of the community functions by introducing branches into the multi-stage chemostats tends to small, and (iii) introduction of branches exponentially increases the computational costs.

The reason that the chain structures tend to maximize the community function can be intuitively explained as follows: by introducing branches into the chemostats, at least one dilution rate between chemostats decreases. Such a decrease in dilution rates weakens allocated species' contribution to a community function. In addition, the decrease of dilution rates weakens species interactions while an optimal allocation tends to choose species that increase population sizes of those downstream (Fig. C.6). Overall, as introducing branches exponentially increase computational costs but decrease or only slightly increase the community functions in most cases, we can conclude that analyzing only in chain network structures is reasonable to find the optimal microbial allocations.

However, one may want to know whether introducing branches improves a community function when (i) computational costs of simulating chemostats' networks with branches are relatively small (e.g., only two stages are available) and/or (ii) increasing community functions produces large benefits. Graph classification (Hamilton, 2020) is one option to predict whether we should search the optimal allocation without introducing branches into a chemostat' network or not given a community. I implemented graph classification with graph neural network using StellarGraph (Data61, 2018) and TensorFlow (Abadi et al., 2015) in Python to predict in which case we should consider the network structures with branches when $\sigma_m = 2$. Because I have only 300 simulation data (100 data with $N = 2, 4, 8$, respectively) and the majority of them maximize their community functions in the chain networks (i.e., biased data), I could not accurately predict when we should analyze the networks with branches using the simulation data(Fig.C.7).

C.4   Prediction in the experimental setup

The experimental conditions in Oliveira Sudário (2022) are different from what is assumed in the main text. Oliveira Sudário (2022) performed the spent medium experiment instead of multi-stage chemostats because building multi-stage chemostats had some technical issues. To fit the spent medium experiments, the dynamics are rewritten from Eqs (4.6a) and (4.6b) as follow:

$$\frac{dS_{ij}}{dt} = r_j S_{ij} \left(K_j - S_{ij} + R_i(t) \sum_{l=1}^{i-1} \sum_{k=1}^{4} a_{jk} S_{kl} \right) - d_j S_{ij} \frac{C_i}{M_j + C_i} - \alpha S_{ij} \tag{C.20a}$$

$$\frac{dC_i}{dt} = \alpha \left(C_{\text{in}} - C_i \right) - \sum_j b_j S_{ij} \frac{C_i}{C_i + M_j} \tag{C.20b}$$

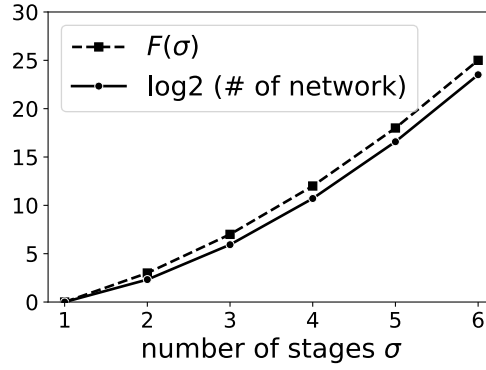$$\frac{dR_i}{dt} = \alpha(1 - R_i) \tag{C.20c}$$

Figure C.4 – Number of feasible chemostat networks with branches

The number of acyclic chemostat networks with branches exponentially increases over σ (the solid line). The increase of the networks is similar to $2^{F(\sigma)}$ (the dashed line), where $F(\sigma)$ is the following quadratic function: $F(\sigma) \equiv (\sigma + 1)(\sigma + 2)/2 - 3$.

where S_{ij} is $\log_{10}$ colony forming unit (CFU) /ml of species j in stage i, C_i is the ampicillin concentration in stage i, C_{in} is the ampicillin concentration in the inflow to stage i, r_j is the growth rate of species j, K_j is the maximum species j's $\log_{10}$ CFU (/ml) in mono-culture (i.e., carrying capacity), α_{jk} is interaction effect from species k to j (scaled so that $\alpha_{jj} = -1$ for $j = 1, \ldots, 4$), d_j is species j's maximum death rate caused by ampicillin, M_j is the concentration of ampicillin that gives the half-max death rates to species j, and b_j is the maximum detoxification rate of ampicillin by species j.

In the analysis of the simulations in the main text, I assumed $C_{\text{in}} = C_{i-1}^*$ and $C_i(0) = C_{i-1}^*$: the dynamics in stage $i - 1$ equilibrate and the initial concentration of ampicillin is given by the equilibrium state in the previous stage. These assumptions need to be relaxed to fit the experimental conditions. In the experiments, the initial medium in the second and third chemostats are the fresh minimal medium without ampicillin ($C_i(0) = 0$ if $i = 2, 3$) and the inflows are the spent media from the previous stage. Then, there are no species interaction effects at the beginning of the cultivation in the second and third chemostats. It would be straightforward to assume that the ratio of the spent medium R_i affects the intensity of species interactions from species in upstream chemostats. The ampicillin concentrations in the inflow media into stages 2 and 3 would be given by the mean ampicillin concentration in the outflow media from stages 1 and 2, respectively:

$$C_{\text{in}} = \bar{C}_{i-1} \equiv \frac{1}{T} \int_0^T C_{i-1}(t)dt, \tag{C.21}$$

where $T = 168$ is the duration of the experiment. In this case, we evaluated the community functions by comparing the inflow ampicillin concentration and the mean ampicillin concentration in the outflow medium over time. It is obvious that if we continue the experiment enough long, we recover the assumption on the inflow ($\lim_{T \to \infty} \bar{C}_{i-1} = C_{i-1}^*$).

The parameter values are estimated by updating the posterior probability distributions as
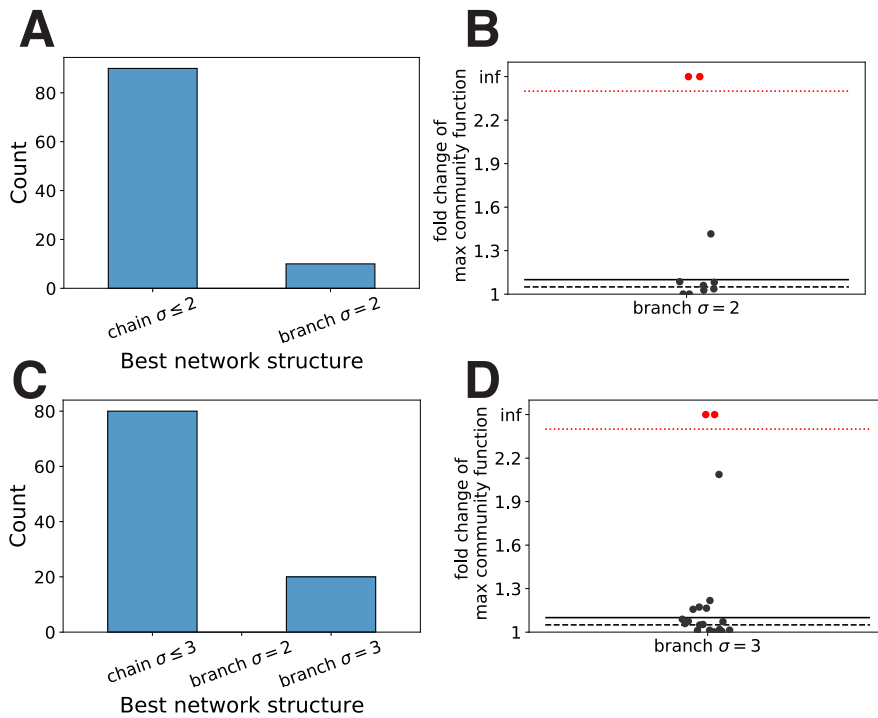
Figure C.5 – Comparing community functions in various network structures

Comparisons of community functions in networks and without branches in 100 $N = 8$-species scenarios. Top row: the maximum number of chemostats is $\sigma_m = 2$. A: about 90% of simulations (95% highest density interval is $[0.849, 0.954]$ assuming the uniform distribution as a prior probability distribution) maximized their community functions in the chain network structures (either $\sigma = 1$ or $\sigma = 2$). B: I evaluated the fold change of maximum community functions, which is the ratio of maxim community functions in the network branches to those in chain networks. This panel only shows the cases when the fold change is larger than one. The black solid line represents 1.1. and the blck dotted line represents 1.05. The two red dots over the red dotted line indicate the fold changes are infinite: in the scenario, we cannot allocate any species in the chain networks and thus the community functions in those cases are zeros. Bottom row: same analyses with $\sigma_m = 3$. C: about 80% of simulations (95% highest density interval is $[0.727, 0.871]$) maximize their community functions in the chain network structures. D: the fold change of community functions due to the introduction of branches. The black dashed and solid lines represent 1.05 and 1.1, respectively.
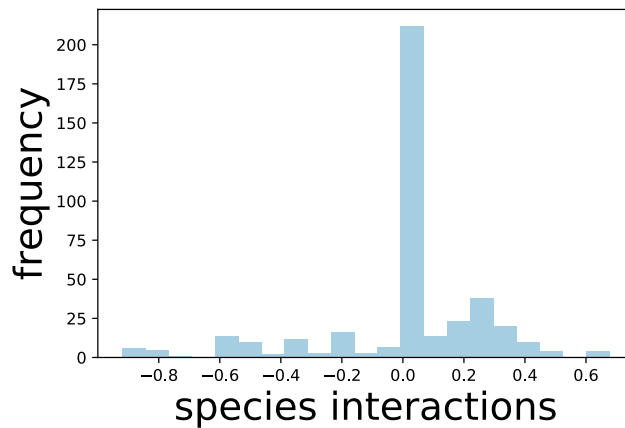
Figure C.6 – Species interactions' distribution under optimal allocations of four-species scenarios

Species interactions are defined by the difference of recipient (downstream) species' population sizes between with and without affecting (upstream) species. The species allocations are biased toward positive (one-sided Wilcoxon's signed-rank test, $p = 5.53 \times 10^{-5}$). We obtained 404 species interactions from 100 with $N = 4$ in the consumer-resource model.
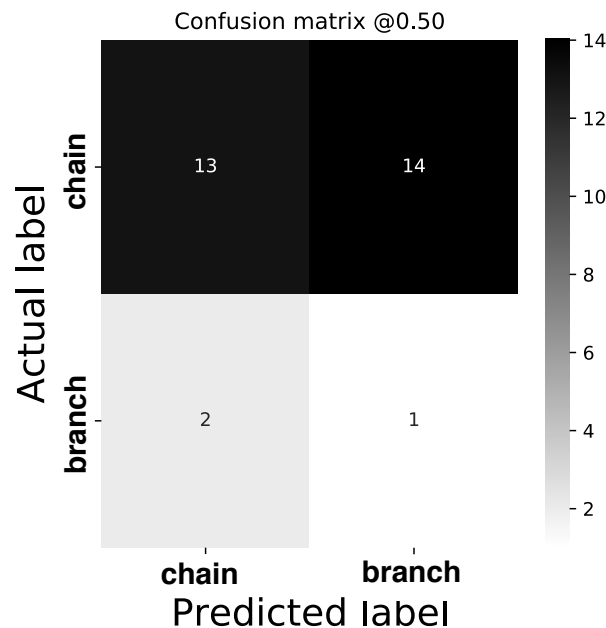
Figure C.7 – Example of graph classification results

Graph classification to predict whether we should introduce branches into the chemostat network or not, given an interaction network between species and compounds. I used three hundreds networks of communities (i.e., species-compounds and compound-community function interaction networks); 270 data were used as training data while the rest 30 were used as test data. The number in each cell represents the number of data corresponding to the actual and predicted network. Because about 90% of communities maximize their community functions in the chain network structures when $\sigma_m = 2$, we weighted the communities that maximize their community functions in the branches and tried F_1 score during the train of the classifier.

follows. First, I estimated the growth rate r_j and the carrying capacity K_j for each species using mono-culture data without ampicillin. The prior probability distributions are given by the following probability distributions:

$$r_j \sim \mathcal{U}(0, 2) \tag{C.22}$$

$$K_j \sim \mathcal{N}(\mu_j, 0.5) \tag{C.23}$$

where μ_j is the estimated value of K_j from the least square estimation using lmfit module (Newville et al., 2014). Because the population sizes converged to equilibrium around 24 hours except for *C. testosteroni*, I used the normal probability distribution as the prior probability distribution of K_j with the moderate standard deviation. On the other hand, I used the uniform distributions for the growth rates because only time points at 0 and 24 hours are informative and it is difficult to estimate the growth rate. Then, using co-culture data, I updated the posterior probability distributions of r_j and K_j, with also estimating the probability distributions of species interaction terms a_{jk} by assuming the following prior probability distributions:

$$a_{jk} \begin{cases} \sim \mathcal{U}(-1, 1) & j \neq k \\ = -1 & \text{otherwise} \end{cases} \tag{C.24}$$

The rationale for this prior probability distribution is as follows: as we scaled the intraspecific interactions as -1, it is very unlikely that the interspecific interactions have a stronger negative effect. We set the upper boundaries of the prior probability distributions as 1 to make the prior probability distribution is symmetric and centered on zero (no interaction). The results of fitting co-culture data are shown in Fig. 4.3.

To estimate how each species interacts with ampicillin, I fitted mono-culture data in the presence of ampicillin. The posterior probability distributions of r_j and K_j are given by merging the posterior distributions of them in the three co-culture experiments for each species.

After estimating posterior probability distributions of r_j, K_j, and a_{jj}, I tried to estimate the parameter values related to ampicillin (b_j, d_j, and M_j). However, due to the many missing time points, I could not use MCMC to estimate these parameter values. Instead, I used the least squares estimation (LSE) assuming that the parameters r_j, K_j, and a_{jj} are given by their means of posterior probability distributions. The residual functions were given by the sum of squared differences between the data and the model, scaled by the initial mean CFUs or initial ampicillin concentration.

Fig. C.8 shows the model prediction on how much ampicillin would be left in the outflow given each species allocation in the spent medium experiment as in Oliveira Sudário (2022). In this experimental setup, the mixtures of three or four species degrade more ampicillin than allocating single species to each chemostat. This indicates that there is no benefit to introducing the spatial structures, which contradicts the results in the main text (Fig 4.6). This difference of the results between the experimental setup in Oliveira Sudário (2022) and the ideal one would be

caused by the fact that the media are replenished in each chemostat in Oliveira Sudário (2022). This protocol weakens the species interaction effects from species in upstream chemostats while diluting the concentration of the ampicillin. Then, the mixtures of the species continues to grow in the second and third chemostats and degrade more ampicillin. In contrast, the ideal condition does not allow the mixtures of species to grow in the second and third chemostats once the dynamics in the first chemostat equilibriate.

