Full title: High immigration rates critical for establishing emigration-driven diversity in microbial communities

Xiaoli Chen[1, 2], Miaoxiao Wang[3,4], Laipeng Luo[1], Liyun An[5], Xiaonan Liu[1], Yuan Fang[6], Ting Huang[6], Yong Nie[1#] and Xiao-Lei Wu[1, 2, 7#]

[1] College of Engineering, Peking University, Beijing 100871, China

[2] Institute of Ocean Research, Peking University, Beijing 100871, China

[3] Microbial Systems Ecology Group, Institute of Biogeochemistry and Pollutant Dynamics, Department of Environmental Systems Sciences, ETH Zurich, 8006, Zurich, Switzerland

[4] Department of Environmental Microbiology, Eawag: Swiss Federal Institute of Aquatic Sciences, 8600, Duebendorf, Switzerland

[5] College of Architecture and Environment, Sichuan University, Chengdu, 610000, China

[6] School of Resource and Environmental Engineering, Hefei University of Technology, Hefei 230000, China

[7] Institute of Ecology, Peking University, Beijing 100871, China

[#]Corresponding author: Research Scientist, College of Engineering, Peking University.
Tel: +86 10-62759047; Fax: +86 10-62759047; E-mail: nieyong@pku.edu.cn

[#]Corresponding author: Professor, College of Engineering, Peking University.
Tel: +86 10-62759047; Fax: +86 10-62759047; E-mail: xiaolei_wu@pku.edu.cn

Abstract

Unraveling the mechanisms governing the diversity of ecological communities is a central goal in ecology. While microbial dispersal (including the emigration and immigration process) constitutes an important ecological process, the coupling effects of dispersal and microbial competition in microbial diversity are poorly understood. Here, we investigated how microbial dispersal affects the diversity of microbial communities in the presence of inter-species competition, using a generalized Lotka-Volterra model in combination with experimental investigations. Our model shows that emigration reduces the diversity induced by immigration at low immigration rates. We surprisingly find that it increases the diversity of the community when the immigration rates cross a defined threshold, which we identified as $I_{neutral}$. We also found that at high immigration rates, emigration weakens the relative abundance of fast-growing species, and thus enhances the mass effect and increases the diversity. We experimentally confirmed this finding using cocultures of 20 bacterial strains isolated from the soil. Our model further showed that $I_{neutral}$ exists over a wide range of species pool sizes, growth rates, and interspecies interactions, and decreases with the increasing of species pool size, growth rate, and interspecies interaction. Our work deepens the understanding of the effects of dispersal on diversity of natural communities.

Introduction

Microorganisms serve essential roles in an enormous range of biological processes, from global biogeochemical cycles to the regulation of human health[1-6]. The maintenance of microbial diversity is crucial to the sustainability of their ecological functions and services[5-7]. However, according to the principle of competitive exclusion, few species can coexist in an unstructured environment with limited resources, which is inconsistent with the observed diversity in natural ecosystems[8]. This contradiction is often referred to as "the paradox of the plankton"[9]. How microbial diversity is maintained remains one of the unanswered fundamental questions in ecology.

Previous studies suggested that both deterministic processes (including species traits, interspecies interactions, and environmental conditions)[10-11] and stochastic processes (including drift, dispersal, extinction, and speciation)[12-13] govern community assembly, and determine species diversity of ecological communities[14]. However, despite recognizing that ecological processes beyond deterministic factors are important for community assembly, a mechanistic understanding of how stochastic processes shape the diversity of ecological communities remains absent. As a key stochastic process within the assembly of microbial communities, dispersal critically affects the composition and functioning of microbial communities[15-16]. In general, the effects of dispersal to a local community are understood to involve two fundamental processes: (1) immigration into and (2) emigration out of a community[17]. Immigration of species from other communities creates mass effects, which is defined as the flow of individuals from areas of high success to unfavorable areas[18]. Thus, mass effect may rescue the species from extinction in communities where they are weak competitors[19] and always increases

microbial diversity[20]. In contrast, emigration drives the loss of local populations in closed communities, a process that greatly diminishes microbial diversity[21]. However, little is known about how precisely immigration and emigration affect microbial diversity in a collective manner[22].

The relationship between dispersal and microbial diversity in natural communities has previously been investigated primarily by using metagenomic sequencing data[23]. However, quantifying key factors during dispersal of microbial communities, a process responsible for microbial diversity, is difficult due to the complex nature of natural communities. Recently, a 'bottom-up' approach was proposed to establish simple principles regarding how ecological factors govern community assembly and diversity[24-25]. This approach uses simple synthetic microbial consortia and contributes to a better understanding of the assembly of microbial communities under controlled conditions. While several 'bottom-up' studies have previously been proposed that assist in investigating the effects of dispersal[26] (including emigration[27] and immigration[28]). It has been reported that species exhibit a tradeoff between growth rate and competitive ability in the biosystem. High emigration favors the species with high growth rates and low competitive abilities, whereas low emigration favors the species with low growth rates and high competitive abilities[27]. However, if species do not exhibit a trade-off between growth rate and competitive ability, and the immigration process is introduced into the community succession, the precise effects of these both dispersal processes on microbial diversity remain to be quantified. The dispersal including immigration and emigration has been evaluated for synthetic bacterial community[26], but the process of immigration and emigration is coupled. In this study, we decoupled the immigration and emigration process to explore how immigration

and emigration individually and jointly affect the diversity of the communities.

Here, we combined a generalized Lotka-Volterra model with a synthetic microbial consortium, to investigate how emigration and immigration in combination regulate the species diversity of a community. We created a synthetic microbial community composed of 20 strains (from 12 genera and 20 species) to emulate an ecological biosystem. To manipulate the dispersal strength (the number of cells dispersal into/out of a community per time, referred to as immigration rate and emigration rate, respectively), coculture experiments were subject to daily growth/dilution cycles across a series of dilution factors and mass resupplies. We find that at low immigration rates, emigration decreases the Shannon diversity of microbial communities; while at high immigration rates, emigration increases the Shannon diversity of microbial communities. Our results highlight that the relative abundance of fast-growing species plays a key role in the regulation of microbial diversity.

Results

The simulation predicts that at high immigration rates, emigration increases Shannon diversity of microbial communities

We focus on the regime where the diversity of the communities is determined by interspecies competition and dispersal (including immigration and emigration). To probe the effects of emigration and immigration on species diversity, we constructed a generalized Lotka-Volterra model that incorporates the emigration term δN_i and immigration term imm[27, 29-30], as follows:

$$\frac{dN_i}{dt} = N_i\, r_i(1 - \frac{N_i}{K_i} - \frac{\sum_j c_{ij} N_j}{K_i}) - \delta N_i + imm \qquad [1]$$

where N_i is the density of species i; r_i is the maximum growth rate of species i; K_i is the carrying capacity of species i; the interaction coefficient c_{ij} is defined as the effect of species j

on species i. The simulations were initialized with a species pool containing S species. We defined the dilution factor as $dif = \frac{1}{1-\delta}$. The magnitude of the dilution factor represents the emigration rate and determines the fraction of cells discarded during each iteration of the simulation[27]. To study the effect on community succession of the emigration and immigration rates, we finished the model by varying the degree of dilution factor dif and immigration rate imm. To guarantee the communities reach steady states, we set the time to the simulation time to 3×10^3 for each simulation. Microbial diversity was assessed by measuring the Shannon index of the community after simulation[31].

