

# Horizontal Gene Transfer Can Help Maintain the Equilibrium of Microbial Communities

Yuhang Fan,<sup>1,2,3</sup> Yandong Xiao,<sup>1,4</sup> Babak Momeni,<sup>5</sup> and Yang-Yu Liu<sup>1,6</sup>

<sup>1</sup>*Channing Division of Network Medicine, Brigham and Women's Hospital,  
Harvard Medical School, Boston, Massachusetts, 02115, USA*

<sup>2</sup>*School of Mathematical Science, Zhejiang University, Hangzhou, 310027, China*

<sup>3</sup>*State Key Laboratory of Industrial Control Technology,  
Zhejiang University, Hangzhou, 310027, China*

<sup>4</sup>*College of Information Systems and Management,  
National University of Defense Technology, Changsha 410073, China*

<sup>5</sup>*Department of Biology, Boston College, Chestnut Hill, Massachusetts, 02467, USA*

<sup>6</sup>*Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, 02115, USA*

Horizontal gene transfer and species coexistence are two focal points in the study of microbial communities. The evolutionary advantage of horizontal gene transfer has not been well-understood and is constantly being debated. Here we propose a simple population dynamics model based on the frequency-dependent interactions between different genotypes to evaluate the influence of horizontal gene transfer on microbial communities. We find that both structural stability and robustness of the microbial community are strongly affected by the gene transfer rate and direction. An optimal gene flux can stabilize the ecosystem, helping it recover from disturbance and maintain the species coexistence.

## I. INTRODUCTION

Diverse microbial communities, consisting of interacting microorganisms, are essential for their hosts [1–4] or environment [5, 6]. The mechanisms that govern the coexistence of many species in such complex ecosystems are intriguing [7–10]. Another intriguing fact is the presence of horizontal gene transfer (HGT), i.e., the movement of genetic material between microorganisms other than by the transmission of DNA from parent to offspring [11]. Conferring greatly different phenotypes on different strains or species, HGT is central to microbial evolution [12–15]. However, with asexual reproduction being a faster process with lower cost, competition between strains or species can make a beneficial allele spread more efficiently via vertical transfer, rendering a barrier to HGT and contradicting to experimental observations [16]. How HGT can so strongly influence microbial communities still remains an open question [17, 18]. Recently, HGT was observed to play an important role in biofilm formation and public good interactions [19–21], suggesting a possible link between HGT and the stable equilibrium within microbial communities. Furthermore, the boosted HGT was found upon inflammation in the mouse gut microbiota, which can be associated with the dysbiosis of gut microbial community [22], suggesting that HGT could be related to the microbiota dysbiosis. To date, the impact of HGT on the coexistence of different strains within microbial communities has not been fully understood.

Here we propose a simple model to evaluate the impact of HGT on microbial population dynamics. This model can help us understand the evolutionary force that drives HGT and explain how different genotypes can coexist stably in microbial communities. Our results can naturally explain the existence of stable gene flux and reveal how gene transfer rate and direction can affect the microbial system's structural stability and robustness.

## II. MODEL

Consider a community composed of two genotypes, one containing a cooperative allele and the other containing a cheating allele. The microbes with cooperative alleles will behave cooperatively to benefit both themselves and the whole community, while microbes with the cheating allele will only benefit themselves. We assume that these two alleles can transfer horizontally within the microbial community, making microbes alter their behavior accordingly [23–25]. Without loss of generality, we let  $C$  be the fraction of cooperators and  $(1 - C)$  be the fraction of cheaters. The

population dynamics of the microbial community can be described by an ordinary differential equation (ODE) [18, 26]:

$$\frac{dC}{dt} = -\mu C(1 - C) + \frac{\delta}{1 + \delta C} C(1 - C). \quad (1)$$

The first term  $-\mu C(1 - C)$  represents the effect of HGT and the second term  $C(1 - C)\delta/(1 + \delta C)$  represents natural selection. Here  $\mu$  is the rate of horizontal gene transfer with positive sign representing a strong cheating allele transfer effect.  $\delta$  is the selective coefficient of the cooperators.

In conventional population genetic studies, gene alleles are simply classified as beneficial or harmful without considering the concrete population structures [27]. Therefore a genotype's fitness is typically constant. However, it has been revealed that the fitness brought by a gene allele can highly depend on the population structure of microbial communities [28, 29]. As a result of complex and nonlinear interactions among different genotypes within the microbial community [30], an initially beneficial gene allele could become harmful for microorganisms carrying it under some specific population structure and vice versa. In other words, we can not simply regard HGT as another way to disperse beneficial allele in evolution. Instead, in our model, we assume that the selective coefficient  $\delta$  is not a constant but a function of the cooperator fraction  $C$  (Fig.1). The exact functional form of  $\delta(C)$  depends on the properties of the microbial community (see Methods). It turns out that many fundamental properties of the dynamical system (1) can be analytically derived, even without knowing the detailed functional form of  $\delta(C)$ .

### III. RESULTS

#### A. Stability Criteria for Steady States

The population dynamics (1) has two trivial steady states:

$$C^* = 0, \quad C^* = 1, \quad (2)$$

and non-trivial ones ( $0 < C^* < 1$ ) corresponding to the coexistence of cooperators and cheaters,

$$r(C^*) = -\mu + \frac{\delta(C^*)}{1 + \delta(C^*)C^*} = 0, \quad (3)$$

that is, the effects of gene transfer and natural selection balance out. For a steady state  $C^*$ , the necessary and sufficient condition of its Lyapunov stability is  $dC/dt = r(C)C(1 - C) < 0$  in the

right neighborhood of  $C^*$  and  $dC/dt = r(C)C(1 - C) > 0$  in the left neighborhood of  $C^*$ . To prevent our microbial system from evolving to non-coexistent states, we just need to ensure that the two trivial steady states  $C^* = 0$  and  $C^* = 1$  are unstable. Mathematically, since  $C(1 - C) > 0$  always holds on interval  $(0, 1)$ , we have the following stability criteria for the two trivial steady states:

if  $\lim_{C \rightarrow 0^+} r(C) < 0$  (or  $> 0$ ), steady state  $C^* = 0$  is stable (or unstable),

if  $\lim_{C \rightarrow 1^-} r(C) > 0$  (or  $< 0$ ), steady state  $C^* = 1$  is stable (or unstable).