Although the efficiency of degrading ampicillin depends on the details of experiments, the ranks of the community functions are robust. I compared the rank orders of the allocations in our experimental setup and in the ideal conditions (Fig. C.9). When I fix the number of chemostats, the rank orders are similar between the two conditions. Therefore, the allocations that degrade large amounts of ampicillin in the experimental conditions of Oliveira Sudário (2022) would also be candidates for the best allocation in the ideal condition.
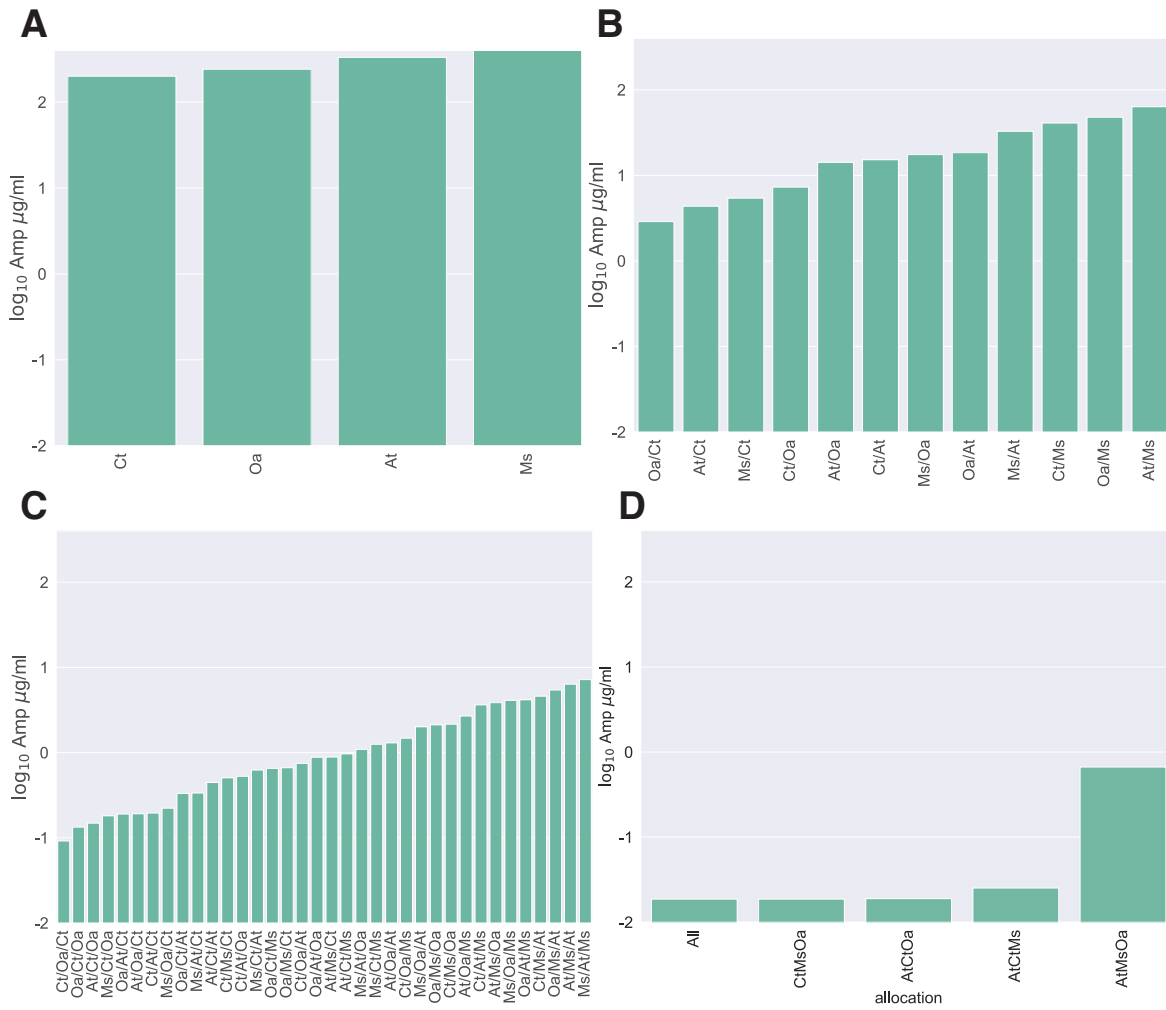
Figure C.8 – Estimated ranks of detoxification in the experimental setup

The predicted concentrations of ampicillin left in the outflow media in the experimental setup. A: $\sigma = 1$, B: $\sigma = 2$, C: $\sigma = 3$, and D: mixtures of three or four species with $\sigma = 3$. Allocation X/Y/Z means we allocate species X to the first chemostat, Y to the second, and Z to the third. In panel C, there are some allocations that do no appear in the ideal set up but appear in the experimental setup. Such allocations are ranked 29 in the ideal setup. At: *A. tumefaciens*, Ct: *C. testosteroni*, Ms: *M. saperdae*, Oa: *O. anthropi*.
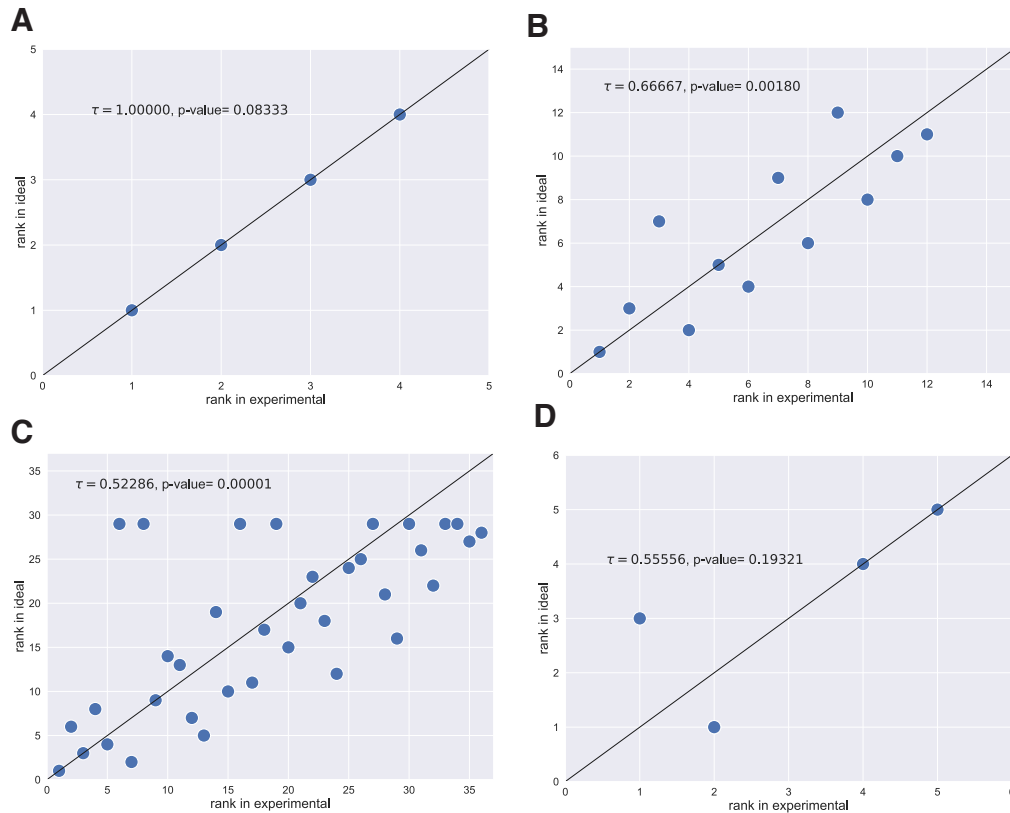
Figure C.9 – Correlations in the two experimental setups

The correlations between the predicted ranks of allocations in the experimental setup (Fig. C.8) and in the ideal setup (Fig. 4.6). The results of Kendall coefficient correlation tests (coefficient τ and p-value) are shown in in each panel. A: $\sigma = 1$, B: $\sigma = 2$, C: $\sigma = 3$, and D: mixtures of three or four species with $\sigma = 1$ because species cannot be allocated to stages $i = 2, 3$ in the ideal condition. The black lines show that the ranks in the two setups are identical.

Appendix D

Appendices of Chapter 5

The appendices will be uploaded on bioRxiv before the private defence.

Supplemental information

SI 1   The details of the model

SI 1.1   Derivation

In the main text, we use a generalized Lotka-Volterra (gLV) model. This model can describe the community dynamics with a small number of parameters but gLV models may not be easy to biologically interpret, especially for non-theoretical biologists. In this section, we explain the derivation of the model so that our results could be easily tested with empirical data.

First, we consider the following Lotka-Volterra model, which would be easier to fit the experimental data than Eqs (1a) and (1b):

$$\frac{dx_{i1}}{dt} = \mu_{i1}x_{i1}\left(K_{i1} + \sum_{k=1}^{N} A_{ik}x_{k1}\right) \tag{S1a}$$

$$\frac{dx_{i2}}{dt} = \mu_{i2}x_{i2}\left\{K_{i2} + \sum_{k=1}^{N} A_{ik}(x_{k1} + x_{k2})\right\} \tag{S1b}$$

where K_{ij} is the maximum population size of species i in patch j in mono-culture, $\mu_{ij}K_{ij}$ represents the maximum *per capita* growth rate in mono-culture of species i in patch j, and A_{ik} is the species interaction from species k to i with $A_{ii} = -1$ for all i. These ordinary differential equations are rewritten as follow

$$\frac{dx_{i1}}{dt} = x_{i1}\left(r_{i1} + \sum_{k=1}^{N} \hat{a}_{ik}^{(1)}x_{k1}\right) \tag{S2a}$$

$$\frac{dx_{i2}}{dt} = x_{i2}\left\{r_{i2} + \sum_{k=1}^{N} \hat{a}_{ik}^{(2)}(x_{k1} + x_{k2})\right\} \tag{S2b}$$

where

$$r_{ij} = \mu_{ij}K_{ij}, \tag{S3}$$

$$\hat{a}_{ik}^{(j)} = \mu_{ij}A_{ik} \tag{S4}$$

As the main text assumes $\vec{r}_1 \neq \vec{r}_2$ and $a_{ik}^{(1)} = a_{ik}^{(2)}$, the patches are assumed to change the maximum population size in mono-culture K_{ij} but not growth rate parameter μ_{ij}. If the growth parameter μ_{ij} differs over the patches, we need to reformulate Eqs (1a) and (1b) using $\hat{a}_{ik}^{(1)}$ and $\hat{a}_{ik}^{(2)}$ instead of a_{ik}. In addition, as we assume $a_{ii} = -1$, we considered the special case of $\mu_{ij} = 1$ in the main text.
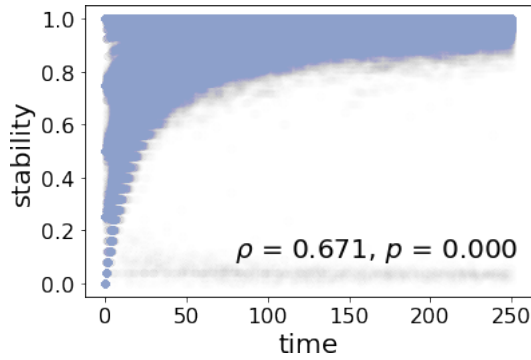
Figure S1: Stability in the assembly scenario over time

Stability of the downstream communities in the assembly scenario over time. Spearman correlation coefficient: 0.671 and $p < 10^{-3}$.

SI 1.2 Sampling parameter values

To see the relationship between the stability and the features of the upstream and downstream communities, we sampled the parameter value from the following probability distributions

$$r_{ij} \sim U(0, 2) \tag{S5}$$

$$a_{ik} \begin{cases} = -1 & \text{for } i = k \\ \sim \mathcal{N}(0, 0.25^2) & \text{with probability } C \text{ for } i \neq k \\ = 0 & \text{otherwise} \end{cases} \tag{S6}$$

where $U(a, b)$ represents the uniform distribution from a to b, $\mathcal{N}(\mu, \sigma^2)$ represents the normal distribution whose mean and standard deviation are μ and σ, respectively, and C represents the connectance of species interaction matrix. We assumed that all species can grow in mono-culture ($r_{ij} > 0$) and that the strength of species interactions is unlikely to be stronger than that of intraspecific interactions ($\sigma = 0.25$). We set $C = 0.25, 0.5, 1.0$ and sampled 20 sets of $\vec{r}_1$, $\vec{r}_2$, and a_{ik} for each C.

For each set of parameters related to the species interactions and the growth rates, we calculated the stability in the two scenarios – assembly and design – over the parameters related to migration: $\rho = 0.0, 0.25, 0.5, 0.75, 1.0$ and $\mu = 0.0, 0.25, 0.5, 1.0, 2.0$. In the stability analysis of the assembly scenario, the downstream stability positively correlates with the time steps the meta-communities were sampled (Fig. S1, Spearman correlation: 0.671 and $p < 10^{-3}$). In the design scenario, the majority of sampled upstream communities cannot coexist with the target downstream communities: one or more species go extinct in either or both of the upstream and downstream communities before the dynamics stabilized (Fig. S2). See SI 3 for the discussion of the coexistence in the meta-communities.
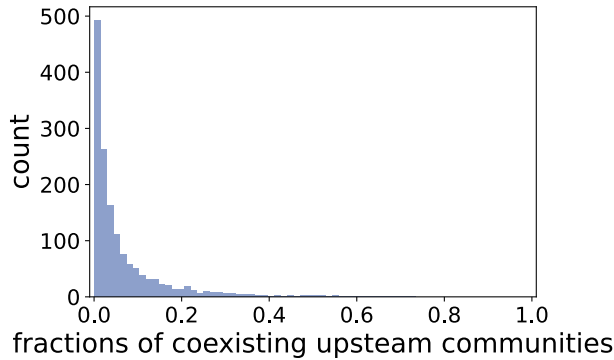
Figure S2: Fractions of upstream communities coexisting with the downstream

Fractions of upstream communities coexisting with the target downstream communities. For each of 1,500 target downstream communities, we randomly generated 200 upstream communities.

SI 2   Statistical analysis without the spatial structure

As a control scenario, we analyzed the cases where we have no spatial structure in this section. We have only one community in this case. e generated such communities using the target downstream communities in the designing scenario. Now, the species always migrate from the outside of the community and thus the stability only represents the resistance to invasion. WIn this case, species richness has the largest positive correlation with the stability (Fig. S3), which support the findings in the previous studies (Case, 1990; Stachowicz et al., 1999; Bonanomi et al., 2014; Hromada et al., 2021).

We also performed the causal inference in this scenario. We assumed the causal relation similar to the main text: species richness increases the total strength of positive and negative interactions, respectively, and these interactions relate to the stability. To satisfy the backdoor criteria, the logistic regression model includes only the richness to see the effect of richness, while we included richness, the total strength of positive interactions, and the total strength of negative interactions to the see effects of the total strength of positive and negative interactions, respectively, on the stability. Table S1 shows that only species richness has a significant effect on the stability.

These results can be used to consider the upstream stability in the spatial structure we considered in the main text. As the upstream patch is not affected by the spatial structure, we can infer that the upstream stability is high when the upstream richness is large.

Table S1: Causal inference without spatial structure

Feature	Coefficient	Standard error	P-value
Richness	0.2813	0.128	0.028
Total_positive	0.1357	0.293	0.643
Total_negative	−0.1733	0.302	0.567

In the main text, we analyzed the downstream stability while the result of this section infer the upstream stability. If we need to consider the stability of both the upstream and downstream communities (i.e., meta-community stability), we would just modify the definition of the stability so that we consider the resistance
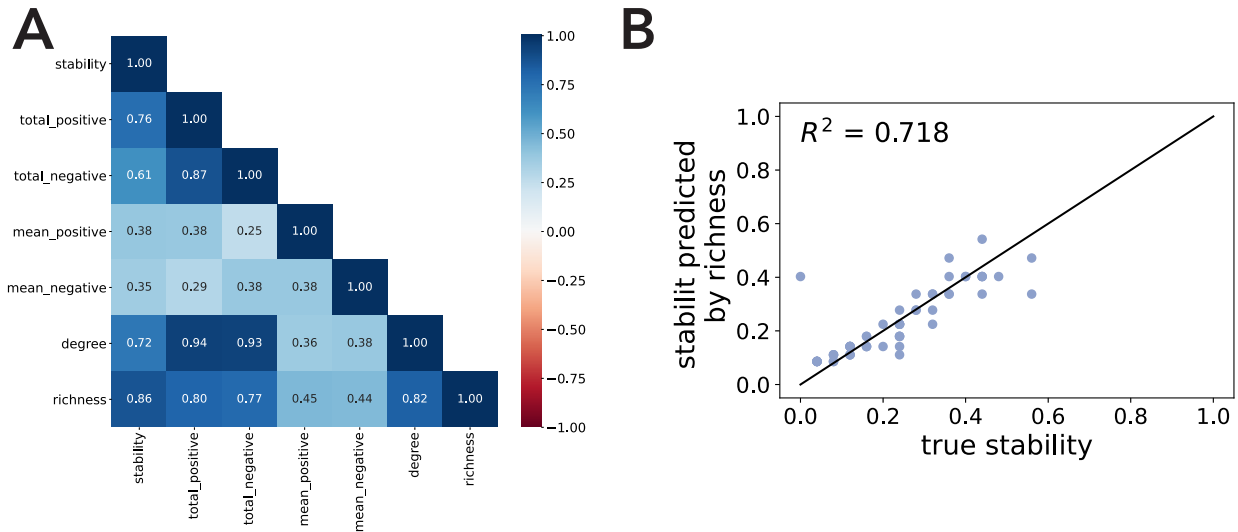
Figure S3: Analysis without spatial structure

A: Pearson correlations in the absence of the spatial structure. The darker blue represents the stronger positive correlation while the darker red indicates the stronger negative correlation. The stability (resistance to invasion) has the largest positive correlation with species richness. B: comparison of the true stability with the prediction by the logistic regression model with the species richness. R^2 between the true stability and the predicted stability is 0.718.

to invasion in the upstream and downstream communities. We do not have to consider the resistance to the environmental changes caused by upstream communities in such cases because the environmental changes do not occur unless invaders colonize the upstream community. In this sense, the downstream stability in the main text can be seen as an upper boundary of the meta-community stability.

SI 3   Coexistence conditions

In the designing scenario, we analyzed the meta-communities where (i) the target downstream communities are feasible and locally stable in the absence of the upstream community, (ii) the accompanying upstream communities are also feasible and locally stable, and (iii) the upstream communities do not drive one or more species extinction in the downstream. Such meta-communities should satisfy the following three necessary conditions

- Hierarchy: because the downstream dynamics stabilize after the stabilization of the upstream dynamics, all species in the downstream should have positive abundances until then. In other words, the upstream community should stabilize enough fast so that no species in the downstream goes extinct.

- Feasibility: after the upstream community stabilizes, we should check whether the downstream dynamics have an equilibrium where all downstream species can coexist. As the upstream species affect the growth rates of the downstream species, such feasible equilibrium may not exist in some cases.