We first created an ecosystem composed of 10 species. We found that when no immigration occurred, the Shannon index was negatively associated with dilution factors (Fig. 1A, Spearman correlation: $R = -0.847$, P-value < 0.001). The rank-abundance curves for the community with high dilution factors were characterized by a more left-skewed abundance distribution (Fig. 1B), indicating that increased emigration of microbial communities resulted in the extinction of low-abundant species within the affected community. Without emigration, the Shannon index positively correlated with the immigration rates (Fig. 1C, Spearman correlation: $R = -0.781$, P-value < 0.001). We found that in communities characterized by higher immigration rates, the gradient of the curves flattened (Fig. 1D). This result points to a higher evenness since the relative abundance of different species was more similar in the community. Our model results indicated that immigration increased microbial diversity by rescuing species from local competitive exclusion, which could be explained by mass effect.

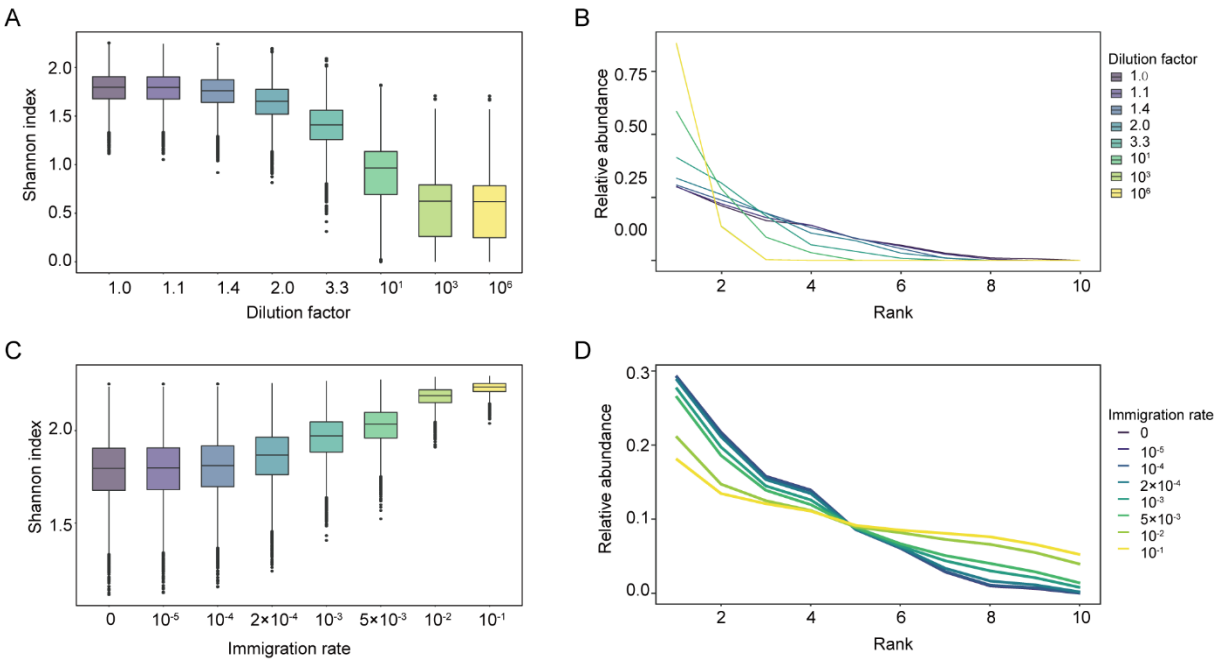

Fig. 1 The Lotka-Volterra model predicts that emigration and immigration had opposite effects on the Shannon diversity of microbial communities. The values of δ were set to 0, 0.1, 0.3, 0.5, 0.7, 0.9, 0.999, and 0.999999. The corresponding values of dif were 1.0, 1.1, 1.4, 2.0, 3.3, 10^1, 10^3, and 10^6 in (A) and (B). (A) The Shannon index of microbial communities negatively correlates with the dilution factors. (B) The effect of emigration on rank-abundance curves for the communities. The X-axis of the curve represents species rank and Y-axis represents the relative abundance of species in a community. (C) The Shannon index of communities positively correlates with immigration rates. (D) The rank-abundance distributions of the communities across a range of immigration rates. Immigration makes the species' abundance distributions even. Figures were drawn using 10^3 independent simulations. For each simulation, we create an ecosystem composed of 10 species. Parameter values used in this simulation: interactions strengths (c_{ij}), independently drawn from normal distribution with a mean of 0.5 and a standard deviation of 0.1; carrying capacities (K_i), independently drawn from normal distribution with a mean of 2 and a standard deviation of 0.2; growth rates (r_i), independently drawn from normal distribution with a mean of 1 and a standard deviation of 0.1.

Notably, when we allowed both emigration and immigration occurred in a microbial community, we observed that increasing the dilution factors resulted in a decrease in Shannon diversity at lower immigration rates ($\leq 3 \times 10^{-3}$). In contrast, increasing the dilution factors resulted in an increase in Shannon diversity at higher immigration rates ($\geq 8 \times 10^{-3}$) (Fig. 2A). As shown in Fig. 2B, at high immigration rates (5×10^{-3} and 10^{-2}), the rank-abundance curves for the community with a high dilution factor (10^6) displayed a less steep decline than

that with a low dilution factor (10^1), indicating higher Shannon diversity for the community with a high dilution factor (10^6). A value of immigration rate (simplified as $I_{neutral}$) was calculated, which represents the Shannon indexes of communities characterized by high dilution factors which are equal to those of communities characterized by low dilution factors. When the immigration rate is over the value of $I_{neutral}$, emigration increases the Shannon diversity of the community. To obtain the value of $I_{neutral}$, we first calculated the slope at each immigration rate by linear fitting of dilution factors and Shannon indexes. Second, we obtained the curve of the slope changing with the immigration rates by nonlinear fitting of the immigration rates and these slopes. The immigration rate on this curve with a slope of zero represents the $I_{neutral}$ (Fig. 2C). When we applied this method to calculate the $I_{neutral}$, we obtained the logarithm of $I_{neutral}$ of -2.25 in these simulations. Together, these results demonstrated that emigration increased the Shannon diversity of the community when the immigration rate was above $I_{neutral}$.