In fact, we can assert from the stability criteria that if neither  $C^* = 0$  nor  $C^* = 1$  is stable, as a result of the intermediate value theorem, system (1) has at least one non-trivial stable steady state, corresponding to the coexistence of cooperators and cheaters.

## B. Stable gene flux

The presence of stable and long-term gene flux has not been fully understood. Previous studies suggest that only very little gene transfer should occur because the carriers of beneficial allele can rapidly outcompete the other microbes before the beneficial allele has chance to transfer [17, 31]. However, in our model the fitness of a genotype is a function of  $C$ , which can naturally lead to the stable gene transfer. Indeed, if Eq. (3) has a non-trivial solution  $C^*$  corresponding to a stable steady state on interval  $(0, 1)$ , then a stable positive gene flux rate  $l = |\mu C^*(1 - C^*)|$  exists around the neighborhood of  $C^*$  (Fig.2b-c). Stable gene flux generally exists in those cases where both natural selection and gene transfer shape population dynamics at equilibrium.

By contrast, in conventional theories the gene transfer direction is always from the genotype that has higher fitness to that with lower fitness (see SI Sec. I). Both  $\mu$  and  $\delta$  are constants and have opposite signs [18]. Moreover, since very large fitness difference is really rare in practice, one can assume that  $\delta > -1$  and  $1 + \delta C > 0$ , rendering  $-\mu + \delta/(1 + \delta C) \neq 0$ . Consequently, the dynamic system (1) has only two trivial steady states:  $C^* = 0$  and  $C^* = 1$ , rendering no coexistence and zero gene flux rate (Fig.2a).

### C. Structural Stability

If starting from any initial state the system can reach a stable steady state with two genotypes coexisting, then we call these two genotypes "globally" coexist. As we discussed above, this occurs if and only if both  $C^* = 0$  and  $C^* = 1$  are unstable. We call the region of the model parameter space that can guarantee the global coexistence of all genotypes as "flexible coexistent region". The volume and shape of this region can be used to quantify the *structural stability* of the dynamic system (1).

Structural stability is a classical notion in dynamic systems theory [32]. A system is considered to be structural stable if smooth change in its model parameters will not change its dynamical behavior. Structural stability plays a key role in maintaining ecosystem equilibrium [33–35]. In this work we use the volume and shape of the flexible coexistent region to quantify a microbial system's structural stability. A larger or more suitable flexible coexistent region will make the microbial system more robust to the environmental disturbance, leading to higher structural stability. We find that HGT can be explained as an efficient way for microbial systems to adjust or even expand their flexible coexistent region to maximize their structural stability in complex environment.

We assume that all the model parameters can affect the selective coefficient of our model, i.e.,  $\delta = \delta(C, \gamma_1, \dots, \gamma_n)$ , where  $(\gamma_1, \dots, \gamma_n)$  are model parameters. Whether a smooth change of  $\gamma_i$  will change the microbial community's dynamical behavior determines its structural stability. Based on the stability criteria we derived above, to insure global coexistence, we just need to keep  $\lim_{C \rightarrow 0^+} r(C) > 0$  and  $\lim_{C \rightarrow 1^-} r(C) < 0$ . Define  $\varphi_1(\gamma_1, \dots, \gamma_n) = \lim_{C \rightarrow 0^+} \frac{\delta(C, \gamma_1, \dots, \gamma_n)}{1 + \delta(C, \gamma_1, \dots, \gamma_n)C}$  and  $\varphi_2(\gamma_1, \dots, \gamma_n) = \lim_{C \rightarrow 1^-} \frac{\delta(C, \gamma_1, \dots, \gamma_n)}{1 + \delta(C, \gamma_1, \dots, \gamma_n)C}$ . We derive that the flexible coexistent region can be expressed by  $\varphi_1(\gamma_1, \dots, \gamma_n) > \mu$  and  $\varphi_2(\gamma_1, \dots, \gamma_n) < \mu$ . This implies that HGT can naturally shape the flexible coexistent region (Fig.3) and hence impact the system's structural stability for arbitrary frequency-dependent fitness form  $\delta(C, \gamma_1, \dots, \gamma_n)$ . Furthermore, we can show that the above analysis can be extended to the case of arbitrary number of genotypes (see Methods).

Define  $A$  to be the volume of flexible coexistent region determined above. The derivation of  $A$  about HGT rate  $\mu$  can be calculated by the following multiple integral of the first kind:

$$\frac{dA}{d\mu} = - \int_{\omega_1} \frac{1}{|\nabla \varphi_1|} dS + \int_{\omega_2} \frac{1}{|\nabla \varphi_2|} dS, \quad (4)$$

where integral regions  $\omega_1$  and  $\omega_2$  are boundaries of flexible coexistent region shaped by  $\varphi_1(\gamma_1, \dots, \gamma_n) > \mu$  and  $\varphi_2(\gamma_1, \dots, \gamma_n) < \mu$ , respectively. Intuitively, a higher (lower) HGT rate can reduce (expand)

the parameter region where cooperators dominate and expand (reduce) the parameter region where cheaters dominate. Therefore the composite effect of the change of HGT rate is the difference of above two effects of cooperators and cheaters (corresponding to the two multiple integral terms in Eq. (4)). As an extended flexible coexistent region corresponds to higher structural stability, we can evaluate the impact of HGT on a microbial community by considering the sign of  $dA/d\mu$ : a positive (or negative)  $dA/d\mu$  suggests a positive (or negative) effect of higher HGT rate on the microbial community's structural stability, respectively.

We validate Eq. (4) for two concrete models: a game theoretical one and an empirical one (see Methods). Interestingly, we find that for the game theoretical fitness model, the volume of the flexible coexistent region spanned by the system parameters  $a_{11}$ ,  $a_{12}$ ,  $a_{21}$  and  $a_{22}$  is invariant with respect to the gene transfer rate  $\mu$ , though its shape does vary over system parameters (Fig.3a-d). For the empirical fitness model (see Methods), increasing  $\mu$  leads to a larger  $A$  and hence stronger structural stability (Fig.3e-g).