- Convergence: if the downstream community has a feasible equilibrium in the presence of the upstream community, we need to know whether the downstream dynamics converge to the feasible equilibrium. In
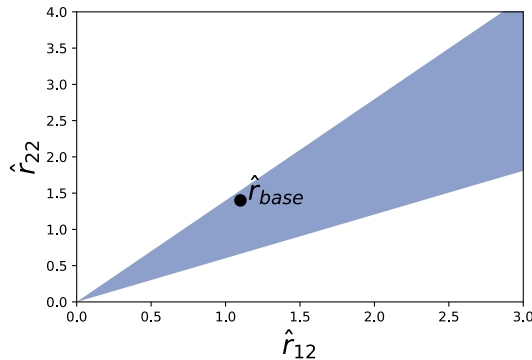
Figure S4: Region of allowing two competitive species coexistence

The colored area represents the region of $(\hat{r}_{12}, \hat{r}_{22})$ that enable two competing species to coexist in the downstream patch. The black dot represents the growth rates of the two species when no species exist in the upstream patch. The upstream community changes the two species' growth rates $(\hat{r}_{12}, \hat{r}_{22})$ from the basal growth rates. Parameter values: $a_{11} = -1$, $a_{12} = -0.715$, $a_{21} = -0.603$, $a_{22} = -1$, $\hat{r}_{12} = 1.1$, and $\hat{r}_{22} = 1.4$. Therefore, this example indicates that species 1 grows slower and less competitive than specie 2.

this case, the local stability may not be sufficient, because the effects from the upstream community can drive the downstream community far from the equilibrium. If the feasible equilibrium is globally stable, on the other hand, the downstream community converges to the equilibrium.

Analyzing these three conditions, in general, would be very difficult. However, if we assume that the initial condition of the upstream community is very close to the equilibrium, we can omit the first condition. In addition, we can analyze the second and third conditions (feasibility and convergence) by assuming that all species interactions in the downstream community are negative (i.e., a competitive community).

Let us start with a simple downstream: we have only two species there and they are competitive ($a_{ik} < 0$ for $i = 1, 2$ and $k = 1, 2$). Assuming that the upstream community is stabilized, we consider the generalized Lotka-Volterra equation given by Eq (4). It is widely known that the feasible equilibrium is globally stable in this classical example if and only if (Saavedra et al., 2017; Godwin et al., 2020)

$$
\underbrace{\sqrt{\frac{a_{11}a_{22}}{a_{12}a_{21}}}}_{\text{Niche overlap}^{-1}} > \underbrace{\frac{\hat{r}_{12}}{\hat{r}_{22}}\sqrt{\frac{a_{22}a_{21}}{a_{11}a_{12}}}}_{\text{fitness difference}} > \underbrace{\sqrt{\frac{a_{12}a_{21}}{a_{11}a_{22}}}}_{\text{Niche overlap}} , \tag{S7}
$$

$\hat{r}_{21}, \hat{r}_{22} > 0$, and $0 < a_{12}a_{21} < a_{11}a_{22}$. In the modern coexistence theory (Chesson, 2000), inequalities (S7) are understood by niche overlap and fitness difference. Recall that the effects from upstream community is summarized in $\hat{r}_{12}$ and $\hat{r}_{22}$, see Eq (5). Then, the upstream community only changes the fitness difference and we can derive the conditions where the upstream community enables the two species coexist in the downstream (Fig S4).

If we have three or more species in the downstream patch and there are only negative interactions in the downstream community, we can use the structural approach suggested by Saavedra et al. (2017). Again the upstream community's effects appear in the fitness difference but not in the niche overlap. Importantly, the fitness difference and the niche overlap defined by Saavedra et al. (2017) show only the feasible condition. The
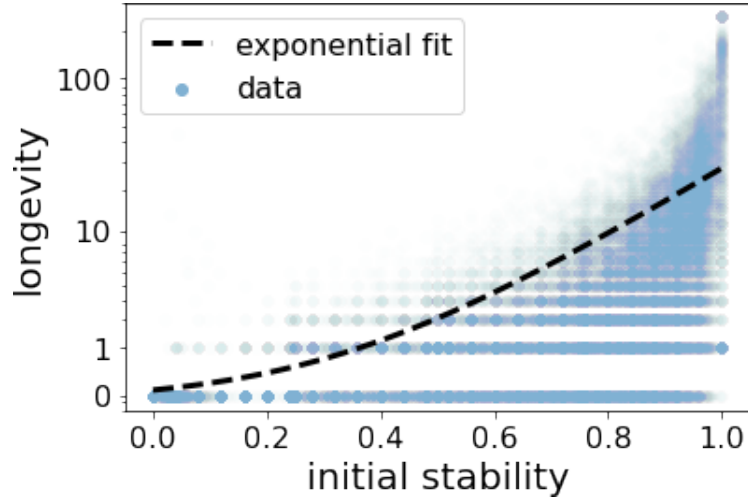
Figure S5: Stability explains the longevity of species compositions

Each blue dot represents the downstream community in the assembly scenarios ($n = 37998$). The initial stability of the downstream community depends on the downstream community as well as the accompanying upstream community when the focal downstream community is composed. Then, we calculated the duration that the downstream species composition is maintained. The black dashed line represents the fitting to the following exponential function: longevity $= 1.780 \times 10^{-1} \times \exp(5 \times \text{initial stability}) + 5.141 \times 10^{-16}$. The coefficients of this function were obtained by scipy.optimize.curve_fit.

sufficient condition for the global stability of the equilibrium is that the interaction matrix in the downstream community is Volterra-dissipative. See Appendix 3 of Saavedra et al. (2017) for more mathematical details.

SI 4   Longevity of downstream communities

The data from assembly scenario were used to measure how long a focal downstream species composition is maintained (Fig. S5). As discussed in the main text, the downstream species composition depends on upstream and downstream communities: the downstream stability can change when it's upstream community changes the species composition. However, if the initial stability (the stability that the focal downstream species composition is generated) is large, the focal downstream species composition is maintained.

SI 5   Prediction analysis

In this section, we build some logistic regression models to predict the stability because measuring the stability in experiments is difficult. Here, we used the data in the designing scenario to train the statistical models because the assembly scenario does not randomly sample the meta-communities. However, we used sub-sampled data of the assembled to see the performance of the statistical prediction. As the stability in the assembly scenario is biased, we sub-sampled 100 data of stability $[0.1 \times i, 0.1 \times (i + 1)]$ fro $i = 0, \ldots, 9$.

First, we justify the choices of community features. We analyzed species richness in the upstream and downstream patches, respectively, because the previous studies suggest that species richness affects the resistance to invasion (Case, 1990; Stachowicz et al., 1999; Bonanomi et al., 2014; Hromada et al., 2021), which is also

supported by our analysis without the spatial structure (SI 2). In addition, because the resistance to the environmental changes is conceptually similar to structural stability, we measured the total strength of positive or negative interactions and the mean degree (i.e., how many species each species interact with), which can correlate with structural stability (Portillo et al., 2021). We, therefore, evaluated the total strength of positive or negative interspecific interactions within the upstream and downstream communities, respectively, the total strength of positive or negative interactions from the upstream to the downstream communities including intraspecific interactions, mean degree within and upstream and downstream communities, respectively, and mean degree between the two communities. We also quantified the mean strength of positive or negative interaction within each patch, and from the upstream to the downstream, respectively. The explanations of features are summarized in Table 1. We emphasize that the models in this section include features that are ignored in the causal inference. This is because certain features may be useful to predict the stability because of correlation, although they are unlikely to have causal effects on the stability.

After calculating these 17 community features, we investigated the best logistic regression model to predict the downstream community's stability using the community features and the two migration parameter values (ρ and μ). The brute force search for the best predictive model is, however, not feasible because we have to analyze $2^{19} - 1 \approx 500,000$ logistic regression models. Instead, we generated some models based on the step-wise algorithms. First, we considered the null model: i.e., the model has only the intercept term (Table S2). This model can be seen as a baseline. Second, we generated a logistic regression model using the forward selection: from the null model, we introduced a single feature into the model that increased McFaddens pseudo-R squared the most and repeated this process until the pseudo-R squared did not increase. We call this model the forward model 1 in this section (Table S3). In addition, we can consider another baseline using the full model1, which has all 19 features and the intercept (Table S4). Then we performed backward selection: we removed each feature from the full model 1 that had the highest p-value and repeated this process until all remaining features had p-values smaller than 0.05. We call this logistic regression model the backward model (Table S5). Then, we calculated the variance inflating factors (VIFs) of the three models – the full model1, the forward model1, and the backward model – using the statsmodels package because VIFs > 5 can cause collinearity (James et al., 2013). Indeed, the full model and the forward model1 have large VIFs while the backward model has small VIFs. We generated additional two models, the full model 2 (Table S6) and the forward model 2 (Table S7), by removing features that have the highest VIF in the full model 1 and the forward model 1, respectively, and continued removing features with the largest VIF until all features in the model had VIFs less than 5.

Within these six logistic regression models – the null model, the full model 1, the full model 2, the forward model 1, the forward model 2, and the backward model– we chose the best predictive model based on Akaike information criteria (AIC). AIC suggested that the backward model is the best model to predict the stability (Table S8). Fig. S6 visualizes the comparisons of the true stability with the predicted stability by the backward model using the data from the designing scenario (A) and the assembly scenario (B). Because we have a small number of data with stability < 0.4, the model cannot predict such small stability. For the other models. see Fig. S7; the models except for the null model has similar predictability to the backward model (Fig. S6).
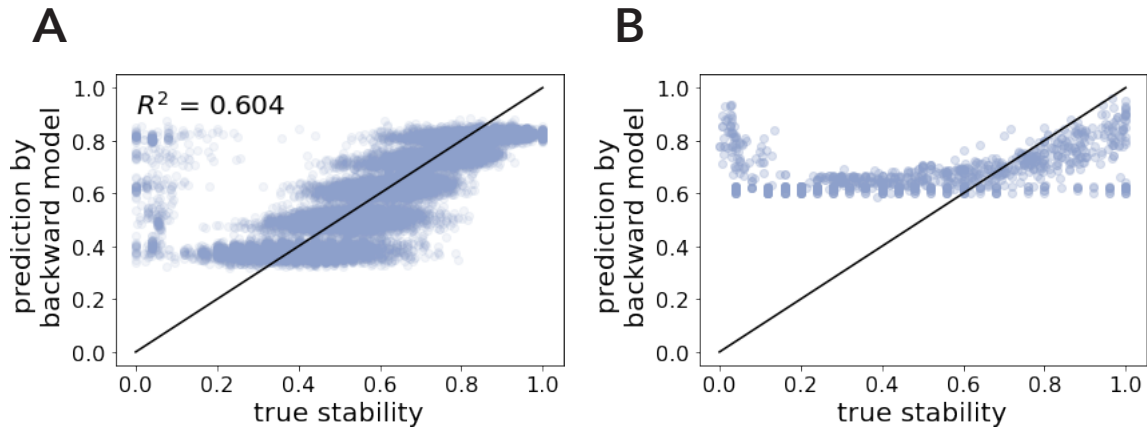
Figure S6: Predicting the stability by the backward model

The prediction of the stability of the downstream after training the backward model with the designing data. A: the competition of the true stability with the prediction by the backward model. The R-squared between the true stability and the estimated stability is 0.604. B: The same analysis but using the sub-sampled data from the assembly scenario so that we have 100 data where stability is $[0.1 \times i, 0.1 \times (i+1)]$, for each $i = 0, \ldots 9$. In both panels, the black diagonal lines represent that the model prediction perfectly matches the true stability

We emphasize that the backward model is the best predictive model, but this model does not suggest causality: we could not tell whether stability increases/decreases or not by changing some features of the communities. This is because AIC is designed to select a model with high predictability.

Table S2: Null model

Feature	Coefficient	Standard error	P-value
Intercept	0.4611	0.015	$< 10^{-3}$

SI 6   Details of causal inference

In the main text, we only show the effect of each controllable feature on the stability in Table 2. As these results come from various regression models, we summarize the results of the regression models in this section. Although the full model of the causal diagram is given in Fig. 3, we show the simpler network for convenience: we excluded the resistance to invasion and environmental changes, degrees, mean strength of positive/negative interactions, and species interactions within the downstream communities (Fig. S8). This network suggests which feature we should include in the logistic regression analysis to satisfy the backdoor criteria. When we analyze the effect of the species richness in the upstream communities, the logistic regression model should exclude the total strength of positive/negative interactions within the upstream community and those from the upstream to the downstream. In contrast, when we analyze the total strength of positive/negative interactions within the upstream community and those from the upstream to the downstream, the logistic regression model should include the species richness in the upstream community. Table S9 summarizes the coefficients and the p-values in each logistic regression. Note that we can also infer the effects on the stability from the species richness in the downstream community, ρ, and μ in the three logistic regression models in Table S9, where the
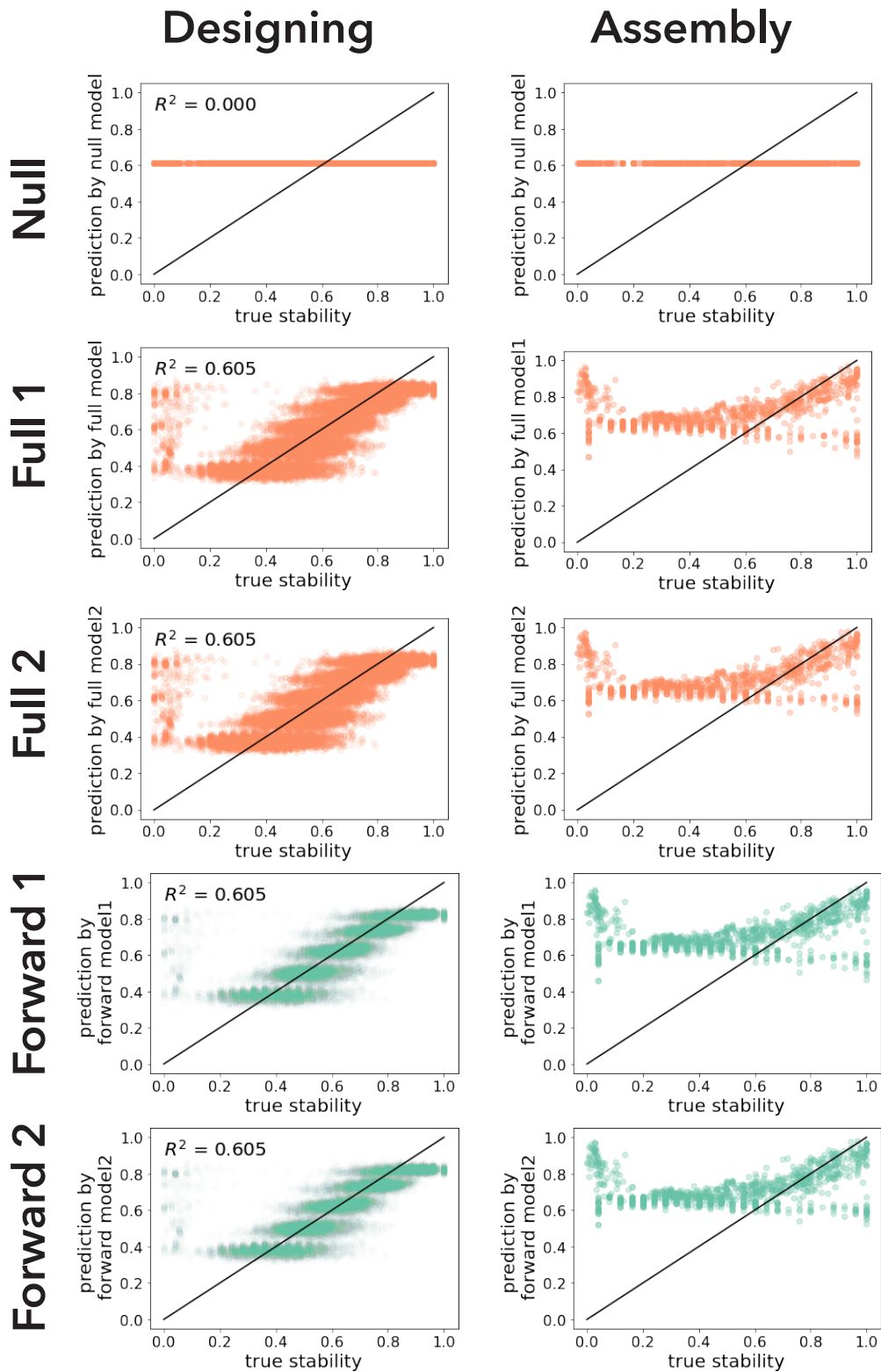
Figure S7: Prediction by the other logistic regression models

Predictions of the stability by each model are summarized. Each row corresponds to each model. The left column shows the comparisons between the true stability in the designing scenario and the prediction by each model. R^2s between the true stability and the predicted one are shown on the top left of each panel. The right panel shows the same analysis for the sub-sampled data from the assembly scenario so that we have 100 data where stability is $[0.1 \times i, 0.1 \times (i+1)]$, for each $i = 0, \ldots 9$. In each panel, the black diagonal lines represent that the model prediction perfectly matches the true stability.