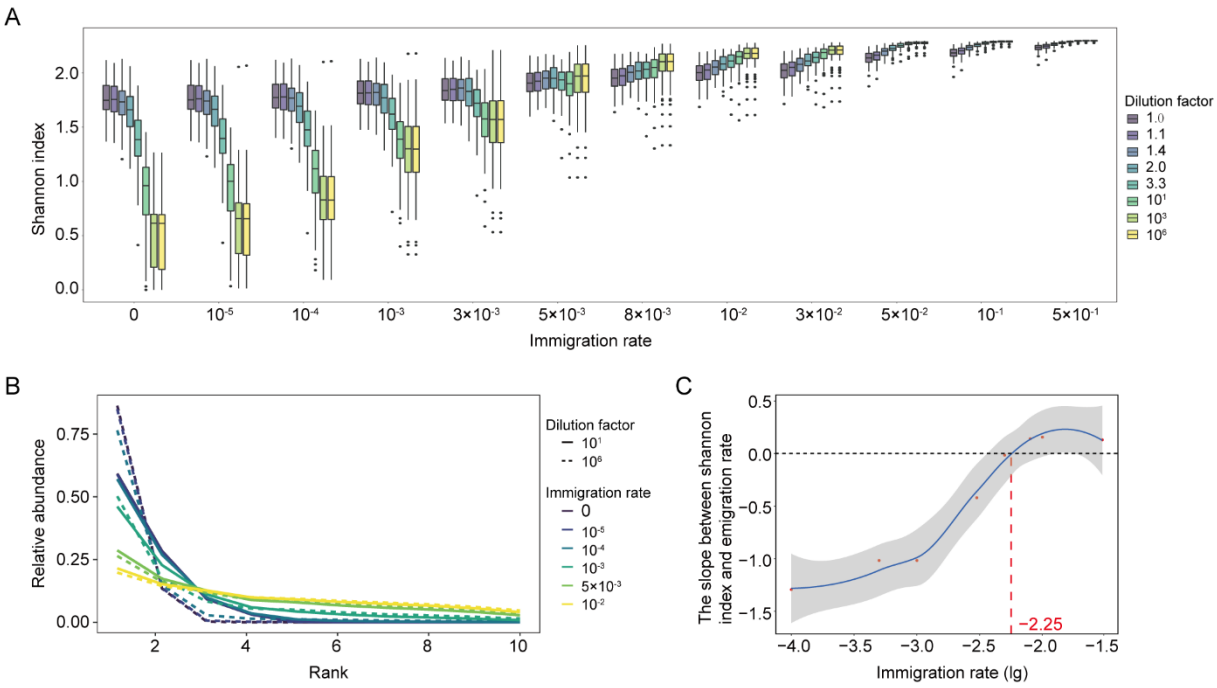

Fig. 2 The Lotka-Volterra model predicts that an increasing immigration rate forces the diversity of the high-emigration community to exceed the diversity of the low-emigration community. (A) At lower immigration rates ($\leqslant 3\times10^{-3}$), Shannon index decreased with increasing dilution factors. At higher immigration rates ($\geqslant 8\times10^{-3}$), Shannon index increased with increasing dilution factors. (B) Effects of both dilution factors and immigration rates on rank-abundance curves for the communities. (C) The X-axis represents the logarithm of the immigration rate. The Y-axis represents the slope of the linear fit between the Shannon index and the dilution factors at each immigration rate. The value of $I_{neutral}$ (2.25) represents the immigration rate at the slope of zero fitted between the Shannon index and dilution factor. Figures were drawn using 10^3 independent simulations. Parameter values used in Fig. 2 are identical to those used in Fig. 1.

Mathematical simulations reveal how dispersal affects species diversity

To better understand why emigration increases the Shannon diversity of the community when the immigration rate is above $I_{neutral}$, we analyzed the role of growth characteristics of species (i.e., growth rate) in community succession. We fitted the relationship between the growth rate of each species in the community and its abundance across a range of emigration and immigration rates. We found that under all simulated immigration and emigration conditions, the slopes obtained by linear fitting of growth rate and relative abundance of species in the community (Fig. 3A) were greater than zero, manifested as significant positive correlations between the growth rate and abundance of species (Fig. 3B, linear regression, $R^2 \geqslant 0.37$, P-value < 0.001). The values of slope reflect the abundance advantages as a result of high growth rates of the species in the community. As shown in Fig. 3B, the slopes decreased with the increasing immigration rates, both under high as well as dilution factors, indicating that immigration weakened the advantages of species as a result of high growth rates in the community. Notably, the slopes with a high dilution factor of 10^6 were larger than that with a low dilution factor of 10^1 when the immigration rates were lower than $\leq 10^{-3}$, which indicated that emigration favored fast-growing species at low immigration rates. In contrast, we found

that when the immigration rates were over 5×10^{-3}, the slopes with a high dilution factor of 10^6 were lower than that with a low dilution factor of 10^1, indicating that emigration reduced the advantage of relative abundance of fast-growing species at high immigration rates. This result indicated that the advantages of the fast-growing species were diminished by high emigration at high immigration rates, which resulted in increased Shannon diversity through emigration at high immigration rates (Fig. 2A).

We next investigated how the advantage as a result of high the growth rate of species affected the relative abundance of species in a community. To this end, we analyzed the changes in the relative abundance of dominant species (the most abundant species in the community)[32], which represents the relative abundance of low-ranking species in of rank-abundance curve. As shown in Fig. S3, the dominant species exhibited faster growth rates compared to the other species of the community assessed. We found that the relative abundance of the dominant species in the community decreased with increasing immigration rates regardless of dilution factors (Fig. 3C), which might be caused by a decline in the advantage of the fast-growing species (Fig. 3B). At low immigration rates ($\leq 10^{-3}$), the relative abundance of dominant species increased with the increasing dilution factors (Fig. 3C). In contrast, at low immigration rates, the Shannon index of the community decreased with the increasing dilution factors (Fig. 2A). As the immigration rate increased to 5×10^{-3}, the relative abundance of dominant species in a community with a high dilution factor of 10^6 was lower than that with a low dilution factor of 10^1 (Fig. 3C), resulting in increasing the Shannon diversity of the high-emigration community more than that of the low-emigration community at high immigration rates (Fig. 2A). Together, our simulation results showed that at high immigration rates, emigration increased the Shannon diversity of

communities, accompanied by a lower relative abundance of dominant species featuring high growth rates.