#### D. Robustness

In the above analysis of structural stability, we focus on the case of global coexistence, in which microbial system will converge to coexistence of multiple genotypes from any initial state. If Eq. (3) has only one solution in  $(0, 1)$ , then the system will reach a steady state that is independent of the initial condition [36] (Fig.4b-e). However, in practice, ecosystems may have multiple non-trivial steady states. Whether an ecosystem will converge to a particular stable coexistence equilibrium may highly depend on the initial state and the level of environmental disturbances. Indeed, it has been observed that the dynamics of microbial systems often subjects to bifurcation [37–39], that is, if the population of one genotype drops below or over some threshold then the system of coexisting multiple genotypes will collapse and only one genotype will dominate (Fig.4a).

For those systems, we can define robustness as the maximal degree of perturbation such systems can sustain around a stable coexistent equilibrium. A wider range of initial states from which microbial system can evolve to this coexistent equilibrium means a more robust microbial system. Mathematically, if Eq. (3) has multiple solutions on interval  $(0, 1)$ , then we can evaluate the robustness of a stable equilibrium by its minimum distance to its neighbouring unstable equilibrium. From Eq. (3) we know that change of gene transfer rate  $\mu$  can adjust the distance between

solutions of  $r(C) = 0$  (Fig.4f, Fig.S1). Hence HGT is naturally linked to the robustness of such microbial systems. For different dynamic systems and system parameters, the optimal direction and rate of gene transfer that maximize the ecosystem utility could be system-dependent.

#### IV. DISCUSSION

Our analysis reveals the key role of HGT in maintaining a microbial system's equilibrium. By introducing the frequency-dependent fitness and discussing the mutation-selection balance, we offer a simple explanation for the presence of stable gene flux. We quantify the impact of HGT on shaping the flexible coexistence region, and reveal how HGT can influence the structural stability and robustness of microbial communities. Moreover, we validate our scheme for two concrete fitness models where global coexistence can exist. Our analysis can be easily extended to the case of multiple genotypes, suggesting that HGT is an indispensable factor when analyzing microbial population dynamics.

Recently, a theoretical model is proposed to explain the horizontal sweeps via migration among microbial communities [18]. We show that our conclusion still holds under the effect of migration (see SI Sec. III). Note that in our model, for the sake of simplicity, we neglect the cost caused by gene transfer, and the gene transfer rate has no direct impact on organisms' fitness. However, since it requires cells to produce gene carriers such as plasmid to make HGT happen in practice [40, 41], a higher gene transfer rate can lead to higher cost for microbes which transfer their genes to others. Therefore their fitness could be associated with the gene transfer rate  $\mu$ . We develop an extended model to evaluate this case (see SI Sec. IV). In this extended model, the impact of HGT on microbial system still exists, and adjusting  $\mu$  can potentially lead to an improvement of microbial system's structural stability and robustness to environmental perturbation. Furthermore, we also develop an individual-based model to confirm our prediction of the impact of horizontal gene transfer on structural stability (see SI Sec. V).

In this work we demonstrate the key link between structural stability, robustness and HGT for a well-mixed and simplified microbial system. The systematic effect of HGT on more complicated and real microbial systems with spatial pattern and various species remains to be studied. We anticipate that introducing the effect of HGT to various processes of microbial systems, especially those associated with coexistence could help us better understand many intriguing phenomena observed in microbial systems.

## V. METHODS

### A. General structural stability model for multi-genotypes microbial system

For a general microbial system with  $m$  genotypes, we can denote their fitness as a function of population fraction  $C_i$  ( $1 \leq i \leq m, \sum_{i=1}^m C_i = 1$ ), model parameters  $\gamma_j$  ( $1 \leq j \leq n$ ) and gene transfer rate matrix  $\mu$ , i.e.  $f_i(C_1, \dots, C_m, \gamma_1, \dots, \gamma_n, \mu)$ ,  $1 \leq i \leq m$ . Therefore we have that the selective coefficient of genotype  $i$

$$\delta_i = \frac{f_i}{\frac{1}{1-C_i} \sum_{j \neq i} C_j f_j} - 1 = \delta_i(C_1, \dots, C_m, \gamma_1, \dots, \gamma_n, \mu) \quad (5)$$

as a function of population fraction and model parameters (see SI Sec. II). Denote the rate of HGT from genotype  $i$  to genotype  $j$  is  $\mu_{ij}$ . Then we have the following dynamic system

$$\frac{dC_i}{dt} = - \sum_{j=1}^m (\mu_{ij} - \mu_{ji}) C_i C_j + \frac{\delta_i}{1 + \delta_i C_i} C_i (1 - C_i), \quad 1 \leq i \leq m. \quad (6)$$

Conditions

$$\frac{d^2 C_i}{dt^2} > 0 \quad \text{for} \quad C_i = 0, \quad 1 \leq i \leq m \quad (7)$$

and

$$\frac{d^2 C_i}{dt^2} < 0 \quad \text{for} \quad C_i = 1, \quad 1 \leq i \leq m \quad (8)$$

determine the coexistent region of parameter space, which can be expressed as

$$\varphi_k(\gamma_1, \dots, \gamma_n, \mu) > 0 \quad \text{for} \quad 1 \leq k \leq l. \quad (9)$$

Therefore the flexible coexistent region, i.e., the structural stability of a microbial system is associated with HGT rate  $\mu$ .

### B. Game theoretical fitness model

The concrete functional forms of the selective coefficient,  $\delta(C)$ , are various for different microbial communities. We first consider a classic game theoretical fitness model with two strategies cooperating and cheating in a microbial community [42, 43]. The payoff matrix  $(a_{ij})$  for this game (where  $a_{ij}$  represents the payoff of the microbes with strategy  $i$  when encountering microbes with strategy  $j$ ) is given by



	Cooperator	Cheater
Cooperator	$a_{11}$	$a_{12}$
Cheater	$a_{21}$	$a_{22}$

For the sake of normalization, we assume that  $0 < a_{ij} < 1$  for  $i, j = 1, 2$ . The fitness of cooperators ( $f_C$ ) and cheaters ( $f_N$ ) are given by

$$f_C = a_{11}C + a_{12}(1 - C) \quad (10)$$

$$f_N = a_{21}C + a_{22}(1 - C). \quad (11)$$

Then the selective coefficient  $\delta$  can be calculated as follows [27]

$$\delta(C) = \frac{f_C}{f_N} - 1 = \frac{a_{11}C + a_{12}(1 - C)}{a_{21}C + a_{22}(1 - C)} - 1. \quad (12)$$