Feature	Coefficient	Standard error	P-value
Intercept	0.5188	0.016	$< 10^{-3}$
Total_positive_up	0.0180	0.040	0.729
Mean_positive_up	0.0132	0.016	0.7456
Total_negative_up	0.0493	0.039	0.203
Mean_negative_up	-0.0026	0.017	0.880
Degree_up	0.0700	0.054	0.195
Richness_up	0.0055	0.016	0.733
Total_positive_down	-0.0031	0.030	0.917
Total_negative_down	-0.0238	0.027	0.375
Richness_down	0.0132	0.016	0.412
Total_positive_trans	0.0661	0.032	0.040
Mean_positive_trans	0.0027	0.018	0.880
Total_negative_trans	0.0378	0.028	0.183
Mean_negative_trans	-0.0160	0.018	0.366
Degree_trans	-0.1388	0.036	$< 10^{-3}$
ρ	0.7113	0.017	$< 10^{-3}$
μ	-0.0603	0.016	$< 10^{-3}$

Table S3: Forward model 1
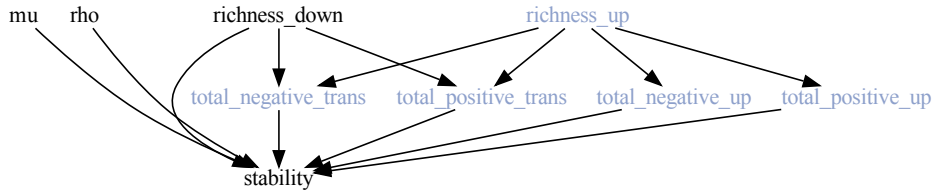
Figure S8: Simplified causal relations

coefficient of each feature is almost identical across the three models. We can see that the ρ has the largest effect on the stability in the analyzed features.

We also inferred the causal effects of the total strength of positive and negative interactions in the downstream patch, respectively. The logistic regression model in this case included the species richness in the upstream and downstream, respectively, the total strength of positive or negative interactions in the downstream, respectively, ρ and μ (Table S10). Although the total strength of negative interaction in the downstream is not statistically significant, the coefficient is as large as the total strength of positive interactions from the upstream to the downstream communities. We could, there, except that the negative interactions (e.g., competitions) in the downstream patch increase the stability. This result is consistent with the previous studies (Case, 1990; Stachowicz et al., 1999; Bonanomi et al., 2014; Hromada et al., 2021), where species-rich communities have high resistance to invasion due to resource competition (Mallon et al., 2015)) because resource competition is implemented as negative interactions in this study.

Table S4: Full model 1

Feature	Coefficient	Standard error	P-value
Intercept	0.5189	0.016	$< 10^{-3}$
Total_positive_up	0.0117	0.041	0.775
Mean_positive_up	0.0133	0.018	0.450
Total_negative_up	0.0486	0.039	0.214
Mean_negative_up	-0.0022	0.017	0.898
Degree_up	0.0722	0.056	0.198
Richness_up	0.0057	0.016	0.723
Total_positive_down	-0.0067	0.036	0.852
Mean_positive_down	0.0058	0.023	0.799
Total_negative_down	-0.0129	0.035	0.710
Mean_negative_down	-0.0119	0.021	0.576
Degree_down	-0.0085	0.048	0.859
Richness_down	0.0131	0.016	0.414
Total_positive_trans	0.0692	0.034	0.042
Mean_positive_trans	0.0024	0.018	0.895
Total_negative_trans	0.0400	0.030	0.179
Mean_negative_trans	-0.0165	0.018	0.355
Degree_trans	-0.1395	0.036	$< 10^{-3}$
ρ	0.7111	0.017	$< 10^{-3}$
μ	-0.0604	0.016	$< 10^{-3}$

Table S5: Backward model

Feature	Coefficient	Standard error	P-value
Intercept	0.5189	0.016	$< 10^{-3}$
Total_negative_up	0.0953	0.025	$< 10^{-3}$
Total_positive_trans	0.0755	0.018	$< 10^{-3}$
Degree_trans	-0.0944	0.026	$< 10^{-3}$
ρ	0.7134	0.017	$< 10^{-3}$
μ	-0.0587	0.016	$< 10^{-3}$

Table S6: Full model 2

Feature	Coefficient	Standard error	P-value
Intercept	0.5188	0.016	$< 10^{-3}$
Total_positive_up	0.0372	0.035	0.293
Mean_positive_up	0.0089	0.017	0.608
Total_negative_up	0.0719	0.035	0.038
Mean_negative_up	-0.0064	0.017	0.709
Richness_up	0.0054	0.016	0.739
Total_positive_down	-0.0059	0.033	0.857
Mean_positive_down	0.0062	0.022	0.779
Total_negative_down	-0.0111	0.031	0.718
Mean_negative_down	-0.0127	0.021	0.551
Richness_down	0.0141	0.016	0.378
Total_positive_trans	0.0675	0.032	0.038
Mean_positive_trans	0.0010	0.018	0.956
Total_negative_trans	0.0308	0.028	0.279
Mean_negative_trans	-0.0171	0.018	0.336
Degree_trans	-0.1120	0.029	$< 10^{-3}$
ρ	0.7113	0.017	$< 10^{-3}$
μ	-0.0600	0.016	$< 10^{-3}$

Table S7: Forward model 2

Feature	Coefficient	Standard error	P-value
Intercept	0.5187	0.016	$< 10^{-3}$
Total_positive_up	0.01379	0.035	0.283
Mean_positive_up	0.0088	0.017	0.610
Total_negative_up	0.0718	0.035	0.038
Mean_negative_up	-0.0066	0.017	0.698
Richness_up	0.0052	0.016	0.749
Total_positive_down	0.0010	0.030	0.974
Total_negative_down	-0.0199	0.027	0.454
Richness_down	0.0142	0.016	0.376
Total_positive_trans	0.0664	0.032	0.039
Mean_positive_trans	0.0009	0.018	0.960
Total_negative_trans	0.0300	0.028	0.279
Mean_negative_trans	-0.0170	0.018	0.326
Degree_trans	-0.1116	0.029	$< 10^{-3}$
ρ	0.7115	0.017	$< 10^{-3}$
μ	-0.0600	0.016	$< 10^{-3}$

Table S8: Model selection

Model	ΔAIC
Backward	0.000
Forward 1	5.212
Forward 2	7.728
Full 1	11.2289
Full 2	11.580
Null	3283.415

Table S9: Coefficients in causal inference upstream

feature(s) of interest	intercept	richness_up	total_positive_up	total_negative_up	total_positive_trans	total_negative_trans	richness_down	ρ	μ
richness_up	$0.518\ (<10^{-3})$	$0.006\ (0.73)$	N.A.	N.A.	N.A.	N.A.	$0.017\ (0.29)$	$0.708\ (<10^{-3})$	$-0.059\ (<10^{-3})$
total_positive/negative_up	$0.518\ (<10^{-3})$	$0.006\ (0.73)$	$0.019\ (0.50)$	$0.032\ (0.27)$	N.A.	N.A.	$0.013\ (0.43)$	$0.709\ (<10^{-3})$	$-0.058\ (<10^{-3})$
total_positive/negative_trans	$0.518\ (<10^{-3})$	$0.005\ (0.76)$	N.A.	N.A.	$0.049\ (0.03)$	$0.019\ (0.39)$	$0.013\ (0.41)$	$0.710\ (<10^{-3})$	$-0.059\ (<10^{-3})$

In each element, the values outside the parentheses indicate the coefficients in the logistic regression while the values inside the parentheses show the p-values. N.A. represents "not applicable."

Table S10: Coefficients in causal inference downstream

feature(s) of interest	intercept	richness_up	richness_down	total_positive_down	total_negative_down	ρ	μ
total_positive/negative_down	$0.518\ (<10^{-3})$	$0.006\ (0.73)$	$0.015\ (0.34)$	$-0.001\ (0.97)$	$0.045\ (0.07)$	$0.709\ (<10^{-3})$	$-0.058\ (<10^{-3})$

SI 7   Additional manipulation of positive interactions

In the main text, we manipulated the strength of the positive effects from upstream species to downstream species to clarify whether these species interactions increase the downstream stability (Figs. 4C and D). However, to balance the mean of the off-diagonal elements of the interaction matrix, all interactions from downstream species to upstream species are assumed negative, and the strength of these negative interactions increase while the positive interactions from upstream species to downstream species get stronger. In this case, the stability would increase over the strength of positive interactions from upstream species to downstream species because upstream species receive stronger negative effects and cannot colonize when they invade the downstream patch (i.e., larger resistance to invasion), but not because of the positive effects from upstream species to downstream species.

To remove this bias, we increased the strength of the positive interactions from upstream species to downstream species while fixing species interaction coefficients from downstream species to upstream species zeros. In this case, upstream species can colonize the downstream patch when their basal growth rates in the downstream patch are larger than in the upstream patch ($r_{i2} > r_{i1}$). In other words, the stability in this scenario cannot increase because of negative interactions that invaders from the upstream patch receive in the downstream patch.

In this case, some simulations shown negative or no changes in the stability over the strength of positive interactions from upstream species to downstream species (Fig. S9A). This results indicates that the increase of negative interactions from downstream species to upstream species increases the stability in Figs. 4C and D. However, about half simulations (52 out of 110 simulations) show positive correlations between the strength of the positive interactions and the stability (Fig. S9B). In addition, one-sided Wilcoxon signed-rank test suggests that the correlation coefficients are biased toward positive ($T = 1620$, $p = 0.042$). These results support our conclusion that the positive interactions from upstream species to downstream species stabilize the downstream community.
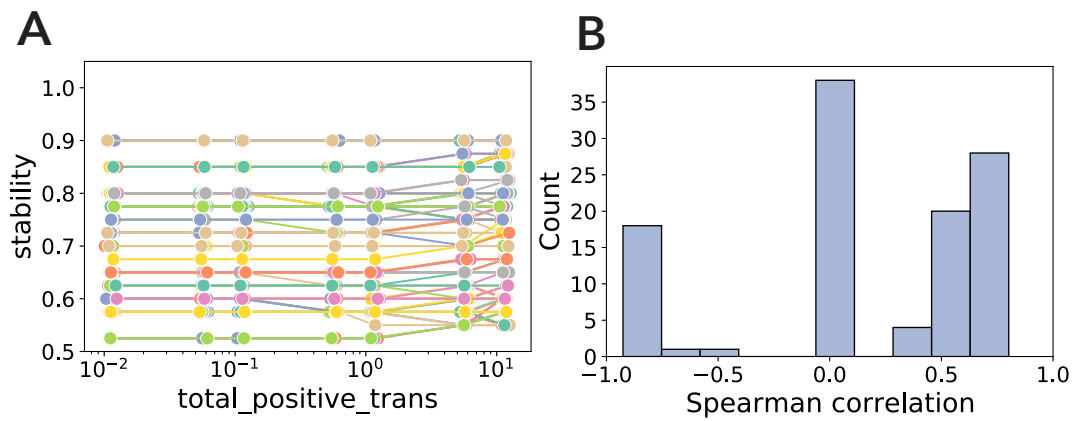
Figure S9: Manipulation of positive effects while fixing negative effects

We manipulate the strength of positive interactions from upstream species to downstream species while fixing the strength of the interactions in the opposite direction zero. A: We plotted the downstream stability over the strength of positive interactions from upstream species to downstream species. Each line corresponds to one of 110 meta-communities. B: The histogram of correlation coefficients between the strength of the positive interactions and the stability in panel A are shown. 52 meta-communities show positive correlations, 20 show negative correlations, and 38 show no correlations.

Bibliography

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL `https://www.tensorflow.org/`. Software available from tensorflow.org.

Abriouel, H., Martín-Platero, A., Maqueda, M., Valdivia, E., and Martínez-Bueno, M. Biodiversity of the microbial community in a Spanish farmhouse cheese as revealed by culture-dependent and culture-independent methods. *International Journal of Food Microbiology*, 127(3):200–208, 10 2008. ISSN 01681605. doi: 10.1016/j.ijfoodmicro.2008.07.004. URL `https://linkinghub.elsevier.com/retrieve/pii/S0168160508003723`.

Alfadaly, R. A., Elsayed, A., Hassan, R. Y. A., Noureldeen, A., Darwish, H., and Gebreil, A. S. Microbial Sensing and Removal of Heavy Metals: Bioelectrochemical Detection and Removal of Chromium(VI) and Cadmium(II). *Molecules*, 26(9):2549, 4 2021. ISSN 1420-3049. doi: 10.3390/molecules26092549. URL `https://www.mdpi.com/1420-3049/26/9/2549`.

Amarasekare, P. The evolution of coexistence theory. *Theoretical Population Biology*, 133:49–51, 2019. ISSN 10960325. doi: 10.1016/j.tpb.2019.09.005. URL `https://doi.org/10.1016/j.tpb.2019.09.005`.

Anderson, R. T., Vrionis, H. A., Ortiz-Bernad, I., Resch, C. T., Long, P. E., Dayvault, R., Karp, K., Marutzky, S., Metzler, D. R., Peacock, A., White, D. C., Lowe, M., and Lovley, D. R. Stimulating the In Situ Activity of Geobacter Species To Remove Uranium from the Groundwater of a Uranium-Contaminated Aquifer. *Applied and Environmental Microbiology*, 69(10):5884–5891, 10 2003. ISSN 0099-2240. doi: 10.1128/AEM.69.10.5884-5891.2003. URL `https://journals.asm.org/doi/10.1128/AEM.69.10.5884-5891.2003`.

Angulo, M. T., Moog, C. H., and Liu, Y. Y. A theoretical framework for controlling complex microbial communities. *Nature Communications*, 10(1):1–12, 2019. ISSN 20411723. doi: 10.1038/s41467-019-08890-y. URL `http://dx.doi.org/10.1038/s41467-019-08890-y`.

Antoni, D., Zverlov, V. V., and Schwarz, W. H. Biofuels from microbes. *Applied Microbiology*

and Biotechnology, 77(1):23–35, 10 2007. ISSN 0175-7598. doi: 10.1007/s00253-007-1163-x. URL http://link.springer.com/10.1007/s00253-007-1163-x.

Archetti, M. and Scheuring, I. COEXISTENCE OF COOPERATION AND DEFECTION IN PUBLIC GOODS GAMES. *Evolution*, 65(4):1140–1148, 4 2011. ISSN 00143820. doi: 10.1111/j.1558-5646.2010.01185.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1558-5646.2010.01185.x.

Bairey, E., Kelsic, E. D., and Kishony, R. High-order species interactions shape ecosystem diversity. *Nature Communications*, 7:1–7, 2016. ISSN 20411723. doi: 10.1038/ncomms12285. URL http://dx.doi.org/10.1038/ncomms12285.

Bakarji, J., Champion, K., Kutz, J. N., and Brunton, S. L. Discovering Governing Equations from Partial Measurements with Deep Delay Autoencoders. *arXiv*, pages 1–23, 1 2022. URL http://arxiv.org/abs/2201.05136.

Barabás, G., D'Andrea, R., and Stump, S. M. Chesson's coexistence theory. *Ecological Monographs*, 0(0):1–27, 2018. ISSN 00129615. doi: 10.1002/ecm.1302. URL http://doi.wiley.com/10.1002/ecm.1302.

Bayrock, D. P. and Michael Ingledew, W. Application of multistage continuous fermentation for production of fuel alcohol by very-high-gravity fermentation technology. *Journal of Industrial Microbiology and Biotechnology*, 27(2):87–93, 8 2001. ISSN 1367-5435. doi: 10.1038/sj.jim.7000167. URL https://academic.oup.com/jimb/article/27/2/87-93/5990373.

Ben Said, S. and Or, D. Synthetic Microbial Ecology: Engineering Habitats for Modular Consortia. *Frontiers in Microbiology*, 8(JUN), 6 2017. ISSN 1664-302X. doi: 10.3389/fmicb.2017.01125. URL http://journal.frontiersin.org/article/10.3389/fmicb.2017.01125/full.

Ben Said, S., Tecon, R., Borer, B., and Or, D. The engineering of spatially linked microbial consortia – potential and perspectives. *Current Opinion in Biotechnology*, 62:137–145, 4 2020. ISSN 09581669. doi: 10.1016/j.copbio.2019.09.015. URL https://doi.org/10.1016/j.copbio.2019.09.015https://linkinghub.elsevier.com/retrieve/pii/S0958166919300904.

Bhunia, A. K., Sahoo, L., and Shaikh, A. A. *Advanced Optimization and Operations Research*, volume 153 of *Springer Optimization and Its Applications*. Springer Singapore, Singapore, 2019. ISBN 978-981-32-9966-5. doi: 10.1007/978-981-32-9967-2. URL http://link.springer.com/10.1007/978-981-32-9967-2_1.

Birch, L. The Meanings of Competition. *The American Naturalist*, 91(856):5–18, 1957. URL https://www.jstor.org/stable/2458507.

Blair, J. M. A., Webber, M. A., Baylay, A. J., Ogbolu, D. O., and Piddock, L. J. V. Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology*, 13(1):42–51, 1 2015. ISSN 1740-1526. doi: 10.1038/nrmicro3380. URL http://www.nature.com/articles/nrmicro3380.