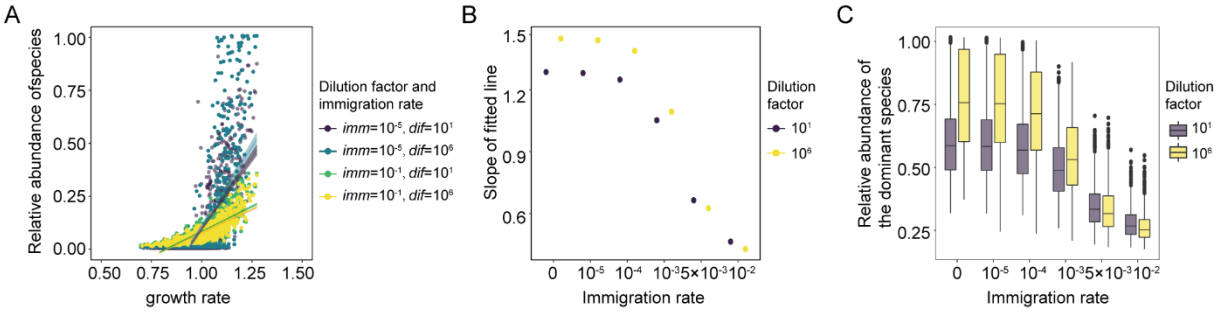

Fig. 3 Results of the 10-species model showing that the dispersal process influences advantages as a result of species' growth rate, and further influences the relative abundance of dominant species featuring high growth rates. (A) Linear regressions between the relative abundance of species and growth rates of species in the community at different dilution factors and immigration rates. (B) Effects of both emigration and immigration on the slopes of linear regressions between the relative abundance of species and growth rates of species in the community. (C) Effects of both emigration and immigration on the relative abundance of dominant species in the communities. Parameter values used in Fig. 3 are identical to those used in Fig. 1.

Confirmation of our Model observations by testing a synthetic consortium

1. Monotonic effects of single dispersal factors on microbial diversity

To experimentally test whether emigration would enhance the Shannon diversity of microbial communities at high immigration levels, we assembled a synthetic consortium containing 20 bacterial isolates from a soil sample taken near Weiming Lake at Peking University (Fig. S1)[33]. To mimic its assembly process, we passaged the consortium across six daily growth/dilution cycles. To regulate the levels of emigration, we changed dilution factor during the passages, with a higher dilution factor mimicking a higher emigration rate (Fig. 4A)[27]. Immigration with different biomass was imposed by adding the 20-strain pool to the local community at the beginning of each growth cycle (Fig. 4B-C).

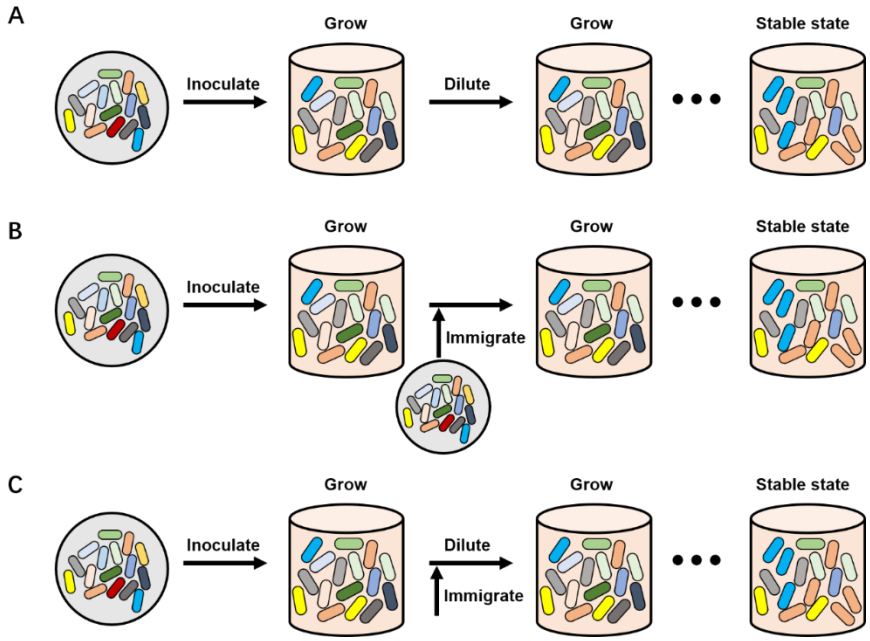

Fig. 4 Schematic of the growth-dilution-immigration scheme employed in co-cultural experiments. The regional pool was composed of 20 species at equal ratios. The microbial communities were grown until a stable equilibrium was reached. (A) Diagram showing how the communities were diluted over a range of dilution factors during succession. The dilution factor determined the fraction of cells discarded, and thus the emigration rates. (B) An immigration community containing all species at equal ratios at multiple immigration rates was added to the cultures during succession. (C) The communities were grown in the presence of both dilution and immigration. Dilution factors and immigration rates were changed over several orders of magnitude.

We first used our system to examine how emigration or immigration individually affects species diversity within the community. When we excluded the immigration process, we found that the Shannon index of the consortia negatively correlated with the dilution factors (Fig. 5A, Spearman correlation: $R = -0.831$, P-value < 0.001). In the rank-abundance curves for the community with different dilution factors, the ranks corresponding to abundance greater than 0 with high dilution factors (10^5 and 10^6) were lower than that with low dilution factors (10^1 and 10^2) (Fig. 5B), suggesting that emigration significantly decreased the Shannon diversity of microbial communities by species loss. To determine how immigration from a source pool affected species diversity, we eliminated "bottleneck effects" by setting the dilution factor to a low level (10^1) to eliminate the, before adding the 20-strain pool to the local community under

different immigration rates. Our results showed that the Shannon index positively correlated with immigration rates (Fig. 5C, Spearman correlation: R = 0.784, P-value < 0.001). The steepness of the abundance rank curves decreased with the increase in immigration rates. In particular, when the immigration rate increased to 10^1, the species richness in the community was significantly higher than that without immigration (Fig. 5D and Fig. S4), suggesting that immigration improved the diversity of the communities through species recolonization. Together, the fundamental effects of a single dispersal process on microbial diversity in our experimental system were in agreement with the effects observed in our model (Fig. 1). Based on these findings, we next investigated the combined effects of the two dispersal processes, including emigration and immigration on the diversity of microbial communities.