For the case of evolutionary matrix game, we regard elements  $a_{ij}$  in payoff matrix as model parameters. Since

$$\lim_{C \rightarrow 0} r(C) = -\mu + \frac{a_{12} - a_{22}}{a_{22}} \quad (13)$$

$$\lim_{C \rightarrow 1} r(C) = -\mu + \frac{a_{11} - a_{21}}{a_{11}}, \quad (14)$$

according to our inference mentioned above, we have the flexible coexistent region as

$$a_{12} > (1 + \mu)a_{22}, \quad 1 < a_{12}, a_{22} < 0, \quad (15)$$

$$a_{21} > (1 - \mu)a_{11}, \quad 1 < a_{11}, a_{21} < 0. \quad (16)$$

It can be seen that  $\mu$  has a clear impact on the flexible coexistent region where the two genotypes can stably coexist. A positive  $\mu$  will expand the flexible region corresponding to Eq. (16) but reduce the flexible region corresponding Eq. (15) (Fig.3).

Interestingly, we find that HGT does not change the total volume of the flexible coexistent region for game theoretical fitness model. From Eq. (15) and (16) we can calculate the volume of the flexible coexistent region's projection on the subspace spanned by  $a_{11}, a_{21}$  and  $a_{12}, a_{22}$ , respectively:

$$A_{a_{11}a_{21}} = 1 - \int_0^1 (1 - \mu)x dx = \frac{1 + \mu}{2}, \quad (17)$$

$$A_{a_{12}a_{22}} = \int_0^{\frac{1}{1+\mu}} (1 + \mu)x dx = \frac{1}{2(1 + \mu)}. \quad (18)$$

Without loss of generality, here we suppose that  $\mu > 0$ . For  $\mu \leq 0$ , a similar derivation can be done. Since two conditions shaping the flexible coexistent region are independent, we can get that the volume of flexible coexistent region in the parameter space as

$$A = A_{a_{11}a_{21}}A_{a_{12}a_{22}} = \frac{1}{4}, \quad (19)$$

which is independent of  $\mu$ .

We can also calculate  $dA/d\mu$  for  $\mu > 0$  according to Eq. (4):

$$\begin{aligned} \frac{dA}{d\mu} &= - \int_{\omega_1} \frac{1}{|\nabla\varphi_1|} dS + \int_{\omega_2} \frac{1}{|\nabla\varphi_2|} dS \\ &= - \int_0^1 \int_{(1-\mu)a_{11}}^1 \int_0^{\frac{1}{1+\mu}} \frac{a_{22}}{\sqrt{1+(1+\mu)^2}} \sqrt{1+(1+\mu)^2} da_{22} da_{21} da_{11} \\ &\quad + \int_0^{\frac{1}{1+\mu}} \int_{(1+\mu)a_{22}}^1 \int_0^1 \frac{a_{11}}{\sqrt{1+(1-\mu)^2}} \sqrt{1+(1-\mu)^2} da_{11} da_{12} da_{22} \\ &= - \frac{1}{4(1+\mu)} + \frac{1}{4(1+\mu)} \\ &= 0. \end{aligned}$$

That is, the positive effect and negative effect of HGT cancel out (and the same result also holds for  $\mu < 0$ ).

### C. Empirical fitness model

Recently, a snowdrift game dynamics was proposed to explain experimental observation in yeast microbial community [28]. The following functional forms of fitness  $f_C$  and  $f_N$  were used to fit the experimental observation [28]:

$$f_C = [\lambda + C(1 - \lambda)]^\alpha - \gamma, \quad (20)$$

$$f_N = [C(1 - \lambda)]^\alpha. \quad (21)$$

Here  $\gamma$  is the cost of cooperation and  $\lambda$  is the efficiency of generating total benefits,  $\alpha$  is a model parameter. With suitable parameter choice, this is a snowdrift game (SG) dynamics. Note that the SG game can result into the coexistence of cooperators and cheaters, which is impossible for the prisoner's dilemma (PD) or the mutually beneficial (MB) game, where one strategy (genotype) dominates. Selective coefficient  $\delta$  can be then calculated as follows:

$$\delta(C) = \frac{f_C}{f_N} - 1 = \frac{[\lambda + C(1 - \lambda)]^\alpha - \gamma}{[C(1 - \lambda)]^\alpha} - 1. \quad (22)$$

To guarantee the global coexistence of two genotypes, we just need to guarantee that  $C^* = 0$  and  $C^* = 1$  are unstable steady states. For this empirical fitness model (22) we have

$$\begin{aligned} & \lim_{C \rightarrow 0^+} r(C) \\ &= \lim_{C \rightarrow 0^+} -\mu + \frac{\delta(C)}{1 + \delta(C)C} \\ &= \lim_{C \rightarrow 0^+} -\mu + \frac{[\lambda + C(1 - \lambda)]^\alpha - \gamma - C^\alpha(1 - \lambda)^\alpha}{C^\alpha(1 - \lambda)^\alpha + C\{[\lambda + C(1 - \lambda)]^\alpha - \gamma - C^\alpha(1 - \lambda)^\alpha\}}. \end{aligned}$$

The limitation tends to infinity, which has the same sign as  $\lambda^\alpha - \gamma$ .  $\lambda^\alpha - \gamma = 0$  is a zero measure on the plane so that we can neglect this case. Moreover, we have

$$\begin{aligned} & \lim_{C \rightarrow 1^-} r(C) \\ &= \lim_{C \rightarrow 0^+} -\mu + \frac{\delta(C)}{1 + \delta(C)C} \\ &= \lim_{C \rightarrow 1^-} -\mu + \frac{[\lambda + C(1 - \lambda)]^\alpha - \gamma - C^\alpha(1 - \lambda)^\alpha}{C^\alpha(1 - \lambda)^\alpha + C\{[\lambda + C(1 - \lambda)]^\alpha - \gamma - C^\alpha(1 - \lambda)^\alpha\}} \\ &= -\mu + 1 - \frac{(1 - \lambda)^\alpha}{1 - \gamma}. \end{aligned}$$

Therefore for this empirical fitness model, with a given  $\alpha$ , there is a flexible region in the parameter space spanned by  $\lambda$  and  $\gamma$  where two genotypes can coexist stably with any initial condition. The coexistent region is described as below:

$$\lambda^\alpha - \gamma > 0, \quad (23)$$

$$-\mu + 1 - \frac{(1 - \lambda)^\alpha}{1 - \gamma} < 0. \quad (24)$$

Due to the singularity of function  $r(C)$  near zero, the gene transfer rate  $\mu$  has no impact on the first condition. But clearly  $\mu$  appears in the second condition and a higher positive HGT rate  $\mu$  can lead to a larger flexible parameter region of coexistence, corresponding a higher structural stability. For this case, we have

$$\frac{dA}{d\mu} = \int_{\omega_2} \frac{1}{|\nabla \phi_2|} dS \geq 0. \quad (25)$$

As what we describe above, for comparatively simple case when Eq. (3) only has at most one solution on  $(0, 1)$ , we understand the corresponding ecosystem dynamics very well (Fig.4). However, if Eq. (3) has more than one solution on the interval  $(0, 1)$ , then the robustness of microbial system may be limited.