Bokulich, N. A., Ohta, M., Lee, M., and Mills, D. A. Indigenous Bacteria and Fungi Drive Traditional Kimoto Sake Fermentations. *Applied and Environmental Microbiology*, 80(17): 5522–5529, 9 2014. ISSN 0099-2240. doi: 10.1128/AEM.00663-14. URL `https://journals.asm.org/doi/10.1128/AEM.00663-14`.

Bonanomi, G., Capodilupo, M., Incerti, G., Gaglione, S. A., and Scala, F. Fungal diversity increases soil fungistasis and resistance to microbial invasion by a non resident species. *Biological Control*, 72:38–45, 2014. ISSN 10499644. doi: 10.1016/j.biocontrol.2014.02.005. URL `http://dx.doi.org/10.1016/j.biocontrol.2014.02.005`.

Bongard, J. and Lipson, H. Automated reverse engineering of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 104(24):9943–9948, 6 2007. ISSN 0027-8424. doi: 10.1073/pnas.0609476104. URL `https://pnas.org/doi/full/10.1073/pnas.0609476104`.

Bossaert, S., Winne, V., Van Opstaele, F., Buyse, J., Verreth, C., Herrera-Malaver, B., Van Geel, M., Verstrepen, K. J., Crauwels, S., De Rouck, G., and Lievens, B. Description of the temporal dynamics in microbial community composition and beer chemistry in sour beer production via barrel ageing of finished beers. *International Journal of Food Microbiology*, 339 (December 2020):109030, 2 2021. ISSN 01681605. doi: 10.1016/j.ijfoodmicro.2020.109030. URL `https://linkinghub.elsevier.com/retrieve/pii/S0168160520305249`.

Bottery, M. J., Wood, A. J., and Brockhurst, M. A. Selective Conditions for a Multidrug Resistance Plasmid Depend on the Sociality of Antibiotic Resistance. *Antimicrobial Agents and Chemotherapy*, 60(4):2524–2527, 4 2016. ISSN 0066-4804. doi: 10.1128/AAC.02441-15. URL `http://aac.asm.org/lookup/doi/10.1128/AAC.02441-15`.

Brännström, , Johansson, J., and von Festenberg, N. The Hitchhiker's Guide to Adaptive Dynamics. *Games*, 4(3):304–328, 2013. ISSN 2073-4336. doi: 10.3390/g4030304. URL `http://www.mdpi.com/2073-4336/4/3/304/`.

Brodie, J., Chan, C. X., De Clerck, O., Cock, J. M., Coelho, S. M., Gachon, C., Grossman, A. R., Mock, T., Raven, J. A., Smith, A. G., Yoon, H. S., and Bhattacharya, D. The Algal Revolution. *Trends in Plant Science*, xx:1–13, 6 2017. ISSN 13601385. doi: 10.1016/j.tplants.2017.05.005. URL `http://linkinghub.elsevier.com/retrieve/pii/S136013851730105X`.

Broom, M. and Rychtar, J. *Game-Theoretical Models in Biology*. Chapman and Hall/CRC, 1 edition, 2013. ISBN 9781439853221. URL `https://www.crcpress.com/Game-Theoretical-Models-in-Biology/Broom-Rychtar/9781439853214`.

Broom, M., Pattni, K., and Rychtář, J. Generalized Social Dilemmas: The Evolution of Cooperation in Populations with Variable Group Size. *Bulletin of Mathematical Biology*, pages 1–32, 12 2018. ISSN 0092-8240. doi: 10.1007/s11538-018-00545-1. URL `http://link.springer.com/10.1007/s11538-018-00545-1`.

Brown, B. L. and Swan, C. M. Dendritic network structure constrains metacommunity properties in riverine ecosystems. *Journal of Animal Ecology*, 79(3):571–580, 5 2010. ISSN 00218790. doi: 10.1111/j.1365-2656.2010.01668.x. URL `http://doi.wiley.com/10.1111/j.1365-2656.2010.01668.x`.

Brown, J. S. Why Darwin would have loved evolutionary game theory. *Proceedings of the Royal Society B: Biological Sciences*, 283(1838):20160847, 9 2016. ISSN 0962-8452. doi: 10.1098/rspb.2016.0847. URL `https://royalsocietypublishing.org/doi/10.1098/rspb.2016.0847`.

Brunton, S. L., Proctor, J. L., and Kutz, J. N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, 4 2016. ISSN 0027-8424. doi: 10.1073/pnas.1517384113. URL `https://pnas.org/doi/full/10.1073/pnas.1517384113`.

Buhlman, P. Bootstraps for time series. *Statistical Science*, 17(1):52–72, 2002. URL `http://www.jstor.org/stable/10.2307/3182810`.

Butler, S. and O'Dwyer, J. P. Stability criteria for complex microbial communities. *Nature Communications*, 9(1):2970, 12 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05308-z. URL `http://www.nature.com/articles/s41467-018-05308-z`.

Carraro, L., Bertuzzo, E., Fronhofer, E. A., Furrer, R., Gounand, I., Rinaldo, A., and Altermatt, F. Generation and application of river network analogues for use in ecology and evolution. *Ecology and Evolution*, 10(14):7537–7550, 7 2020. ISSN 2045-7758. doi: 10.1002/ece3.6479. URL `https://onlinelibrary.wiley.com/doi/10.1002/ece3.6479`.

Case, T. J. Invasion resistance arises in strongly interacting species-rich model competition communities. *Proceedings of the National Academy of Sciences*, 87(24):9610–9614, 12 1990. ISSN 0027-8424. doi: 10.1073/pnas.87.24.9610. URL `https://pnas.org/doi/full/10.1073/pnas.87.24.9610`.

Case, T. J. and Gilpin, M. E. Interference Competition and Niche Theory. *Proceedings of the National Academy of Sciences*, 71(8):3073–3077, 8 1974. ISSN 0027-8424. doi: 10.1073/pnas.71.8.3073. URL `https://pnas.org/doi/full/10.1073/pnas.71.8.3073`.

Cason, E. D., Mahlomaholo, B. J., Taole, M. M., Abong, G. O., Vermeulen, J. G., de Smidt, O., Vermeulen, M., Steyn, L., Valverde, A., and Viljoen, B. Bacterial and Fungal Dynamics During the Fermentation Process of Sesotho, a Traditional Beer of Southern Africa. *Frontiers in Microbiology*, 11(June):1–14, 2020. ISSN 1664302X. doi: 10.3389/fmicb.2020.01451.

Cenci, S. and Saavedra, S. Structural stability of nonlinear population dynamics. *Physical Review E*, 97(1):012401, 1 2018. ISSN 2470-0045. doi: 10.1103/PhysRevE.97.012401. URL `https://link.aps.org/doi/10.1103/PhysRevE.97.012401`.

Chamberlain, S. A., Bronstein, J. L., and Rudgers, J. A. How context dependent are species interactions? *Ecology Letters*, 17(7):881–890, 7 2014. ISSN 1461023X. doi: 10.1111/ele.12279. URL `https://onlinelibrary.wiley.com/doi/10.1111/ele.12279`.

Chen, Y., Lin, C. J., Jones, G., Fu, S., and Zhan, H. Enhancing biodegradation of wastewater by microbial consortia with fractional factorial design. *Journal of Hazardous Materials*, 171(1-3): 948–953, 2009. ISSN 03043894. doi: 10.1016/j.jhazmat.2009.06.100.

Chesson, P. MacArthur's consumer-resource model. *Theoretical Population Biology*, 37(1):26–38, 2 1990. ISSN 00405809. doi: 10.1016/0040-5809(90)90025-Q. URL `https://linkinghub.elsevier.com/retrieve/pii/004058099090025Q`.

Chesson, P. Multispecies Competition in Variable Environments. *Theoretical Population Biology*, 45(3):227–276, 6 1994. ISSN 00405809. doi: 10.1006/tpbi.1994.1013. URL `https://linkinghub.elsevier.com/retrieve/pii/S0040580984710136`.

Chesson, P. Mechanisms of Maintenance of Species Diversity. *Annual Review of Ecology and Systematics*, 31(1):343–366, 11 2000a. ISSN 0066-4162. doi: 10.1146/annurev.ecolsys.31.1.343. URL `http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.31.1.343`.

Chesson, P. General Theory of Competitive Coexistence in Spatially-Varying Environments. *Theoretical Population Biology*, 58(3):211–237, 11 2000b. ISSN 00405809. doi: 10.1006/tpbi. 2000.1486. URL `https://linkinghub.elsevier.com/retrieve/pii/S0040580900914862`.

Cheung, S. G., Goldenthal, A. R., Uhlemann, A.-C., Mann, J. J., Miller, J. M., and Sublette, M. E. Systematic Review of Gut Microbiota and Major Depression. *Frontiers in Psychiatry*, 10(FEB), 2 2019. ISSN 1664-0640. doi: 10.3389/fpsyt.2019.00034. URL `https://www.frontiersin.org/article/10.3389/fpsyt.2019.00034/full`.

Chisholm, C., Lindo, Z., and Gonzalez, A. Metacommunity diversity depends on connectivity and patch arrangement in heterogeneous habitat networks. *Ecography*, 34(3):415–424, 6 2011. ISSN 09067590. doi: 10.1111/j.1600-0587.2010.06588.x. URL `https://onlinelibrary.wiley.com/doi/10.1111/j.1600-0587.2010.06588.x`.

Cinquin, C., Le Blay, G., Fliss, I., and Lacroix, C. New three-stage in vitro model for infant colonic fermentation with immobilized fecal microbiota. *FEMS Microbiology Ecology*, 57 (2):324–336, 8 2006. ISSN 01686496. doi: 10.1111/j.1574-6941.2006.00117.x. URL `https://academic.oup.com/femsec/article-lookup/doi/10.1111/j.1574-6941.2006.00117.x`.

Clark, R. L., Connors, B. M., Stevenson, D. M., Hromada, S. E., Hamilton, J. J., Amador-Noguez, D., and Venturelli, O. S. Design of synthetic human gut microbiome assembly and butyrate production. *Nature Communications*, 12(1):3254, 12 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-22938-y. URL `http://www.nature.com/articles/s41467-021-22938-y`.

Cocolin, L. Direct profiling of the yeast dynamics in wine fermentations. *FEMS Microbiology Letters*, 189(1):81–87, 8 2000. ISSN 03781097. doi: 10.1016/S0378-1097(00)00257-3. URL `http://doi.wiley.com/10.1016/S0378-1097(00)00257-3`.

Cornforth, D. M. and Foster, K. R. Competition sensing: the social side of bacterial stress responses. *Nature Reviews Microbiology*, 11(4):285–293, 4 2013. ISSN 1740-1526. doi: 10.1038/nrmicro2977. URL `http://www.nature.com/articles/nrmicro2977`.

Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J. M., and Relman, D. A. The Application of Ecological Theory Toward an Understanding of the Human Microbiome. *Science*, 336(6086):1255–1262, 6 2012. ISSN 0036-8075. doi: 10.1126/science.1224203. URL `https://www.science.org/doi/10.1126/science.1224203`.

Coyte, K. Z., Schluter, J., and Foster, K. R. The ecology of the microbiome: Networks, competition, and stability. *Science*, 350(6261):663–666, 2015. ISSN 0036-8075. doi: 10.1126/science.aad2602. URL `http://www.sciencemag.org/cgi/doi/10.1126/science.aad2602`.

Daniels, B. C. and Nemenman, I. Automated adaptive inference of phenomenological dynamical models. *Nature Communications*, 6:1–8, 2015. ISSN 20411723. doi: 10.1038/ncomms9133.

Data61, C. Stellargraph machine learning library. `https://github.com/stellargraph/stellargraph`, 2018.

Dell'Anno, A., Beolchini, F., Rocchetti, L., Luna, G. M., and Danovaro, R. High bacterial biodiversity increases degradation performance of hydrocarbons during bioremediation of contaminated harbor marine sediments. *Environmental Pollution*, 167:85–92, 2012. ISSN 02697491. doi: 10.1016/j.envpol.2012.03.043. URL `http://dx.doi.org/10.1016/j.envpol.2012.03.043`.

Demeter, M. *Harnessing Mixed Species Biofilms for Hydrocarbon Biodegradation*. PhD thesis, University of Calgary, 2015. URL `http://hdl.handle.net/11023/2424`.

Diekmann, O. A BEGINNER'S GUIDE TO ADAPTIVE DYNAMICS. *Banach Center Publications*, 63(1):47–86, 2 2004. ISSN 1740-1526. URL `https://eudml.org/doc/281592`.

Donohue, I., Petchey, O. L., Montoya, J. M., Jackson, A. L., McNally, L., Viana, M., Healy, K., Lurgi, M., O'Connor, N. E., and Emmerson, M. C. On the dimensionality of ecological stability. *Ecology Letters*, 16(4):421–429, 4 2013. ISSN 1461023X. doi: 10.1111/ele.12086. URL `https://onlinelibrary.wiley.com/doi/10.1111/ele.12086`.

Dos Santos, A. Deciphering between-species interactions within a bacterial consortium designed to degrade metal working fluids. Master's thesis, Faculty of Biology and Medicine, University of Lausanne, 2019.

Dos Santos, A. R., Di Martino, R., Testa, S., and Mitri, S. Classifying interactions in a synthetic bacterial community is hindered by inhibitory growth. *bioRxiv*, pages 1–11, 2022.

doi: doi.org/10.1101/2022.03.02.482509. URL `https://www.biorxiv.org/content/10.1101/2022.03.02.482509v1`.

Dua, M., Singh, A., Sethunathan, N., and Johri, A. Biotechnology and bioremediation: successes and limitations. *Applied Microbiology and Biotechnology*, 59(2-3):143–152, 7 2002. ISSN 0175-7598. doi: 10.1007/s00253-002-1024-6. URL `http://link.springer.com/10.1007/s00253-002-1024-6`.

Dunne, J. A., Williams, R. J., and Martinez, N. D. Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, 5(4):558–567, 7 2002. ISSN 1461-023X. doi: 10.1046/j.1461-0248.2002.00354.x. URL `http://doi.wiley.com/10.1046/j.1461-0248.2002.00354.x`.

Economo, E. P. and Keitt, T. H. Species diversity in neutral metacommunities: a network approach. *Ecology Letters*, 11(1):071117033013001–???, 11 2007. ISSN 1461-023X. doi: 10.1111/j.1461-0248.2007.01126.x. URL `https://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2007.01126.x`.

Eiteman, M. A., Lee, S. A., and Altman, E. A co-fermentation strategy to consume sugar mixtures effectively. *Journal of Biological Engineering*, 2(1):3, 2008. ISSN 1754-1611. doi: 10.1186/1754-1611-2-3. URL `http://jbioleng.biomedcentral.com/articles/10.1186/1754-1611-2-3`.

Ellis, R. J., Lilley, A. K., Lacey, S. J., Murrell, D., and Godfray, H. C. J. Frequency-dependent advantages of plasmid carriage by Pseudomonas in homogeneous and spatially structured environments. *The ISME Journal*, 1(1):92–95, 5 2007. ISSN 1751-7362. doi: 10.1038/ismej.2007.11. URL `http://www.nature.com/articles/ismej200711`.

Estrela, S., Libby, E., Van Cleve, J., Débarre, F., Deforet, M., Harcombe, W. R., Peña, J., Brown, S. P., and Hochberg, M. E. Environmentally Mediated Social Dilemmas. *Trends in Ecology & Evolution*, 34(1):6–18, 1 2019. ISSN 01695347. doi: 10.1016/j.tree.2018.10.004. URL `https://linkinghub.elsevier.com/retrieve/pii/S0169534718302490`.

Evrensel, A. and Ceylan, M. E. The gut-brain axis: The missing link in depression. *Clinical Psychopharmacology and Neuroscience*, 13(3):239–244, 2015. ISSN 20934327. doi: 10.9758/cpn.2015.13.3.239.

Gardner, M. R. and Ashby, W. R. Connectance of Large Dynamic (Cybernetic) Systems: Critical Values for Stability. *Nature*, 228(5273):784–784, 11 1970. ISSN 0028-0836. doi: 10.1038/228784a0. URL `https://www.nature.com/articles/228784a0`.

Gause, G. F. Experimental analysis of Vito Volterra's mathematical theory of the struggle for existence. *Science*, 79(2036):16–17, 1 1934. doi: 10.1126/science.79.2036.16-a. URL `http://www.sciencemag.org/cgi/doi/10.1126/science.79.2036.16-a`.

George, A. B. and Korolev, K. S. Ecological landscapes guide the assembly of optimal microbial communities. *arXiv*, pages 1–33, 10 2021. URL `http://arxiv.org/abs/2110.11310`.

Georgianna, D. R. and Mayfield, S. P. Exploiting diversity and synthetic biology for the production of algal biofuels. *Nature*, 488(7411):329–335, 8 2012. ISSN 0028-0836. doi: 10.1038/nature11479. URL `https://www.nature.com/articles/nature11479`.

Geritz, S., Kisdi, , Meszéna, G., and Metz, J. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, 12(1):35–57, 1 1998. ISSN 0269-7653. doi: 10.1023/A:1006554906681. URL `http://link.springer.com/10.1023/A:1006554906681`.