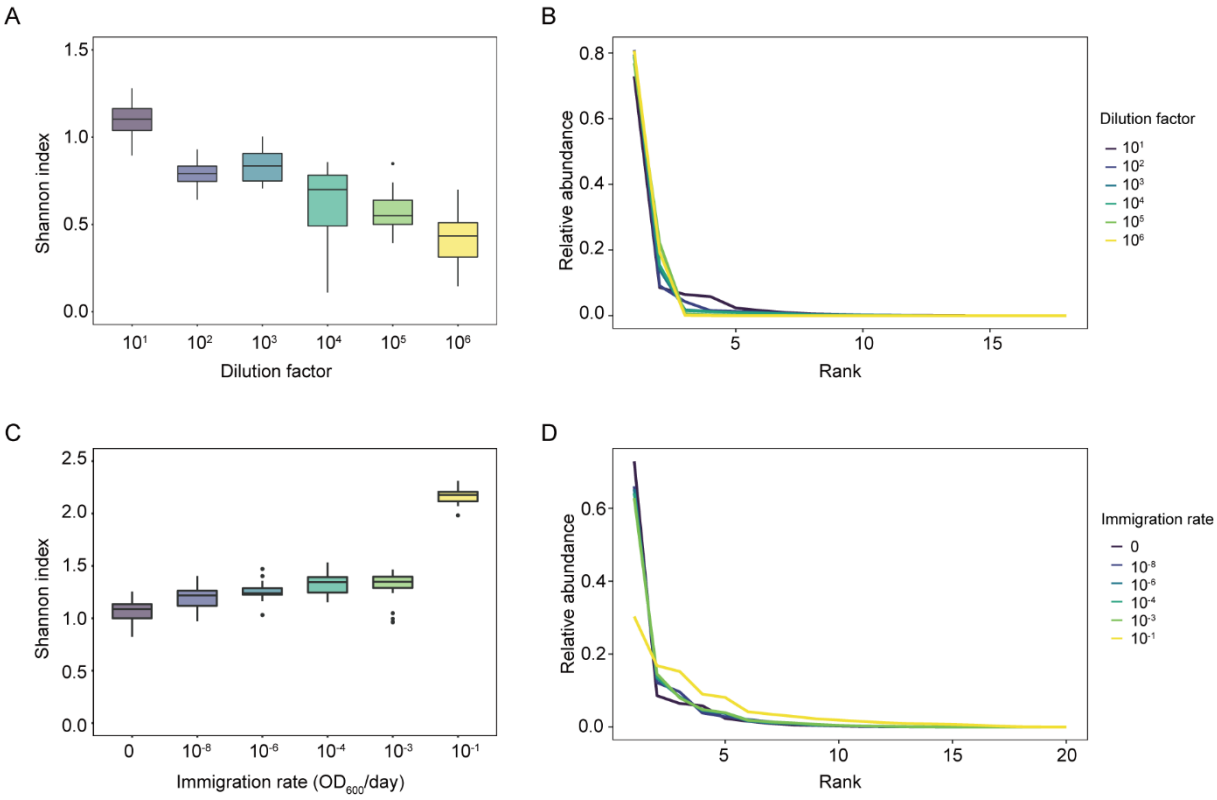

Fig. 5 Impact of the emigration alone and immigration alone on the diversity of bacterial communities and the rank-abundance curves for the communities in the experimental system.

(A) Emigration decreases the Shannon diversity of bacterial communities in the absence of immigration. (B) Rank-abundance plot of the communities under a series of dilution factors. (C) Immigration increases the Shannon diversity of bacterial communities. The consortium was cultured through daily growth-dilution cycles under the dilution factor of 10. (D) The effect of immigration on rank-abundance curves for the communities.

2. Combined effects of immigration and emigration on the diversity of bacterial communities

To assess the combined effects of immigration and emigration in combination on the species diversity of the community, we set up a synthetic consortium consisting of 20 species as described above and passaged the co-cultures through daily growth-dilution cycles under two dilution factors (10^1 and 10^6) and a series of immigration rates. We found that the Shannon index of the microbial community with a lower dilution factor of 10^1 was higher than that with a higher dilution factor of 10^6 in the co-cultures without immigration or with low immigration rates ($\leq 10^{-6}$). However, when we increased the immigration rates to 10^{-4}, the Shannon diversity of the community with a high dilution factor of 10^6 was higher than that with a low dilution factor of 10^1 (Fig. 6A, Wilcoxon test, P-value < 0.001). In agreement with the simulation results (Fig. 2B), under the condition of high immigration rates ($\geq 10^{-3}$), the rank-abundance curves for the high-emigration community were less steep than those for the low-emigration community (Fig. 6B). In this regime, the relative abundance of low-ranking species at the head of the curve in low-emigration communities was lower than those in high-emigration communities. In addition, the relative abundance of rare species at the tail of the curve in high-emigration communities was higher than those in low-emigration communities. These

abundance changes of dominant species and rare species under high immigration resulted in higher Shannon diversity of high-emigration communities. The experimental results confirmed the presence of $I_{neutral}$ over which emigration favored the diversity of microbial communities. Among the 20 strains in the experimental biosystem, the growth rate of Ka was significantly higher than that of the other 19 strains (Fig. S5). Ka dominated the microbial community under the condition of low immigration rates ($\leq 10^{-6}$) (Fig. 6C and Fig. S4), which indicates that a high growth rate resulted in an advantage in the relative abundance of the species in the community. To investigate the relationship between the relative abundance of fast-growing species and the Shannon diversity in the community, the abundance of Ka was calculated. At low immigration rates ($\leq 10^{-6}$), the relative abundance of fast-growing species (i.e., Ka in the experimental biosystem) remained at a high level in the microbial community with a dilution factor of 10^6 (Fig. 6C), while the Shannon index of the microbial community was low. In comparison, at high immigration rates ($\geq 10^{-4}$), the relative abundance of the fast-growing species in the microbial community with a dilution factor of 10^6 was lower than that with a dilution factor of 10^1, resulting in the Shannon index at a high dilution factor of 10^6 higher than that at a low dilution factor of 10^1. This result suggested that high immigration rates decreased the relative abundance of fast-growing species in a high dilution factor more greatly than that in a low dilution factor, and the diversity of high emigration community was greater than that in low emigration community. Overall, our experimental results confirmed our model observations, which predicted that under the high immigration rates, the abundance of fast-growing bacteria decreased with increasing dilution factors, resulting in higher Shannon diversity in the high-emigration community than in the low-emigration community.

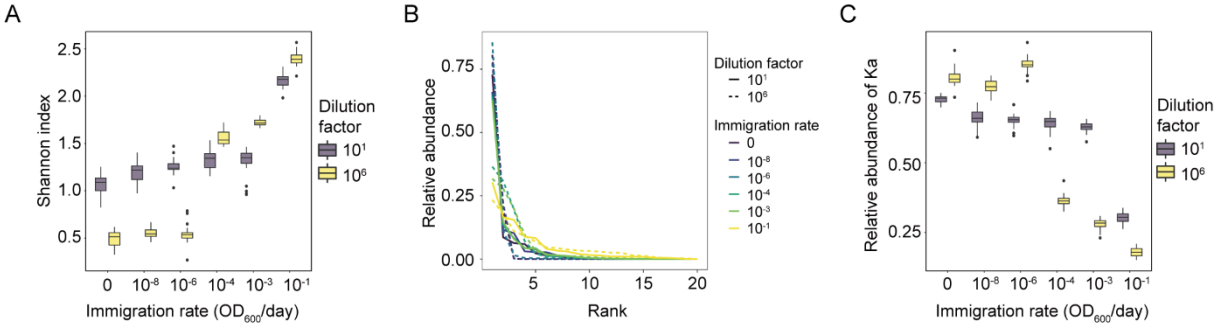

Fig. 6 Experimental results showing effects of the dilution and immigration rates on the diversity of bacterial communities and the relative abundance of the fast-growing species (Ka). (A) Effects of both dilution and immigration rates on the Shannon diversity of bacterial communities. (B) The effect of both emigration and immigration on species abundance distributions of the microbial communities. (C) Relative abundance of the fast-growing species under different dilution factors and across a range of immigration rates.