**Author Contribution** Y.-Y.L. conceived and designed the project. Y.F. performed all the analytical calculations. Y.F. and Y.X. performed numerical simulations. All authors analyzed the results. Y.F. and Y.-Y.L. wrote the manuscript. B.M. and Y.X. edited the manuscript.

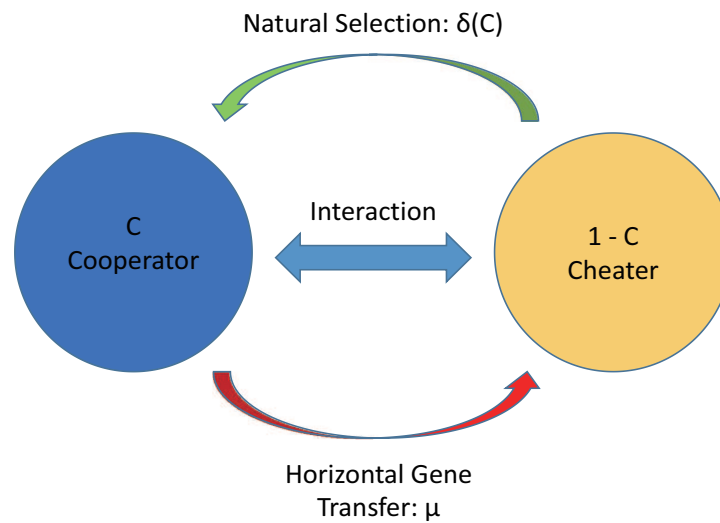
**Acknowledgements** This work was partly supported by the John Templeton Foundation under Award No. 51977. Y. F. is sponsored by Undergraduate Overseas Research Fellowship of Cho Kochen Honors College, Zhejiang University.

- 
- [1] Human Microbiome Project Consortium et al. A framework for human microbiome research. *Nature*, 486(7402):215–221, 2012.
  - [2] Amir Bashan, Travis E Gibson, Jonathan Friedman, Vincent J Carey, Scott T Weiss, Elizabeth L Hohmann, and Yang-Yu Liu. Universality of human microbial dynamics. *Nature*, 534(7606):259–262, 2016.
  - [3] Ruth E Ley, Daniel A Peterson, and Jeffrey I Gordon. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell*, 124(4):837–848, 2006.
  - [4] Ilseung Cho and Martin J Blaser. The human microbiome: at the interface of health and disease. *Nature Reviews Genetics*, 13(4):260–270, 2012.
  - [5] Irina Dana Ofițeru, Mary Lunn, Thomas P Curtis, George F Wells, Craig S Criddle, Christopher A Francis, and William T Sloan. Combined niche and neutral effects in a microbial wastewater treatment community. *Proceedings of the National Academy of Sciences*, 107(35):15345–15350, 2010.
  - [6] Farooq Azam and Richard A Long. Oceanography: Sea snow’ microcosms. *Nature*, 414(6863):495–498, 2001.
  - [7] Joseph Rauch, Jane Kondev, and Alvaro Sanchez. Cooperators trade off ecological resilience and evolutionary stability in public goods games. *Journal of The Royal Society Interface*, 14(127):20160967, 2017.
  - [8] Alvaro Sanchez and Jeff Gore. Feedback between population and evolutionary dynamics determines the fate of social microbial populations. *PLoS biology*, 11(4):e1001547, 2013.
  - [9] Martin J Blaser and Denise Kirschner. The equilibria that allow bacterial persistence in human hosts. *Nature*, 449(7164):843–849, 2007.

- [10] Sarah P Hammarlund, Brian D Connelly, Katherine J Dickinson, and Benjamin Kerr. The evolution of cooperation by the hankshaw effect. *Evolution*, 70(6):1376–1385, 2016.
- [11] Ravi Jain, Maria C Rivera, and James A Lake. Horizontal gene transfer among genomes: the complexity hypothesis. *Proceedings of the National Academy of Sciences*, 96(7):3801–3806, 1999.
- [12] Otto X Cordero, Hans Wildschutte, Benjamin Kirkup, Sarah Proehl, Lynn Ngo, Fatima Hussain, Frederique Le Roux, Tracy Mincer, and Martin F Polz. Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. *Science*, 337(6099):1228–1231, 2012.
- [13] Rachel J Whitaker, Dennis W Grogan, and John W Taylor. Recombination shapes the natural population structure of the hyperthermophilic archaeon *sulfolobus islandicus*. *Molecular biology and evolution*, 22(12):2354–2361, 2005.
- [14] J Peter W Young. Bacteria are smartphones and mobile genes are apps. *Trends in microbiology*, 24(12):931–932, 2016.
- [15] Jacob D Cooper and Benjamin Kerr. Evolution at sutures and centers: Recombination can aid adaptation of spatially structured populations on rugged fitness landscapes. *PLoS computational biology*, 12(12):e1005247, 2016.
- [16] J Peter Gogarten and Jeffrey P Townsend. Horizontal gene transfer, genome innovation and evolution. *Nature Reviews Microbiology*, 3(9):679–687, 2005.
- [17] B Jesse Shapiro, Lawrence A David, Jonathan Friedman, and Eric J Alm. Looking for darwin’s footprints in the microbial world. *Trends in microbiology*, 17(5):196–204, 2009.
- [18] Rene Niehus, Sara Mitri, Alexander G Fletcher, and Kevin R Foster. Migration and horizontal gene transfer divide microbial genomes into multiple niches. *Nature communications*, 6, 2015.
- [19] Jonas Stenl kke Madsen, Mette Burm lle, Lars Hestbjerg Hansen, and S ren Johannes S rensen. The interconnection between biofilm formation and horizontal gene transfer. *FEMS Immunology & Medical Microbiology*, 65(2):183–195, 2012.
- [20] Thomas A Richards and Nicholas J Talbot. Horizontal gene transfer in osmotrophs: playing with public goods. *Nature Reviews Microbiology*, 11(10):720–727, 2013.
- [21] Martin F Polz, Eric J Alm, and William P Hanage. Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends in Genetics*, 29(3):170–175, 2013.
- [22] B rbel Stecher, R my Denzler, Lisa Maier, Florian Bernet, Mandy J Sanders, Derek J Pickard, Manja Barthel, Astrid M Westendorf, Karen A Krogfelt, Alan W Walker, et al. Gut inflammation can boost horizontal gene transfer between pathogenic and commensal enterobacteriaceae. *Proceedings of the*