Ghoul, M. and Mitri, S. Microbial Communities The Ecology and Evolution of Microbial Competition. *Trends in Microbiology*, xx:1–13, 2016. ISSN 0966-842X. doi: 10.1016/j.tim.2016.06.011. URL `http://dx.doi.org/10.1016/j.tim.2016.06.011`.

Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 12 1977. ISSN 0022-3654. doi: 10.1021/j100540a008. URL `https://pubs.acs.org/doi/10.1021/j100540a008`.

Godsoe, W., Holland, N. J., Cosner, C., Kendall, B. E., Brett, A., Jankowski, J., and Holt, R. D. Interspecific interactions and range limits: contrasts among interaction types. *Theoretical Ecology*, 10(2):167–179, 6 2017. ISSN 1874-1738. doi: 10.1007/s12080-016-0319-7. URL `http://link.springer.com/10.1007/s12080-016-0319-7`.

Godwin, C. M., Chang, F., and Cardinale, B. J. An empiricist's guide to modern coexistence theory for competitive communities. *Oikos*, 129(8):1109–1127, 8 2020. ISSN 0030-1299. doi: 10.1111/oik.06957. URL `https://onlinelibrary.wiley.com/doi/10.1111/oik.06957`.

Gong, L., Gao, J., and Cao, M. Evolutionary Game Dynamics for Two Interacting Populations in A Co-evolving Environment. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 3535–3540. IEEE, 12 2018. ISBN 978-1-5386-1395-5. doi: 10.1109/CDC.2018.8619801. URL `http://arxiv.org/abs/1806.03194https://ieeexplore.ieee.org/document/8619801/`.

Grainger, T. N., Letten, A. D., Gilbert, B., and Fukami, T. Applying modern coexistence theory to priority effects. *Proceedings of the National Academy of Sciences*, 116(13):201803122, 2019. ISSN 0027-8424. doi: 10.1073/pnas.1803122116. URL `http://www.pnas.org/lookup/doi/10.1073/pnas.1803122116`.

Hamilton, W. L. Graph Representation Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 14(3):1–159, 9 2020. ISSN 1939-4608. doi: 10.2200/S01045ED1V01Y202009AIM046. URL `https://www.morganclaypool.com/doi/10.2200/S01045ED1V01Y202009AIM046`.

Han, B.-Z., Rombouts, F. M., and Nout, M. A Chinese fermented soybean food. *International Journal of Food Microbiology*, 65(1-2):1–10, 4 2001. ISSN 01681605. doi: 10.1016/S0168-1605(00)00523-7. URL https://linkinghub.elsevier.com/retrieve/pii/S0168160500005237.

Hanski, I. and Gyllenberg, M. Two General Metapopulation Models and the Core-Satellite Species Hypothesis. *The American Naturalist*, 142(1):17–41, 1993. URL https://www.jstor.org/stable/2462632.

Hardin, G. The competitive exclusion principle. *Science*, 131(3409):1292–1297, 1960. ISSN 00368075. doi: 10.1126/science.131.3409.1292.

Hooten, M. B. and Hobbs, N. T. A guide to Bayesian model selection for ecologists. *Ecological Monographs*, 85(1):3–28, 2 2015. ISSN 0012-9615. doi: 10.1890/14-0661.1. URL http://doi.wiley.com/10.1890/14-0661.1.

Hromada, S., Clark, R. L., Qian, Y., Watson, L., Safdar, N., and Venturelli, O. S. Species richness determines C. difficile invasion outcome in synthetic human gut communities. *bioRxiv*, page 2021.03.23.436677, 2021. doi: 10.1101/2021.03.23.436677. URL https://www.biorxiv.org/content/10.1101/2021.03.23.436677v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.03.23.436677v1.abstract.

Ives, A. R. and Carpenter, S. R. Stability and Diversity of Ecosystems. *Science*, 317(5834):58–62, 2007. ISSN 0036-8075. doi: 10.1126/science.1133258. URL http://www.sciencemag.org/cgi/doi/10.1126/science.1133258.

Jacquet, C., Moritz, C., Morissette, L., Legagneux, P., Massol, F., Archambault, P., and Gravel, D. No complexity–stability relationship in empirical ecosystems. *Nature Communications*, 7(1):12573, 11 2016. ISSN 2041-1723. doi: 10.1038/ncomms12573. URL http://www.nature.com/articles/ncomms12573.

Kalantary, R. R., Mohseni-Bandpi, A., Esrafili, A., Nasseri, S., Ashmagh, F. R., Jorfi, S., and Ja'fari, M. Effectiveness of biostimulation through nutrient content on the bioremediation of phenanthrene contaminated soil. *Journal of Environmental Health Science and Engineering*, 12(1):143, 12 2014. ISSN 2052-336X. doi: 10.1186/s40201-014-0143-1. URL https://link.springer.com/10.1186/s40201-014-0143-1.

Kang, C.-h., Kwon, Y.-j., and So, J.-s. Bioremediation of heavy metals by using bacterial mixtures. *Ecological Engineering*, 89:64–69, 4 2016. ISSN 09258574. doi: 10.1016/j.ecoleng.2016.01.023. URL http://dx.doi.org/10.1016/j.ecoleng.2016.01.023https://linkinghub.elsevier.com/retrieve/pii/S0925857416300234.

Kang, Y. H., Kim, S., Choi, S. K., Lee, H. J., Chung, I. K., and Park, S. R. A comparison of the bioremediation potential of five seaweed species in an integrated fish-seaweed aquaculture system: implication for a multi-species seaweed culture. *Reviews in Aquaculture*, 13(1):353–364,

1 2021. ISSN 1753-5123. doi: 10.1111/raq.12478. URL `https://onlinelibrary.wiley.com/doi/10.1111/raq.12478`.

Karkman, A., Lehtimäki, J., and Ruokolainen, L. The ecology of human microbiota: dynamics and diversity in health and disease. *Annals of the New York Academy of Sciences*, 1399(1): 78–92, 7 2017. ISSN 00778923. doi: 10.1111/nyas.13326. URL `https://onlinelibrary.wiley.com/doi/10.1111/nyas.13326`.

Kazamia, E., Riseley, A. S., Howe, C. J., and Smith, A. G. An Engineered Community Approach for Industrial Cultivation of Microalgae. *Industrial Biotechnology*, 10(3):184–190, 6 2014. ISSN 1550-9087. doi: 10.1089/ind.2013.0041. URL `http://www.liebertpub.com/doi/10.1089/ind.2013.0041`.

Kim, H. J., Boedicker, J. Q., Choi, J. W., and Ismagilov, R. F. Defined spatial structure stabilizes a synthetic multispecies bacterial community. *Proceedings of the National Academy of Sciences*, 105(47):18188–18193, 11 2008. ISSN 0027-8424. doi: 10.1073/pnas.0807935105. URL `http://www.pnas.org/cgi/doi/10.1073/pnas.0807935105`.

Kitagaki, H. and Kitamoto, K. Breeding Research on Sake Yeasts in Japan: History, Recent Technological Advances, and Future Perspectives. *Annual Review of Food Science and Technology*, 4(1):215–235, 2 2013. ISSN 1941-1413. doi: 10.1146/annurev-food-030212-182545. URL `https://www.annualreviews.org/doi/10.1146/annurev-food-030212-182545`.

Kong, W., Meldgin, D. R., Collins, J. J., and Lu, T. Designing microbial consortia with defined social interactions. *Nature Chemical Biology*, 14(8):821–829, 8 2018. ISSN 1552-4450. doi: 10.1038/s41589-018-0091-7. URL `http://www.nature.com/articles/s41589-018-0091-7`.

Krebs, C. J. *Ecology: The Experimental Analysis of Distribution and Abundance.* Pearson, 6th edition, 7 2013. ISBN 9781292026275.

Lakshmanan, D., Clifford, D., and Samanta, G. Arsenic Removal by Coagulation With Aluminum, Iron, Titanium, and Zirconium. *Journal - American Water Works Association*, 100(2):76–88, 2 2008. ISSN 0003150X. doi: 10.1002/j.1551-8833.2008.tb08144.x. URL `https://onlinelibrary.wiley.com/doi/10.1002/j.1551-8833.2008.tb08144.x`.

Landi, P., Minoarivelo, H. O., Brännström, , Hui, C., and Dieckmann, U. Complexity and stability of ecological networks: a review of the theory. *Population Ecology*, 60(4):319–345, 10 2018. ISSN 1438-3896. doi: 10.1007/s10144-018-0628-3. URL `https://onlinelibrary.wiley.com/doi/10.1007/s10144-018-0628-3`.

Lea-Smith, D. J. and Howe, C. J. The Use of Cyanobacteria for Biofuel Production. In Love, J. and Bryant, J. A., editors, *Biofuels and Bioenergy*, chapter 9, pages 143–155. John Wiley & Sons, Ltd, Chichester, UK, 1 edition, 3 2017. ISBN 9781118350560. doi: 10.1002/9781118350553.ch9. URL `http://doi.wiley.com/10.1002/9781118350553.ch9`.

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Holt, R. D., Shurin, J. B., Law, R., Tilman, D., Loreau, M., and Gonzalez, A. The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, 7(7):601–613, 6 2004. ISSN 1461023X. doi: 10.1111/j.1461-0248.2004.00608.x. URL `https://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2004.00608.x`.

Lenhart, S. and Workman, J. T. *Optimal Control Applied to Biological Models.* Chapman & Hall/CRC, 2007. ISBN 9781420011418. doi: 10.1201/9781420011418.

León, J. A. and Tumpson, D. B. Competition between two species for two complementary or substitutable resources. *Journal of Theoretical Biology*, 50(1):185–201, 3 1975. ISSN 00225193. doi: 10.1016/0022-5193(75)90032-6. URL `https://linkinghub.elsevier.com/retrieve/pii/0022519375900326`.

Letten, A. D., Ke, P.-J., and Fukami, T. Linking modern coexistence theory and contemporary niche theory. *Ecological Monographs*, 87(2):161–177, 5 2017. ISSN 00129615. doi: 10.1002/ecm.1242. URL `http://doi.wiley.com/10.1002/ecm.1242`.

Levin, S. A. Community Equilibria and Stability, and an Extension of the Competitive Exclusion Principle. *The American Naturalist*, 104(939):413–423, 9 1970. ISSN 0003-0147. doi: 10.1086/282676. URL `https://www.journals.uchicago.edu/doi/10.1086/282676`.

Ley, R. E., Turnbaugh, P. J., Klein, S., and Gordon, J. I. Human gut microbes associated with obesity. *Nature*, 444(7122):1022–1023, 12 2006. ISSN 0028-0836. doi: 10.1038/4441022a. URL `http://www.nature.com/articles/4441022a`.

Li, S.-p., Tan, J., Yang, X., Ma, C., and Jiang, L. Niche and fitness differences determine invasion success and impact in laboratory bacterial communities. *The ISME Journal*, 13(2):402–412, 2 2019. ISSN 1751-7362. doi: 10.1038/s41396-018-0283-x. URL `http://dx.doi.org/10.1038/s41396-018-0283-xhttp://www.nature.com/articles/s41396-018-0283-x`.

Lin, Y. H., Bayrock, D. P., and Ingledew, W. M. Evaluation of Saccharomyces cerevisiae grown in a multistage chemostat environment under increasing levels of glucose. *Biotechnology Letters*, 24(6):449–453, 2002. doi: https://doi.org/10.1023/A:1014501125355. URL `https://rdcu.be/cMKvQ`.

Liu, L., Wang, J., Rosenberg, D., Zhao, H., Lengyel, G., and Nadel, D. Fermented beverage and food storage in 13,000 y-old stone mortars at Raqefet Cave, Israel: Investigating Natufian ritual feasting. *Journal of Archaeological Science: Reports*, 21(May):783–793, 2018. ISSN 2352409X. doi: 10.1016/j.jasrep.2018.08.008. URL `https://doi.org/10.1016/j.jasrep.2018.08.008`.

Lotka, A. J. Contribution to the Theory of Periodic Reactions. *The Journal of Physical Chemistry*, 14(3):271–274, 3 1910. ISSN 0092-7325. doi: 10.1021/j150111a004. URL `https://pubs.acs.org/doi/abs/10.1021/j150111a004`.

MacArthur, R. Fluctuations of Animal Populations and a Measure of Community Stability Author. *Ecology*, 36(3):533–536, 1955. URL `https://www.jstor.org/stable/1929601`.

MacArthur, R. Species packing and competitive equilibrium for many species. *Theoretical Population Biology*, 1(1):1–11, 5 1970. ISSN 00405809. doi: 10.1016/0040-5809(70)90039-0. URL `https://linkinghub.elsevier.com/retrieve/pii/0040580970900390`.

Marchesi, J. R., Adams, D. H., Fava, F., Hermes, G. D., Hirschfield, G. M., Hold, G., Quraishi, M. N., Kinross, J., Smidt, H., Tuohy, K. M., Thomas, L. V., Zoetendal, E. G., and Hart, A. The gut microbiota and host health: A new clinical frontier. *Gut*, 65(2):330–339, 2016. ISSN 14683288. doi: 10.1136/gutjnl-2015-309990.

May, R. M. Will a Large Complex System be Stable? *Nature*, 238(5364):413–414, 8 1972. ISSN 0028-0836. doi: 10.1038/238413a0. URL `https://www.nature.com/articles/238413a0`.

Mayfield, M. M. and Stouffer, D. B. Higher-order interactions capture unexplained complexity in diverse communities. *Nature Ecology and Evolution*, 1(3):0062, 2 2017. ISSN 2397334X. doi: 10.1038/s41559-016-0062. URL `http://www.ncbi.nlm.nih.gov/pubmed/28812740`.

Maynard Smith, J. and Price, G. R. The Logic of Animal Conflict. *Nature*, 246(5427):15–18, 11 1973. ISSN 0028-0836. doi: 10.1038/246015a0. URL `https://www.nature.com/articles/246015a0`.

McCann, K. S. The diversity–stability debate. *Nature*, 405(6783):228–233, 5 2000. ISSN 0028-0836. doi: 10.1038/35012234. URL `http://www.nature.com/articles/35012234`.

McClure, S. B., Magill, C., Podrug, E., Moore, A. M. T., Harper, T. K., Culleton, B. J., Kennett, D. J., and Freeman, K. H. Fatty acid specific δ13C values reveal earliest Mediterranean cheese production 7,200 years ago. *PLOS ONE*, 13(9):e0202807, 9 2018. ISSN 1932-6203. doi: 10.1371/journal.pone.0202807. URL `https://dx.plos.org/10.1371/journal.pone.0202807`.

McGovern, P., Jalabadze, M., Batiuk, S., Callahan, M. P., Smith, K. E., Hall, G. R., Kvavadze, E., Maghradze, D., Rusishvili, N., Bouby, L., Failla, O., Cola, G., Mariani, L., Boaretto, E., Bacilieri, R., This, P., Wales, N., and Lordkipanidze, D. Early Neolithic wine of Georgia in the South Caucasus. *Proceedings of the National Academy of Sciences*, 114(48):E10309–E10318, 11 2017. ISSN 0027-8424. doi: 10.1073/pnas.1714728114. URL `https://pnas.org/doi/full/10.1073/pnas.1714728114`.

Meszéna, G., Gyllenberg, M., Pásztor, L., and Metz, J. A. Competitive exclusion and limiting similarity: A unified theory. *Theoretical Population Biology*, 69(1):68–87, 2 2006. ISSN 00405809. doi: 10.1016/j.tpb.2005.07.001. URL `https://linkinghub.elsevier.com/retrieve/pii/S004058090500095X`.

Michel, R. H., McGovern, P. E., and Badler, V. R. Chemical evidence for ancient beer. *Nature*, 360(6399):24–24, 11 1992. ISSN 0028-0836. doi: 10.1038/360024b0. URL `http://www.nature.com/articles/360024b0`.

Miller, Burns, Munguia, Walters, Kneitel, Richards, Mouquet, and Buckley. A Critical Review of Twenty Years' Use of the Resource-Ratio Theory. *The American Naturalist*, 165(4):439, 2005. ISSN 00030147. doi: 10.2307/3473474. URL `https://www.jstor.org/stable/10.2307/3473474`.

Momeni, B., Xie, L., and Shou, W. Lotka-Volterra pairwise modeling fails to capture diverse pairwise microbial interactions. *eLife*, 6:1–34, 2017. ISSN 2050084X. doi: https://doi.org/10.7554/eLife.25051.001. URL `https://elifesciences.org/articles/25051#content`.

Mougi, A. and Kondoh, M. Diversity of Interaction Types and Ecological Community Stability. *Science*, 337(6092):349–351, 2012. ISSN 0036-8075. doi: 10.1126/science.1220529. URL `http://www.sciencemag.org/cgi/doi/10.1126/science.1220529`.

Murray, J. D. *Mathematical Biology : I . An Introduction.* Springer, New York, 3 edition, 2002. ISBN 0387952233.

Narwani, A., Alexandrou, M. A., Oakley, T. H., Carroll, I. T., and Cardinale, B. J. Experimental evidence that evolutionary relatedness does not affect the ecological mechanisms of coexistence in freshwater green algae. *Ecology Letters*, 16(11):1373–1381, 11 2013. ISSN 1461023X. doi: 10.1111/ele.12182. URL `https://onlinelibrary.wiley.com/doi/10.1111/ele.12182`.