Extended mathematical simulations reveal that $I_{neutral}$ exists cover a wide range of parameter values

As our model identified an I_{neural} at which emigration rates did not affect the ecosystem diversity in one parameter set, we tested whether the observed $I_{neutral}$ existed in the communities by varying key parameters, including the number of species, interaction strengths, and growth rates. First, we employed the Lotka-Volterra model with different species pool sizes. As shown in Fig. 7A, the $I_{neutral}$ decreased with increasing species pool size. Next, to identify how interaction strength c_{ij} and growth rate r_i of species i determined the $I_{neutral}$ associated with the assembly of a given community, we used a 10-species biosystem. The model results showed that $I_{neutral}$ negatively correlated with c_{ij} while fixing the growth rate of species (Fig. 7B). We also found that $I_{neutral}$ significantly changed whether the values for interaction strength were high or low. In contrast, $I_{neutral}$ had a slight downward trend at intermediate levels of interaction strength ($c_{ij} =$

0.3~0.8), suggesting that species' interaction strength over a wide range had a negligible effect on $I_{neutral}$. When we changed growth rates of species, we found that the increasing growth rates of species for a given interaction strength resulted in a lower $I_{neutral}$. In addition, we found that as the growth rate increased, $I_{neutral}$ steadily decreased (Fig. 7C). Our results indicated that $I_{neutral}$ exists over a wide range of ecological parameters (i.e., varying species pool size, interaction strength, and growth rate), despite the fact that several key ecological parameters (species pool size, interaction strength, and growth rate) of a microbial community can regulate the effects of dispersal on biodiversity.

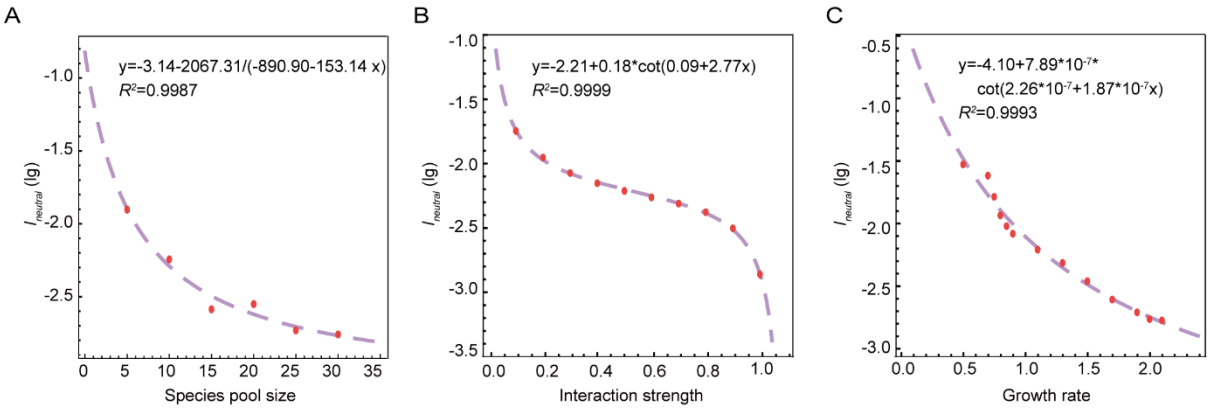

Fig. 7 Species pool size, interaction strength, and growth rate determined $I_{neutral}$. $I_{neutral}$ negatively correlated with the parameters of species pool size (A), interaction strength (B), and growth rate (C). (A) Effects of species pool size on Shannon diversity while fixing interaction strength ($<c_{ij}>$=0.5, std (c_{ij})=0.1) and growth rate ($<r_i>$=1, std (r_i)=0.1). (B) Effects of interaction strength on Shannon diversity while fixing the species pool size (i.e, 10 species) and growth rate ($<r_i>$=1, std (r_i)=0.1). (C) Effects of growth rate on Shannon diversity while fixing the species pool size (i.e, 10 species) and interaction strength ($<c_{ij}>$=0.5, std (c_{ij})=0.1). The dots in the plots represent the mean value over 1000 simulations.

Discussion

Here, we used both model and experiment to show that the loss of species with increasing dilution factors led to a decrease in the Shannon diversity of an ecosystem in the absence of immigration. However, with increasing immigration, the effect of emigration levels on microbial diversity shifted from decreasing to increasing with increasing immigration. At high

immigration levels, emigration weakened the relative abundance of fast-growing species, thus improving the Shannon diversity of the community. Using the generalized Lotka-Volterra model, we further quantified the impact of key parameters (i.e., species pool size, interaction strength, and growth rate) on $I_{neutral}$, which represents the Shannon indexes of communities characterized by high dilution factors which are equal to those of communities characterized by low dilution factors.

Empirical and theoretical studies of biosystems have previously demonstrated that the size of communities is negatively correlated with ecological drift[34-36]. It is well-established that low-abundance microorganisms are more vulnerable to the effects of ecological drift, as slight negative changes in their abundance are likely to result in their extinction[37-38]. In our system, ecological drift plays a stronger role when low-abundance species in communities go randomly extinct with increasing dilution factors, driving microbial diversity with high dilution factors smaller than those with low dilution factors. Our model predicts that fast growers have the advantage of being the majority in the community regardless of emigration levels, which is inconsistent with previous observations[27]. A previous study showed that at high emigration levels, slow growers with high competition coefficients dominated the community[27]. In our model, however, the competition coefficients of interaction forces of fast growers and slow growers are similar. As a result, the slow growers do not have the advantage at low levels of emigration out of the community.

Our findings are consistent with previous studies. These studies claimed that immigration allows bacterial populations to constantly colonize the localities within the local community, causing the rescue effects, and thus substantially mitigating diversity loss. In previous studies,

immigration was a major contributor to buffering against extinction, primarily by recolonization. In those studies, as immigration rates increase, an organism with a competitive disadvantage can be sustained in the microbial system. Conversely, the species with a competitive advantage is more able to dominate the local community[26, 39-41].