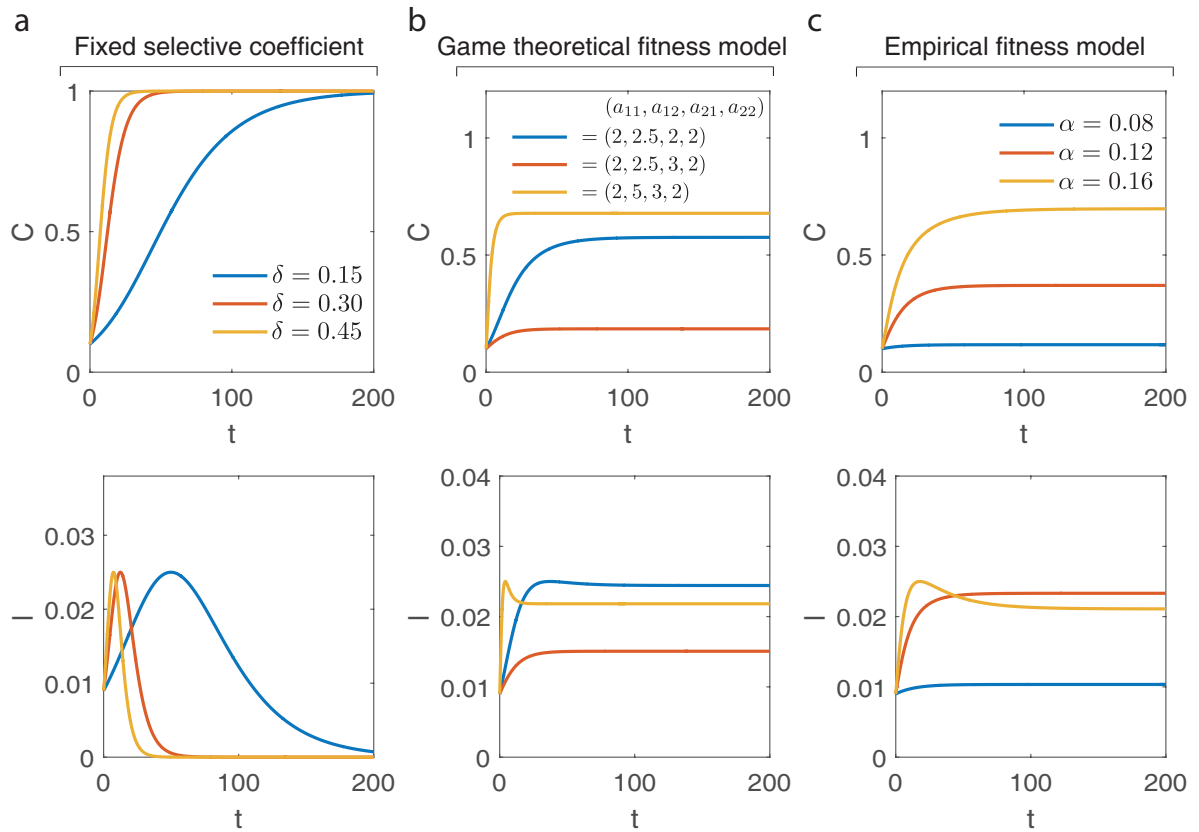
- National Academy of Sciences*, 109(4):1269–1274, 2012.
- [23] Jeff Smith. The social evolution of bacterial pathogenesis. *Proceedings of the Royal Society of London B: Biological Sciences*, 268(1462):61–69, 2001.
- [24] Sorcha E Mc Ginty, Daniel J Rankin, and Sam P Brown. Horizontal gene transfer and the evolution of bacterial cooperation. *Evolution*, 65(1):21–32, 2011.
- [25] Bruce R Levin. Periodic selection, infectious gene exchange and the genetic structure of e. coli populations. *Genetics*, 99(1):1–23, 1981.
- [26] Mario Recker, Oliver G Pybus, Sean Nee, and Sunetra Gupta. The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. *Proceedings of the National Academy of Sciences*, 104(18):7711–7716, 2007.
- [27] John H Gillespie. *Population genetics: a concise guide*. JHU Press, 2010.
- [28] Jeff Gore, Hyun Youk, and Alexander Van Oudenaarden. Snowdrift game dynamics and facultative cheating in yeast. *Nature*, 459(7244):253–256, 2009.
- [29] Sara Mitri, João B Xavier, and Kevin R Foster. Social evolution in multispecies biofilms. *Proceedings of the National Academy of Sciences*, 108(Supplement 2):10839–10846, 2011.
- [30] Richard Law and Ulf Dieckmann. Symbiosis through exploitation and the merger of lineages in evolution. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1402):1245–1253, 1998.
- [31] Frederick M Cohan. The effects of rare but promiscuous genetic exchange on evolutionary divergence in prokaryotes. *The American Naturalist*, 143(6):965–986, 1994.
- [32] Yuri A Kuznetsov. *Elements of applied bifurcation theory*, volume 112. Springer Science & Business Media, 2013.
- [33] Rudolf P Rohr, Serguei Saavedra, and Jordi Bascompte. On the structural stability of mutualistic systems. *Science*, 345(6195):1253497, 2014.
- [34] Ricard V Solé and Joaquim Valls. On structural stability and chaos in biological systems. *Journal of theoretical Biology*, 155(1):87–102, 1992.
- [35] Ian Donohue, Helmut Hillebrand, José M Montoya, Owen L Petchey, Stuart L Pimm, Mike S Fowler, Kevin Healy, Andrew L Jackson, Miguel Lurgi, Deirdre McClean, et al. Navigating the complexity of ecological stability. *Ecology Letters*, 19(9):1172–1185, 2016.
- [36] Vladimir Igorevich ArnolĚzd. *Geometrical methods in the theory of ordinary differential equations*, volume 250. Springer Science & Business Media, 2012.