Newville, M., Stensitzki, T., Allen, D. B., and Ingargiola, A. Lmfit: Non-linear least-square minimization and curve-fitting for python, 2014. URL `https://zenodo.org/record/11813#.Yfqy3i3vBz8`.

Nowak, M. A. *Evolutionary dynamics exploring the equations of life.* Belknap Press of Harvard University Press, Cambridge, 2006. ISBN 9780674023383.

Nozzi, N. E., Oliver, J. W. K., and Atsumi, S. Cyanobacteria as a Platform for Biofuel Production. *Frontiers in Bioengineering and Biotechnology*, 1(September):1–6, 2013. ISSN 2296-4185. doi: 10.3389/fbioe.2013.00007. URL `http://journal.frontiersin.org/article/10.3389/fbioe.2013.00007/abstract`.

O'Brien, S. and Buckling, A. The sociality of bioremediation. *EMBO reports*, 16(10):1241–1245, 10 2015. ISSN 1469-221X. doi: 10.15252/embr.201541064. URL `https://onlinelibrary.wiley.com/doi/10.15252/embr.201541064`.

O'Brien, S., Hodgson, D. J., and Buckling, A. Social evolution of toxic metal bioremediation in Pseudomonas aeruginosa. *Proceedings of the Royal Society B: Biological Sciences*, 281 (1787):20140858, 7 2014. ISSN 0962-8452. doi: 10.1098/rspb.2014.0858. URL `https://royalsocietypublishing.org/doi/10.1098/rspb.2014.0858`.

Oliveira Sudário, M. A bacterial assembly line to degrade polluting compounds. Master's thesis, Faculty of Biology and Medicine, University of Lausanne, 2022.

O'Sullivan, J. D., Knell, R. J., and Rossberg, A. G. Metacommunity-scale biodiversity regulation and the self-organised emergence of macroecological patterns. *Ecology Letters*, 22(9):1428–1438, 9 2019. ISSN 1461-023X. doi: 10.1111/ele.13294. URL `https://onlinelibrary.wiley.com/doi/10.1111/ele.13294`.

Payne, A. N., Zihler, A., Chassard, C., and Lacroix, C. Advances and perspectives in in vitro human gut fermentation modeling. *Trends in Biotechnology*, 30(1):17–25, 2012a. ISSN 01677799. doi: 10.1016/j.tibtech.2011.06.011. URL `http://dx.doi.org/10.1016/j.tibtech.2011.06.011`.

Payne, A. N., Zihler, A., Chassard, C., and Lacroix, C. Advances and perspectives in in vitro human gut fermentation modeling. *Trends in Biotechnology*, 30(1):17–25, 2012b. ISSN 01677799. doi: 10.1016/j.tibtech.2011.06.011. URL `http://dx.doi.org/10.1016/j.tibtech.2011.06.011`.

Pham, H. L., Ho, C. L., Wong, A., Lee, Y. S., and Chang, M. W. Applying the design-build-test paradigm in microbiome engineering. *Current Opinion in Biotechnology*, 48:85–93, 2017. ISSN 18790429. doi: 10.1016/j.copbio.2017.03.021. URL `http://dx.doi.org/10.1016/j.copbio.2017.03.021`.

Piccardi, P., Vessman, B., and Mitri, S. Toxicity drives facilitation between 4 bacterial species. *Proceedings of the National Academy of Sciences*, 116(32):15979–15984, 8 2019. ISSN 0027-8424. doi: 10.1073/pnas.1906172116. URL `http://www.pnas.org/lookup/doi/10.1073/pnas.1906172116`.

Pimm, S. L. The complexity and stability of ecosystems. *Nature*, 307(5949):321–326, 1 1984. ISSN 0028-0836. doi: 10.1038/307321a0. URL `http://www.nature.com/articles/307321a0`.

Portillo, M. d. C. and Mas, A. Analysis of microbial diversity and dynamics during wine fermentation of Grenache grape variety by high-throughput barcoding sequencing. *LWT - Food Science and Technology*, 72:317–321, 2016. ISSN 00236438. doi: 10.1016/j.lwt.2016.05.009. URL `http://dx.doi.org/10.1016/j.lwt.2016.05.009`.

Portillo, José, R., Soler-Toscano, F., and Analysis, N. Cooperation enhances structural stability in mutualistic systems. *bioRxiv*, pages 1–26, 2021. URL `http://biorxiv.org/cgi/content/short/2021.04.23.441179v1?rss=1&utm_source=researcher_app&utm_medium=referral&utm_campaign=RESR_MRKT_Researcher_inbound`.

Post, W. and Pimm, S. Community assembly and food web stability. *Mathematical Biosciences*, 64(2):169–192, 6 1983. ISSN 00255564. doi: 10.1016/0025-5564(83)90002-0. URL `https://linkinghub.elsevier.com/retrieve/pii/0025556483900020`.

Raninger, A. and Steiner, W. Accelerated process development for protease production in continuous multi-stage cultures. *Biotechnology and Bioengineering*, 82(5):517–524, 6 2003. ISSN 0006-3592. doi: 10.1002/bit.10597. URL `https://onlinelibrary.wiley.com/doi/10.1002/bit.10597`.

Rohr, R. P., Saavedra, S., and Bascompte, J. On the structural stability of mutualistic systems. *Science*, 345(6195):1253497–1253497, 7 2014. ISSN 0036-8075. doi: 10.1126/science.1253497. URL `http://www.sciencemag.org/cgi/doi/10.1126/science.1253497`.

Rojo-Molinero, E., Macià, M. D., and Oliver, A. Social Behavior of Antibiotic Resistant Mutants Within Pseudomonas aeruginosa Biofilm Communities. *Frontiers in Microbiology*, 10(MAR):1–11, 3 2019. ISSN 1664-302X. doi: 10.3389/fmicb.2019.00570. URL `https://www.frontiersin.org/article/10.3389/fmicb.2019.00570/full`.

Rugbjerg, P., Myling-Petersen, N., Porse, A., Sarup-Lytzen, K., and Sommer, M. O. A. Diverse genetic error modes constrain large-scale bio-based production. *Nature Communications*, 9(1): 787, 12 2018a. ISSN 2041-1723. doi: 10.1038/s41467-018-03232-w. URL `http://www.nature.com/articles/s41467-018-03232-w`.

Rugbjerg, P., Sarup-Lytzen, K., Nagy, M., and Sommer, M. O. A. Synthetic addiction extends the productive life time of engineered Escherichia coli populations. *Proceedings of the National Academy of Sciences*, 115(10):2347–2352, 3 2018b. ISSN 0027-8424. doi: 10.1073/pnas.1718622115. URL `https://pnas.org/doi/full/10.1073/pnas.1718622115`.

Saavedra, S., Rohr, R. P., Bascompte, J., Godoy, O., Kraft, N. J. B., and Levine, J. M. A structural approach for understanding multispecies coexistence. *Ecological Monographs*, 87(3): 470–486, 8 2017. ISSN 00129615. doi: 10.1002/ecm.1263. URL `http://doi.wiley.com/10.1002/ecm.1263`.

Salque, M., Bogucki, P. I., Pyzel, J., Sobkowiak-Tabaka, I., Grygiel, R., Szmyt, M., and Evershed, R. P. Earliest evidence for cheese making in the sixth millennium bc in northern Europe. *Nature*, 493(7433):522–525, 1 2013. ISSN 0028-0836. doi: 10.1038/nature11698. URL `http://www.nature.com/articles/nature11698`.

Salvatier, J., Wiecki, T. V., and Fonnesbeck, C. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, 2:e55, 4 2016. ISSN 2376-5992. doi: 10.7717/peerj-cs.55. URL `http://arxiv.org/abs/1507.08050https://peerj.com/articles/cs-55`.

Samuelson, P. A. The Pure Theory of Public Expenditure. *The Review of Economics and Statistics*, 36(4):387, 11 1954. ISSN 00346535. doi: 10.2307/1925895. URL `https://www.jstor.org/stable/1925895?origin=crossref`.

Sanchez, A. and Gore, J. Feedback between Population and Evolutionary Dynamics Determines the Fate of Social Microbial Populations. *PLoS Biology*, 11(4):e1001547, 4 2013. ISSN 1545-7885. doi: 10.1371/journal.pbio.1001547. URL `https://dx.plos.org/10.1371/journal.pbio.1001547`.

Sánchez, B., Delgado, S., Blanco-Míguez, A., Lourenço, A., Gueimonde, M., and Margolles, A. Probiotics, gut microbiota, and their influence on host health and disease. *Molecular Nutrition & Food Research*, 61(1):1600240, 1 2017. ISSN 16134125. doi: 10.1002/mnfr.201600240. URL `https://onlinelibrary.wiley.com/doi/10.1002/mnfr.201600240`.

Schmidt, M. and Lipson, H. Distilling Free-Form Natural Laws from Experimental Data. *Science*, 324(5923):81–85, 4 2009. ISSN 0036-8075. doi: 10.1126/science.1165893. URL `https://www.science.org/doi/10.1126/science.1165893`.

Selber-Hnatiw, S., Rukundo, B., Ahmadi, M., Akoubi, H., Al-Bizri, H., Aliu, A. F., Ambeaghen, T. U., Avetisyan, L., Bahar, I., Baird, A., Begum, F., Ben Soussan, H., Blondeau-Éthier, V., Bordaries, R., Bramwell, H., Briggs, A., Bui, R., Carnevale, M., Chancharoen, M., Chevassus, T., Choi, J. H., Coulombe, K., Couvrette, F., D'Abreau, S., Davies, M., Desbiens, M.-P., Di Maulo, T., Di Paolo, S.-A., Do Ponte, S., dos Santos Ribeiro, P., Dubuc-Kanary, L.-A., Duncan, P. K., Dupuis, F., El-Nounou, S., Eyangos, C. N., Ferguson, N. K., Flores-Chinchilla, N. R., Fotakis, T., Gado Oumarou H D, M., Georgiev, M., Ghiassy, S., Glibetic, N., Grégoire Bouchard, J., Hassan, T., Huseen, I., Ibuna Quilatan, M.-F., Iozzo, T., Islam, S., Jaunky, D. B., Jeyasegaram, A., Johnston, M.-A., Kahler, M. R., Kaler, K., Kamani, C., Karimian Rad, H., Konidis, E., Konieczny, F., Kurianowicz, S., Lamothe, P., Legros, K., Leroux, S., Li, J., Lozano Rodriguez, M. E., Luponio-Yoffe, S., Maalouf, Y., Mantha, J., McCormick, M., Mondragon, P., Narayana, T., Neretin, E., Nguyen, T. T. T., Niu, I., Nkemazem, R. B., O'Donovan, M., Oueis, M., Paquette, S., Patel, N., Pecsi, E., Peters, J., Pettorelli, A., Poirier, C., Pompa, V. R., Rajen, H., Ralph, R.-O., Rosales-Vasquez, J., Rubinshtein, D., Sakr, S., Sebai, M. S., Serravalle, L., Sidibe, F., Sinnathurai, A., Soho, D., Sundarakrishnan, A., Svistkova, V., Ugbeye, T. E., Vasconcelos, M. S., Vincelli, M., Voitovich, O., Vrabel, P., Wang, L., Wasfi, M., Zha, C. Y., and Gamberi, C. Human Gut Microbiota: Toward an Ecology of Disease. *Frontiers in Microbiology*, 8(JUL), 7 2017. ISSN 1664-302X. doi: 10.3389/fmicb.2017.01265. URL `http://journal.frontiersin.org/article/10.3389/fmicb.2017.01265/full`.

Seth, E. C. and Taga, M. E. Nutrient cross-feeding in the microbial world. *Frontiers in Microbiology*, 5:350, 7 2014. doi: 10.3389/fmicb.2014.00350. URL `http://journal.frontiersin.org/article/10.3389/fmicb.2014.00350/abstract`.

Shibasaki, S. and Mitri, S. Controlling evolutionary dynamics to optimize microbial bioremediation. *Evolutionary Applications*, 13(9):2460–2471, 10 2020. ISSN 1752-4571. doi: 10.1111/eva.13050. URL `https://onlinelibrary.wiley.com/doi/10.1111/eva.13050`.

Shibasaki, S., Mobilia, M., and Mitri, S. Exclusion of the fittest predicts microbial community diversity in fluctuating environments. *Journal of the Royal Society Interface*, 18(183), 10 2021. ISSN 17425662. doi: 10.1098/rsif.2021.0613. URL `https://royalsocietypublishing.org/doi/10.1098/rsif.2021.0613`.

Shin, H.-D., McClendon, S., Vo, T., and Chen, R. R. Escherichia coli Binary Culture Engineered for Direct Fermentation of Hemicellulose to a Biofuel. *Applied and Environmental Microbiology*, 76(24):8150–8159, 12 2010. ISSN 0099-2240. doi: 10.1128/AEM.00908-10. URL `https://journals.asm.org/doi/10.1128/AEM.00908-10`.

Shurtleff, W. and Aoyagi, A. *HISTORY OF NATTO AND ITS RELATIVES (1405-2012)*. SoyInfo Center, 2012. ISBN 9781928914426.

Simutis, R. and Lübbert, A. Bioreactor control improves bioprocess performance. *Biotechnology Journal*, 10(8):1115–1130, 8 2015. ISSN 18606768. doi: 10.1002/biot.201500016. URL `https://onlinelibrary.wiley.com/doi/10.1002/biot.201500016`.

Smith, N. W., Shorten, P. R., Altermann, E., Roy, N. C., and McNabb, W. C. The Classification and Evolution of Bacterial Cross-Feeding. *Frontiers in Ecology and Evolution*, 7(May):1–15, 5 2019. ISSN 2296-701X. doi: 10.3389/fevo.2019.00153. URL `https://www.frontiersin.org/article/10.3389/fevo.2019.00153/full`.

Soares, E. V. and Soares, H. M. V. M. Bioremediation of industrial effluents containing heavy metals using brewing cells of Saccharomyces cerevisiae as a green technology: a review. *Environmental Science and Pollution Research*, 19(4):1066–1083, 5 2012. ISSN 0944-1344. doi: 10.1007/s11356-011-0671-5. URL `http://link.springer.com/10.1007/s11356-011-0671-5`.

Somerville, V., Berthoud, H., Schmidt, R. S., Bachmann, H.-P., Meng, Y. H., Fuchsmann, P., von Ah, U., and Engel, P. Functional strain redundancy and persistent phage infection in Swiss hard cheese starter cultures. *The ISME Journal*, 16(2):388–399, 2 2022. ISSN 1751-7362. doi: 10.1038/s41396-021-01071-0. URL `https://www.nature.com/articles/s41396-021-01071-0`.

Song, C. and Saavedra, S. Structural stability as a consistent predictor of phenological events. *Proceedings of the Royal Society B: Biological Sciences*, 285(1880):20180767, 6 2018. ISSN 0962-8452. doi: 10.1098/rspb.2018.0767. URL `https://royalsocietypublishing.org/doi/10.1098/rspb.2018.0767`.

Sontag, E. D. *Mathematical Control Theory*. Springer New York, NY, 2nd edition, 1998. ISBN 978-1-4612-0577-7. URL `http://www.math.rutgers.edu/`.

Stachowicz, J. J., Whitlatch, R. B., and Osman, R. W. Species Diversity and Invasion Resistance in a Marine Ecosystem. *Science*, 286(5444):1577–1579, 11 1999. ISSN 0036-8075. doi: 10.1126/science.286.5444.1577. URL `https://www.science.org/doi/10.1126/science.286.5444.1577`.

Steel, H., Habgood, R., Kelly, C., and Papachristodoulou, A. Chi.Bio: An open-source automated experimental platform for biological science research. *bioRxiv*, 2019. doi: 10.1101/796516.

Steel, H., Habgood, R., Kelly, C. L., and Papachristodoulou, A. In situ characterisation and manipulation of biological systems with Chi.Bio. *PLOS Biology*, 18(7):e3000794, 7 2020. ISSN 1545-7885. doi: 10.1371/journal.pbio.3000794. URL `https://dx.plos.org/10.1371/journal.pbio.3000794`.

Stein, R. R., Tanoue, T., Szabady, R. L., Bhattarai, S. K., Olle, B., Norman, J. M., Suda, W., Oshima, K., Hattori, M., Gerber, G. K., Sander, C., Honda, K., and Bucci, V. Computer-guided design of optimal microbial consortia for immune system modulation. *eLife*, 7:1–17, 4 2018. ISSN 2050-084X. doi: 10.7554/eLife.30916. URL `https://elifesciences.org/articles/30916`.

Swann, J. R., Rajilic-Stojanovic, M., Salonen, A., Sakwinska, O., Gill, C., Meynier, A., Fança-Berthon, P., Schelkle, B., Segata, N., Shortt, C., Tuohy, K., and Hasselwander, O. Considerations for the design and conduct of human gut microbiota intervention studies relating to foods. *European Journal of Nutrition*, 59(8):3347–3368, 2020. ISSN 14366215. doi: 10.1007/s00394-020-02232-1. URL https://doi.org/10.1007/s00394-020-02232-1.