Our study examined the role of both emigration and immigration simultaneously. We found that emigration decreased the diversity of the microbial communities at low immigration rates, but increased the diversity of the microbial communities at high immigration rates, different from previous studies focusing on just one single factor[42-43]. Our model and experimental results demonstrate that with the increase in immigration rates, the diversity of the high-emigration community was higher than that of the low-emigration community. High rates of propagule input coupled with small local communities provide a mechanism whereby inferior competitors enter resident communities, thus increasing microbial diversity[35]. This enabled us to observe that under high immigration rates, the diversity of high-emigration communities was higher than that of low-emigration communities. We found that when a new community was added, at low immigration rates, the small input of external communities cannot homogenize the resident community. However, our results clearly show that at high immigration rates, the resident community with a high dilution factor is overwhelmed by mass effects, resulting in the existence of $I_{neutral}$. Specifically, the main reason responsible for the existence of $I_{neutral}$ is that high dilution factors correspond to a small inoculation community size, which is more vulnerable to mass effects[34]. As a consequence of mass effects, the relative abundance of superior native competitors declines, and more inferior competitors colonize, resulting in high Shannon diversity of the microbial community.

Previous studies revealed that the steady-state diversity was strongly dependent on the dynamics of the microbial community on both the species pool size and interaction strength[44]. In our study, in addition to these two parameters, we also determined the effects of dispersal on community diversity when species in communities were set at different growth rates. $I_{neutral}$ is present over a wide range of these parameters. $I_{neutral}$ decreased with increasing these parameters. However, we found here that $I_{neutral}$ differs as a function of species pool size, interaction strengths, and growth rates. Specifically, $I_{neutral}$ changes slightly with the interaction strength over a wide range of interaction strengths ($c_{ij} = 0.3\sim0.8$), while $I_{neutral}$ changes largely with the growth rate over a wide range of growth rates ($r_i = 0.5\sim1.5$). One reason for the observed difference in $I_{neutral}$ as a function of interaction strengths or growth rate may result from the fact that the dominant species show an advantage in growth rate rather than an advantage in interaction strength within the community. Our results suggest that it is possible to quantitatively predict the effect of dispersal on microbial diversity by several key features, such as the species pool size, the statistics of interaction strengths and growth rates.

Although we have found that the dispersal process played a crucial role in controlling microbial diversity, several limitations in our system must be addressed in future work. Our study is based on simple synthetic communities composed of 20 species, which cannot reproduce the complexity of natural systems. For example, our experimental system only includes cooperation and competition interactions. In particular, the Lotka-Volterra model only contains competition interactions, while other interactions such as predation and symbiosis are not included. In addition, in the presence of immigration, only two dilution factors (10^1 and 10^6) were performed in our experiments. To calculate the $I_{neutral}$ in the experimental systems, it is necessary to do

more experiments with more dilution factors. The results of species pool size, growth rate, and interaction strength affecting $I_{neutral}$ should also be further verified experimentally.

Conclusion

In summary, we provide new insight into how microbial diversity depends on both emigration and immigration. At high immigration levels, emigration decreases an advantage in the relative abundance of fast growers, resulting in the increased Shannon diversity of the community. We demonstrate that the effect of immigration and emigration on community assembly is not independent, which enlightens us that when we study both immigration and emigration process, we should consider the two process as a whole. We further quantified the coupling effect of immigration and emigration on the Shannon diversity of microbial communities using mathematical model. Our study suggests that emigration accompanied by high immigration levels may be a strategy for natural microbial communities to regain diversity.

Materials and Methods

Lotka-Volterra simulations

In agreement with previous reports[27], our model assumes that the growth of each species within a microbial community is determined by the combination of its intrinsic growth rate (r_i), its carrying capacity (K_i), and the combination of any interactions each species has with other community members (c_{ij}). Beginning with equal abundances of all species, the dynamics of microbial communities were simulated with different species pool sizes S ranged from 5 to 30, interaction matrices, and the maximum growth rate of species i. For each simulation, the parameter c_{ij} were independently drawn from a normal distribution with a mean ranging from 0.1 to 0.9 and a standard deviation of 0.1[29], and a set of r_i were created from a normal distribution with a mean ranged from 0.2 to 2 and a standard deviation of 0.1.

Calculation of Shannon index

We determined the diversity of each microbial community by calculating its Shannon index H:

$$H = -\sum_{i=1}^{N} p_i \, ln \, p_i \qquad [1]$$

where p_i is the frequency of the *ith* bacterium, N is the total number of bacteria in a community.

Calculation of the value of immigration rate ($I_{neutral}$) at which the effects of emigration levels on microbial diversity switch from a decrease to an increase

Linear fitting was performed between Shannon indexes and dilution factors at each immigration rate and the slope at each immigration rate was obtained. Next, locally weighted regression was performed between slopes and immigration rates. The immigration rate at slope 0 was defined as the $I_{neutral}$ for the specific simulation condition. By fitting species pool size, interaction coefficient, or the growth rate of species i and $I_{neutral}$, the function relation between them was

449 obtained.

450 **Strain isolation and identification**

451 Bacterial strains were isolated from one identical location of soil environment near Weiming
452 Lake at Peking University. Isolated colonies were cultured in Luria Bertani (LB) medium (0.5%
453 yeast extract, 1% tryptone, 1% NaCl) at 30°C until strains grew to stationary phases. All strains
454 were cultured while shaking at 220 rpm. Strains that had been cultured were stored in 20%
455 glycerol at -80°C until further use.

456 The 16S rRNA gene was sequenced via Sanger sequencing using the universal 16S rRNA
457 primers (i.e., 8F and 1492R), yielding assembled sequences of ~1200 nt in usable length.
458 Species names were assigned using the National Center for Biotechnology Information (NCBI)
459 database based on the type strain with the highest sequence similarity relative to the query strain.
460 For phylogenetic analysis, sequences of 16S rRNA gene were aligned using MEGA 7.0.21 and
461 a tree was constructed by the neighbor-joining method.

462 **Correction for differences in DNA extraction efficiency**

463 Since the selected 20 strains, namely 3 strains of Gram-positive bacteria and 17 strains of Gram-
464 negative bacteria, originated from different genera, we found a bias in DNA extraction due to
465 different lysis efficiencies. Therefore, a reference strain was introduced when DNA was
466 extracted. The DNA extraction efficiency of the test strain was corrected by mixing a certain
467 proportion of the test strain and the reference strain, extracting DNA, and measuring the ratio
468 of the number of DNA amplicons between the test strain and the reference strain. The correction
469 coefficient of extraction efficiency was calculated to facilitate data correction in the following
470 data analysis. The DNA extractions were performed by combining steel ball crushing and