- [37] Marten Scheffer, Steve Carpenter, Jonathan A Foley, Carl Folke, and Brian Walker. Catastrophic shifts in ecosystems. *Nature*, 413(6856):591–596, 2001.
- [38] Eugene Anatoly Yurtsev, Arolyn Conwill, and Jeff Gore. Oscillatory dynamics in a bacterial cross-protection mutualism. *Proceedings of the National Academy of Sciences*, 113(22):6236–6241, 2016.
- [39] Lei Dai, Daan Vorselen, Kirill S Korolev, and Jeff Gore. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science*, 336(6085):1175–1177, 2012.
- [40] Aya Brown Kav, Goor Sasson, Elie Jami, Adi Doron-Faigenboim, Itai Benhar, and Itzhak Mizrahi. Insights into the bovine rumen plasmidome. *Proceedings of the National Academy of Sciences*, 109(14):5452–5457, 2012.
- [41] James PJ Hall, A Jamie Wood, Ellie Harrison, and Michael A Brockhurst. Source–sink plasmid transfer dynamics maintain gene mobility in soil bacterial communities. *Proceedings of the National Academy of Sciences*, page 201600974, 2016.
- [42] Josef Hofbauer and Karl Sigmund. *Evolutionary games and population dynamics*. Cambridge university press, 1998.
- [43] Josef Hofbauer and Karl Sigmund. Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 40(4):479–519, 2003.



**FIG. 1. Both natural selection and horizontal gene transfer shape microbial population dynamics.** Sub-communities  $C$  (blue) and  $1 - C$  (yellow) represent the fractions of cooperators and cheaters in the microbial community, respectively. Their fitness typically changes according to the interaction among them, rendering a frequency dependent selective coefficient  $\delta(C)$ . Green and red arrows represent the effect of natural selection and horizontal gene transfer, respectively.





**FIG. 2. Alterable selective coefficient leads to a stable gene flux.** The fraction of cooperators  $C$  (top) and the gene flux rate  $l = |\mu C(1 - C)|$  (bottom) are represented as a function of time. **(a)** Fixed selective coefficient ( $\delta = \text{constant}$ ). Genotype carrying beneficial allele quickly dominates the whole population, leading to zero gene flux. Here we take  $\mu = 0.1$  in the simulations. **(b)** Game theoretical fitness model. Stable gene flux exists with frequency dependent selective coefficient ( $\delta(C) = \frac{a_{11}C + a_{12}(1-C)}{a_{21}C + a_{22}(1-C)} - 1$ , see Methods). Here we take  $\mu = 0.1$  in the simulations. **(c)** Empirical fitness model. Stable gene flux also exists with  $\delta(C) = \frac{[\lambda + C(1-\lambda)]^{\alpha} - \gamma}{[C(1-\lambda)]^{\alpha}} - 1$  (see Methods). Here we take  $\lambda = \gamma = 0.5$  and  $\mu = 0.1$  in the simulations.