Tan, J., Yang, X., and Jiang, L. Species ecological similarity modulates the importance of colonization history for adaptive radiation. *Evolution*, 71(6):1719–1727, 2017. ISSN 15585646. doi: 10.1111/evo.13249.

Tilman, A. R., Plotkin, J. B., and Akçay, E. Evolutionary games with environmental feedbacks. *Nature Communications*, 11(1):915, 12 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-14531-6. URL http://www.nature.com/articles/s41467-020-14531-6.

Tilman, D. Resources: A Graphical-Mechanistic Approach to Competition and Predation. *The American Naturalist*, 116(3):362–393, 1980. URL https://www.jstor.org/stable/2463311.

Tilman, D. A consumer-resource approach to community structure. *Integrative and Comparative Biology*, 26(1):5–22, 1986. ISSN 15407063. doi: 10.1093/icb/26.1.5.

Treloar, N. J., Fedorec, A. J. H., Ingalls, B., and Barnes, C. P. Deep reinforcement learning for the control of microbial co-cultures in bioreactors. *PLOS Computational Biology*, 16(4):e1007783, 4 2020. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1007783. URL https://dx.plos.org/10.1371/journal.pcbi.1007783.

Tyakht, A., Kopeliovich, A., Klimenko, N., Efimova, D., Dovidchenko, N., Odintsova, V., Kleimenov, M., Toshchakov, S., Popova, A., Khomyakova, M., and Merkel, A. Characteristics of bacterial and yeast microbiomes in spontaneous and mixed-fermentation beer and cider. *Food Microbiology*, 94(October 2020):103658, 2021. ISSN 10959998. doi: 10.1016/j.fm.2020.103658. URL https://doi.org/10.1016/j.fm.2020.103658.

Vidali, M. Bioremediation. An overview. *Pure and Applied Chemistry*, 73(7):1163–1172, 7 2001. ISSN 1365-3075. doi: 10.1351/pac200173071163. URL http://www.degruyter.com/view/j/pac.2001.73.issue-7/pac200173071163/pac200173071163.xml.

Volterra, V. Fluctuations in the Abundance of a Species considered Mathematically. *Nature*, 118:558–560, 1926. URL http://www.natureasia.com/ja-jp/research/highlight/8406#.UZLPLiubXFA.mendeley.

Vrieze, A., Van Nood, E., Holleman, F., Salojärvi, J., Kootte, R. S., Bartelsman, J. F., Dallinga-Thie, G. M., Ackermans, M. T., Serlie, M. J., Oozeer, R., Derrien, M., Druesne, A., Van Hylckama Vlieg, J. E., Bloks, V. W., Groen, A. K., Heilig, H. G., Zoetendal, E. G., Stroes, E. S., De Vos, W. M., Hoekstra, J. B., and Nieuwdorp, M. Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology*,

143(4):913–916, 2012. ISSN 15280012. doi: 10.1053/j.gastro.2012.06.031. URL http://dx.doi.org/10.1053/j.gastro.2012.06.031.

Wang, J., Liu, L., Ball, T., Yu, L., Li, Y., and Xing, F. Revealing a 5,000-y-old beer recipe in China. *Proceedings of the National Academy of Sciences*, 113(23):6444–6448, 6 2016. ISSN 0027-8424. doi: 10.1073/pnas.1601465113. URL https://pnas.org/doi/full/10.1073/pnas.1601465113.

Wang, Z., Jones, E. W., Mueller, J. M., and Carlson, J. M. Control of ecological outcomes through deliberate parameter changes in a model of the gut microbiome. *Physical Review E*, 101(5):052402, 5 2020. ISSN 2470-0045. doi: 10.1103/PhysRevE.101.052402. URL http://arxiv.org/abs/1912.03412http://dx.doi.org/10.1103/PhysRevE.101.052402.

Watanabe, S. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *Journal of Machine Learning Research*, 11: 3571–3594, 4 2010. ISSN 15324435. URL http://jmlr.org/papers/v11/watanabe10a.html.

Waxman, D. and Gavrilets, S. 20 Questions on Adaptive Dynamics. *Journal of Evolutionary Biology*, 18(5):1139–1154, 8 2005. ISSN 1010061X. doi: 10.1111/j.1420-9101.2005.00948.x. URL http://doi.wiley.com/10.1111/j.1420-9101.2005.00948.x.

Weitz, J. S., Eksin, C., Paarporn, K., Brown, S. P., and Ratcliff, W. C. An oscillating tragedy of the commons in replicator dynamics with game-environment feedback. *Proceedings of the National Academy of Sciences*, 113(47):E7518–E7525, 2016. ISSN 0027-8424. doi: 10.1073/pnas.1604096113. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1604096113.

Xu, X., Xu, P., Ma, C., Tang, J., and Zhang, X. Gut microbiota, host health, and polysaccharides. *Biotechnology Advances*, 31(2):318–337, 2013. ISSN 07349750. doi: 10.1016/j.biotechadv.2012.12.009. URL http://dx.doi.org/10.1016/j.biotechadv.2012.12.009.

Xu, X., Zarecki, R., Medina, S., Ofaim, S., Liu, X., Chen, C., Hu, S., Brom, D., Gat, D., Porob, S., Eizenberg, H., Ronen, Z., Jiang, J., and Freilich, S. Modeling microbial communities from atrazine contaminated soils promotes the development of biostimulation solutions. *ISME Journal*, 13(2):494–508, 2019. ISSN 17517370. doi: 10.1038/s41396-018-0288-5. URL http://dx.doi.org/10.1038/s41396-018-0288-5.

Yunita, D. and Dodd, C. E. Microbial community dynamics of a blue-veined raw milk cheese from the United Kingdom. *Journal of Dairy Science*, 101(6):4923–4935, 2018. ISSN 15253198. doi: 10.3168/jds.2017-14104. URL http://dx.doi.org/10.3168/jds.2017-14104.

Yurtsev, E. A., Conwill, A., and Gore, J. Oscillatory dynamics in a bacterial cross-protection mutualism. *Proceedings of the National Academy of Sciences*, 113(22):201523317, 2016. ISSN 0027-8424. doi: 10.1073/pnas.1523317113. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1523317113.

Curriculum Vitae

The CV of Shota Shibasaki appears in the next pages.

Shota Shibasaki (柴﨑祥太)

Curriculum Vitae. last update: July 5th, 2022

PERSONAL DETAILS

Birth	April 20, 1993
Address	Lausanne, Switzerland
Nationality	Japanese
Mail 1	shibasaki.sh[at]gmail.com
Mail 2	shota-shibasaki[at]18.alumni.u-tokyo.ac.jp[1]
Mail 3	shota.shibasaki[at]unil.ch
ORCID	https://orcid.org/0000-0002-8196-0745
Publons	https://publons.com/researcher/1728635/shota-shibasaki/

EDUCATION

PhD, Life science　　　　　　　　　　　　　　　　　　　Sep. 2018 -

University of Lausanne, Switzerland

> Department of Fundamental Microbiology, Faculty of Biology and Medicine. Ph.D. thesis:
> "Two design frameworks for optimizing microbial community functions"
> Supervisor: Prof. Sara Mitri

Master, Multidisciplinary science　　　　　　　　　　　Apr. 2016 - Mar. 2018

The University of Tokyo, Japan

> Department of general system studies, Graduate School of Arts and Sciences.
> MS thesis: "Cyclic dominance emerges from the two cooperative behaviors in the social amoeba" (based on Shibasaki et al, 2017, Shibasaki and Shimada 2017, and Shibasaki and Shimada 2018. See PUBLICATIONS with †).
> Supervisor: Prof. Masakazu Shimada

Bachelor, Arts and Sciences　　　　　　　　　　　　　Apr. 2012 - Mar. 2016

The University of Tokyo, Japan

> Department of Integrated sciences, College of Arts and Sciences.
> Major: Integrated life sciences (incl. molecular biology, cellular biology, mathematical biology, systems biology, etc).
> Minor: Cognitive and behavioral sciences.

SKILLS

Languages	Japanese (mother tongue)
	English
	Italian (better than French) and French (A1)
Programming	Python 3, C, R, Matlab, Scilab
Software	LaTeX, Illustrator

---

[1]receiving only

PUBLICATIONS

Original paper (peer-reviewed)

*: corresponding author. **: co-corresponding author.

1. [¶] **S. Shibasaki****, M, Mobilia, and S. Mitri. (2021). Exclusion of the fittest predicts microbial community diversity in fluctuating environments [2] *J. R. Soc. Interface.*DOI: 10.1098/rsif.2021.0613 preprint

2. A. Yamauchi, K. Ito, and **S. Shibasaki** (2021). Colonization process determines species diversity via competitive quasi-exclusion. *Ecol. Evol.* DOI:10.1002/ece3.7342

3. [¶] **S. Shibasaki** and S. Mitri (2020). Controlling evolutionary dynamics to optimize microbial bioremediation. *Evol. Appl.* 2020;00:1–12. preprint DOI: 10.1101/2020.07.22.216010

4. **S. Shibasaki*** (2019). The evolutionary game of interspecific mutualism in the multi-species model. *J. Theo. Biol.* (471), 51-58. preprint. DOI: 10.1016/j.jtbi.2019.03.026

5. [†] **S. Shibasaki*** and M. Shimada. (2018). Cyclic dominance emerges from the evolution of two inter-linked cooperative behaviours in the social amoeba. *Proc. R. Soc. B.* 285. 20180905. preprint. DOI: 10.1098/rspb.2018.0905

6. [†] **S. Shibasaki***, Y. Shirokawa, and M. Shimada. (2017). Cooperation Induces Other Cooperation: Fruiting Bodies Promote the Evolution of Macrocysts in *Dictyostelium discoideum. J. Theo. Biol.* (421), 136-145. DOI: 10.1016/j.jtbi.2017.04.002

Conference proceedings (peer-reviewed)

1. [†] **S. Shibasaki*** and M. Shimada. (2017). Stochastic sexual interaction facilitates the evolution of asexual cooperation in the social amoeba. *Proceedings of European Conference of Artificial Life 2017.* 372-379. DOI:10.7551/ecal_a_064

[¶]: relates to Ph.D. thesis. [†]: relates to M.S. thesis

PRESENTATIONS

*: presenter

Contributed talks

- A. Yamauchi, K. Ito, **S. Shibasaki**, and T. Namba (2021). "Continuous irregular dynamics with multiple neutral trajectories permit species coexistence in competitive communities" Annual meeting of Japanese Society of Mathematical Biology. 2021, Sept. 14-16. online.

- **S. Shibasaki***, M. Mobilia, and S. Mitri (2021). "Exclusion of the fittest predicts microbial community diversity in fluctuating environments" Mathematical Population Dynamics, Ecology and Evolution. 2021, April 26-30. Online. recorded talk

---

[2]renamed from older preprint: "Microbial species interactions determine community diversity in fluctuating environments"

- A. Yamauchi, K. Ito, and **S. Shibasaki** (2020). "Species diversity resulted from colonization process with trade-off - Test in rank-abundance relationship" Annual meeting of Japanese Society of Mathematical Biology. 2020, Sept. 20-22. Nagoya, Japan.

- A. Yamauchi, K. Ito, and **S. Shibasaki** (2019). "Theory of evolution of programmed cell death in unicellular organisms" The 67th Annual Meeting of the Ecological Society of Japan. 2020, Mar. 4-8. Nagoya, Japan.[3]

- A. Yamauchi, K. Ito, and **S. Shibasaki** (2019). "Mathematical model of programmed cell death of unicellular organisms" Annual Meeting of Japanese Society of Mathematical Biology. 2019, Sept. 14-16. Tokyo, Japan.

- **S.Shibasaki*** and S. Mitri (2019). "Microbial public goods games in a toxic environment: to degrade or to resist?" The 2019 Congress of the European Society for Evolutionary Biology. 2019. Aug. 19-24. Turku, Finland.

- **S.Shibasaki*** and S. Mitri (2019). "Microbial public goods games in a toxic environment: to degrade or to resist?" Mathematical Models in Ecology and Evolution. 2019, Jul. 16-19 Lyon, France.

- **S.Shibasaki*** and S. Mitri (2019). "Microbial public goods game in a toxic environment; to degrade or to be resistant?" The 66th Annual Meeting of the Ecological Society of Japan. 2019, Mar. 15-19. Kobe, Japan.

- **S. Shibasaki*** (2018). The complex effect of the evolutionary rates on generalized mutualistic communities. 2018 Annual Meeting of the Society for Mathematical Biology and the Japanese Society of Mathematical Biology. 2018, Jul. 8 – 12. Sydney, Australia.

- **Shibasaki*** and M. Shimada. (2018). Cyclic dominance emerges from the two cooperative behaviors in the social amoeba. Game Theory Work shop 2018. Mar. 2-4. Osaka, Japan.

- **S. Shibasaki*** and M. Shimada. (2017). Stochastic sexual interaction facilitates the evolution of asexual cooperation in the social amoeba. European Conference on Artificial Life 2017. 2017, Sep. 4-9. Lyon, France.

Invited talks

- **S. Shibasaki***, M. Mobilia, and S. Mitri "Environmental and demographic stochasticity together changes microbial interactions and diversity" mini-symposium of "Predicting ecological dynamics in fluctuating environments" at SMB2021 (USA). 2021, June 15th. Online.

- **S. Shibasaki***, "Finding the best spatial structures to maximize microbial community functions", Curiosity Seminars, Centre de Biologie Intégrative (France), 2021, May 26th. Online

Contributed poster presentations

- **S. Shibasaki***, M. Sudário, A.Dos. Santos, and S. Mitri. "Finding the best spatial structures to maximize microbial community functions" Annual meeting of Japanese Society of Mathematical Biology. 2021, Sept. 14-16. online.

---

[3]cancelled conference because of COVID-19, but the abstract is available

- **S. Shibasaki***, M. Mobilia, and S, Mitri. (2020). "Fluctuating environments affect the strength of species interactions and diversity in microbial communities similarly", 17-20, Aug., eSMB. (SMB Poster prize) Available here

- **S. Shibasaki***, M. Mobilia, and S, Mitri. (2019). "Intermediate Environmental Switching Rate Maximizes Competitive Exclusion", Kick-off of the Theoretical Biology Network in Western Switzerland, Nov. 26, Lausanne, Switzerland.

- **S. Shibasaki*** and M. Shimada. (2017). "Evolution of cooperation with the multi-game dynamics in the social amoeba". Annual Meeting of Japanese Society of Mathematical Biology. 2017, Oct. 6-8. Hokkaido, Japan. (best poster award)

- **S. Shibasaki***, Y. Shirokawa, and M. Shimada. (2017). "The Evolutionary Game in Sexual Reproduction of the social amoeba." Annual Meeting of Japanese Society of Mathematical Biology. 2016, Sept. 7-9. Fukuoka, Japan.

- **S. Shibasaki***, Y. Shirokawa, and M. Shimada. (2017). "Cooperation in the sexual reproduction of *Dictyostelium discoideum*." Annual Meeting of Society of Evolutionary Studies, Japan. Aug. 25-28. Tokyo, Japan.

GRANTS, PRIZES, ETC.

Grants and/or Fellowships

- PhD fellowship abroad by Nakajima foundation (Sep. 2018- Aug. 2022)

- PhD fellowship in Life Science by University of Lausanne (Sep. 2018- August 2022)

- DC1 by JSPS (Apr. 2018 - Aug. 2018)

- Division of Multi-Disciplinary Sciences 2017 International Conference Travel Grants, 2017, the University of Tokyo.

Prizes

- SMB Poster Prize: Population dynamics Ecology and Evolution. 2020 Annual meeting of Society for Mathematical Biology (SMB).

- Outstanding Master's Thesis Award (2018), Department of Multidisciplinary Sciences, the University of Tokyo

- Best poster award, 27th (2017) Annual meeting of Japanese Society of Mathematical Biology (JSMB)

SUPERVISION

Master students

- Omar Keshk at University of Lausanne, Oct. –Dec. 2021, (First-step project).Topic: fitting the Lotka-Voltera model to data from the lab, understanding species interactions, and analyzing best order of species in the multi-stage chemostat to maximize species diversity. My responsibility: designing projects, consulting programming in R, discussing the results, and advising on academic writing and presentation.

- Marina Oliveira Sudário at University of Lausanne, Mar. 2021– Jan. 2022, (Master project co-supervising with Andrea Dos Santos[4]). Topic: what are the best order of species in the multi-stage chemostat to maximize the degradation of antibiotics?

REVIEW

See my Publon

MEMEBERSHIP

- Japanese Society for Mathematical Biology (JSMB)

- European Society for Mathematical and Theoretical Biology (ESMTB)

- Ecological Society of Japan (ESJ)

- European Society for Evolutionary Biology (ESEB)

OUTREACH

- Doing PhD abroad, available on Youtube (start from 35min, in Japanese). See also this related interview by UmeeT (in Japanese)

---

[4]My responsibilities: supervising statistics, programming, presentations, and academic writing