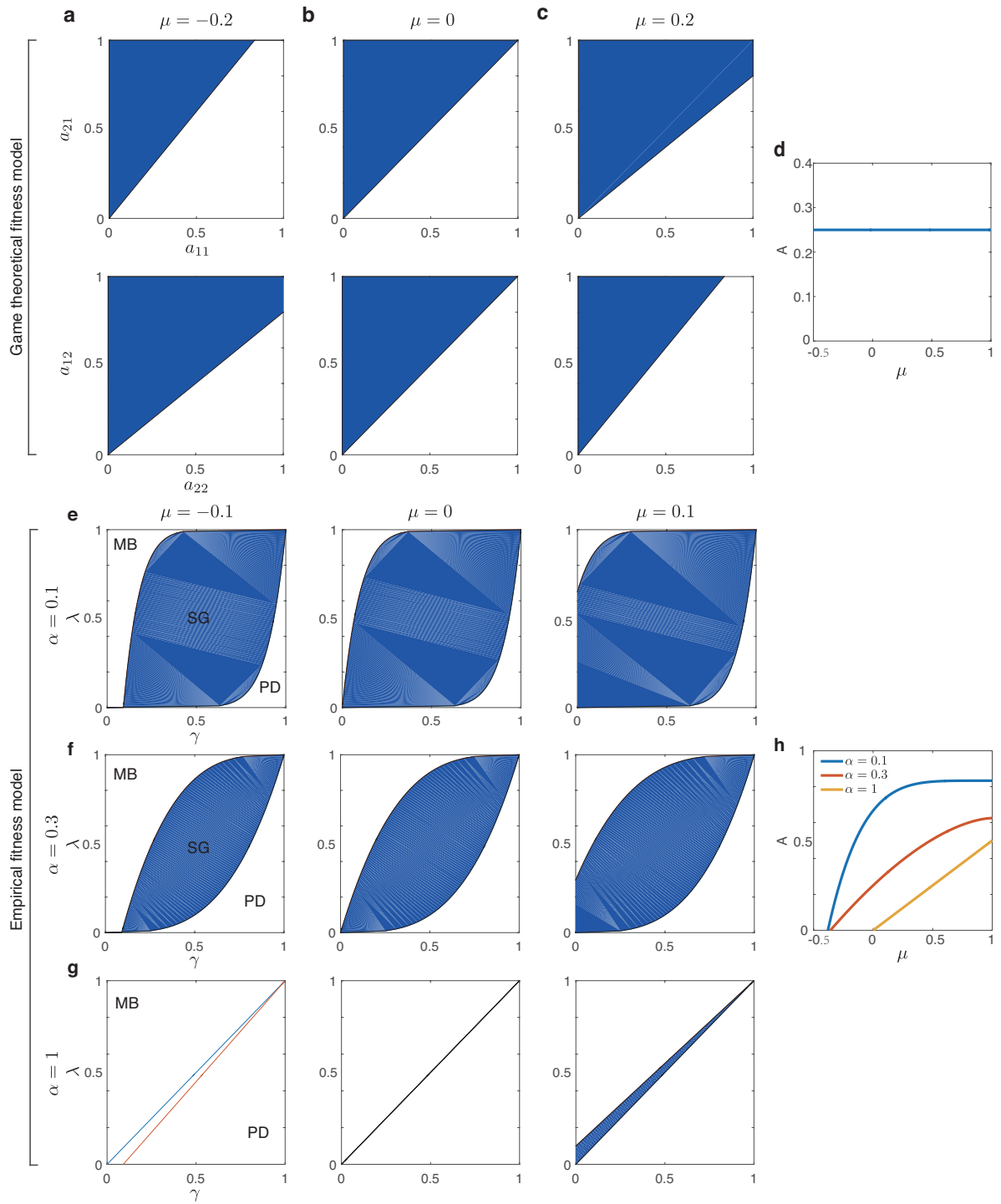
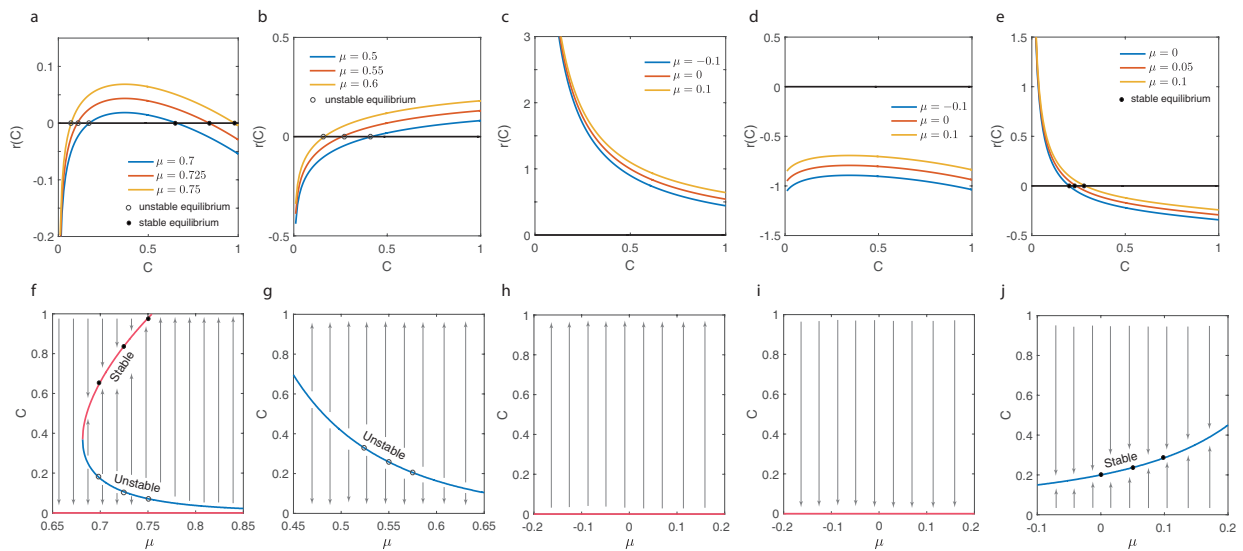


FIG. 3.

**FIG.3: Horizontal gene transfer can adjust the flexible coexistent region of microbial systems. (a-d)** Game theoretical fitness model. The flexible coexistent regions correspond to the blue areas when the gene transfer rate  $\mu = -0.20$  (a), 0 (b) and 0.2 (c). It can be seen that a positive gene transfer rate expands the flexible region spanned by  $a_{11}$  and  $a_{21}$  but reduces the flexible region spanned by  $a_{22}$  and  $a_{12}$ . The volume of the flexible coexistent region ( $A$ , see Methods) spanned by model parameters  $(a_{11}, a_{21}, a_{21}, a_{22})$  is plotted as a function of the gene transfer rate  $\mu$  in (d). **(e-h)** Empirical fitness model. The flexible coexistent region for the case of  $\alpha = 0.1$  (e), 0.3 (f) and 1 (g) with  $\mu = -0.1$ ,  $\mu = 0$  and  $\mu = 0.1$ , respectively. Note that low efficiency ( $\lambda$ ) and high cost ( $\gamma$ ) lead to the prisoner's dilemma (PD) regime, while high efficiency and low cost lead to the mutually beneficial (MB) regime. Both regimes have no coexistence. Only in the snowdrift game (SG) regime can the two genotypes coexist. The volume of the flexible coexistent region is plotted as a function of the gene transfer rate  $\mu$  for different  $\alpha$  in (h). It can be seen that a higher positive gene transfer rate leads to larger flexible coexistent region, thus enhancing the system's structural ability.



**FIG. 4. The impact of horizontal gene transfer on the robustness of microbial systems with an empirical fitness model.** (a-e) The solutions of Eq. (3) correspond to the intersection points with x-axis. The solid circles represent the stable equilibria while the hollow circles mean the unstable steady states, i.e., threshold of according dynamic system. (f-j) With the different model parameters ( $\alpha, \gamma, \lambda$ ), the selection coefficient  $\delta(C)$  (see Eq. (22) in Methods) can lead to various scenarios on the non-trivial equilibria. (a, f) Partial Coexistence. Eq. (3) has two non-trivial solutions corresponding to one stable and one unstable equilibrium, respectively. Whether the microbial system will coverage to a stable coexistent equilibrium highly depends on the initial condition. (b, g) Bi-stable. Only one nontrivial equilibrium exists, but it is unstable. The system will converge to a trivial equilibrium. Whether  $C^* = 0$  or 1 depends on the initial condition. (c, d, h, i) One species dominates. No coexistent equilibrium exists. The system will converge to a trivial equilibrium, regardless of the initial condition. (e, j) Global coexistence. Any initial condition will lead to the only non-trivial equilibrium. Hence the two genotypes always coexist. Parameters used for different cases are : ( $\alpha, \gamma, \lambda$ )=(0.765,0.799,0.744) (a, f); (0.69,0.46,0.32) (b, g); (0.55,0.1,0.8) (c, h); (0.3,0.5,0.1) (d, i); (0.5,0.5,0.55) (e, j).