471 magnetic bead purification from Beijing Biotech Biotechnology Co., Ltd. Details about the
472 DNA extraction efficiency correction are described in Supporting Information: S1.1–S1.2. The
473 obtained DNA was used for 16S amplicon sequencing of the V3-V4 region.
474 Our reference strain was selected to meet both the following two requirements: 1) The reference
475 strain must be easy to obtain and cultivate; 2) The V3V4 region of the 16S rRNA sequence of
476 the reference strain can be distinguished from the 20 strains used in the experiment. Strain
477 *Pseudomonas geniculata*, which is the most similar to *Agrobacterium tumefaciens* GV 3101
478 among the 20 strains, has only 82.97% sequence similarity in the 16S rRNA V3-V4 region.
479 Based on these conditions, *A. tumefaciens* GV 3101 was chosen as the reference strain.
480 **Emigration and immigration passaging experiments**
481 Frozen stocks of individual species were streaked out on LB agar Petri dishes, grown at 30°C
482 until strains grew up, 36 h for *Cellulosimicrobium cellulans* LMG 16121(T), *Flavobacterium*
483 *ginsengisoli* DCY54(T), *Pseudomonas kunmingensis* HL22-2(T), and *Pseudomonas*
484 *alcaligenes* NBRC 14159(T); 72 h for *Kocuria rosea* DSM 20447(T), 12 h for all other stains.
485 Before we performed emigration and immigration passaging experiments, single colonies were
486 picked and each species was grown separately in 3 ml LB broth for 12-24 h, and then 2%
487 inoculum was transferred to 20 ml LB broth for 24 h. Cells were then collected by centrifuging
488 at 3000 g for 10 minutes and were washed three times using M9 minimal medium (containing
489 M9 salts, 13.52 mg/L $FeCl_3 \cdot 6H_2O$, 1.98 mg/L $MnCl_2 \cdot 4H_2O$, 1.36 mg/L $ZnCl_2$, 0.48 mg/L
490 $CoCl_2 \cdot 6H_2O$, 0.34 mg/L $CuCl_2 \cdot 2H_2O$, 0.48 mg/L $NiCl_2 \cdot 6H_2O$, 0.48 mg/L $Na_2MoO_4 \cdot 2H_2O$,
491 0.35 mg/L Na_2O_3Se, 0.12 mg/L H_3BO_3, 0.24 g/L $MgSO_4$, 0.01 g/L $CaCl_2$) without carbon source.
492 The individual cultures were then adjusted to an OD_{600} =2 and mixed at equal proportions to

493 obtain pre-cultures. The pre-cultures were diluted in M9 minimal medium (containing 4.5 g/L
494 glucose) with a final OD_{600} of 0.1. During the emigration and immigration passaging
495 experiments, cell cultures were incubated in 400 μL 96-well plates (NEST), with each well
496 containing a 120-μL culture. They were propagated via serial passaging once per day across a
497 broad range of dilutions at 30°C and shaken at 600 rpm (Thermal Shaker, AOSHENG, China).
498 After each passage, the remainder of the cultures were stored in 20% glycerol and frozen at -
499 80°C for DNA extraction, amplification, and sequencing. These transfers were repeated 24
500 times for each independent experiment carried out. Community composition was measured at
501 each transfer.

502 **DNA amplification and high-throughput sequencing**

503 To reveal the composition of each culture, 72~96 culture samples were merged into one mixed
504 sample for library construction and sequencing. Barcodes were added to the front of the forward
505 primer for each sample in the library, which were used to split samples in the library later. For
506 the library preparation, amplification of the V3-V4 region of the bacterial 16S rRNA gene was
507 performed to assess the bacterial community using the barcoded primers 338 F: 5′-
508 Barcode+ACTCCTACGGGAGGCAGCA-3′ and 805 R: 5′-
509 GACTACHVGGGTATCTAATCC-3′. After PCR amplification of the extracted DNA, DNA
510 electrophoresis was performed. The brightness of the bands running through electrophoresis
511 reflects the DNA concentration. The concentration of all samples was adjusted to the same by
512 adding ddH_2O. 72~96 samples in the same plate were then mixed in equal volumes. DNA in
513 samples was amplified, purified, and finally sequenced on Illumina HiSeq 2500 platform at
514 Guangdong Magigene Biotechnology, China[45].

515 **16S rRNA gene sequence processing and statistical analyses**

516 The barcodes of the raw sequencing reads were removed using QIIME (v1.9)[46], generating fastq
517 files with the forward and reverse reads. The sequencing data were then demultiplexed,
518 denoised, and clustered into amplicon sequence variants (ASVs) using the QIIME 2 pipeline
519 (v2020.2)[47]. The ASVs were aligned to the 16S rRNA gene sequence of 21 strains (including
520 20 strains and a reference strain) to obtain the species composition information of each sample.
521 The α-diversity analysis was performed using R (v3.6.2). Unless indicated otherwise, the
522 number of replicates was 10^3 for each simulation and 24 for each experiment. For comparative
523 statistics, the *P* values were obtained using an unpaired, two-tailed, Student's t-test.

524 **Competing Interests**

525 The authors declare that they have no conflict of interest.

526 **Acknowledgments**

527 We thank the entire Wu lab for critical discussion of our manuscript. We wish to thank Dr. T.
528 Juelich (UCAS, Beijing) for linguistic assistance during the preparation of this manuscript. This
529 work was supported by the National Key R&D Program of China (2018YFA0902100 and
530 2021YFA0910300), the National Natural Science Foundation of China (91951204, 32130004,
531 32161133023, and 32170113), Sino Swiss Science and Technology Cooperation (SSSTC)
532 Program (IZLCZ0_206044), and the High-performance Computing Platform of Peking
533 University.

534

535 **References**

536 1. Goetghebuer, L.; Bonal, M.; Faust, K.; Servais, P.; George, I. F., The Dynamic of a River
537 Model Bacterial Community in Two Different Media Reveals a Divergent Succession and
538 an Enhanced Growth of Most Strains Compared to Monocultures. *Microb Ecol* **2019**.
539 2. Carlstrom, C. I.; Field, C. M.; Bortfeld-Miller, M.; Muller, B.; Sunagawa, S.; Vorholt, J. A.,
540 Synthetic microbiota reveal priority effects and keystone strains in the Arabidopsis
541 phyllosphere. *Nat Ecol Evol* **2019,** *3* (10), 1445-1454.
542 3. Medlock, G. L.; Carey, M. A.; McDuffie, D. G.; Mundy, M. B.; Giallourou, N.; Swann, J.
543 R.; Kolling, G. L.; Papin, J. A., Inferring Metabolic Mechanisms of Interaction within a
544 Defined Gut Microbiota. *Cell Syst* **2018,** *7* (3), 245-257 e7.
545 4. Flemming, H. C.; Wuertz, S., Bacteria and archaea on Earth and their abundance in biofilms.
546 *Nat Rev Microbiol* **2019,** *17* (4), 247-260.
547 5. Sunagawa, S.; Coelho, L. P.; Chaffron, S.; Kultima, J. R.; Labadie, K.; Salazar, G.;
548 Djahanschiri, B.; Zeller, G.; Mende, D. R.; Alberti, A.; Cornejo-Castillo, F. M.; Costea, P. I.;
549 Cruaud, C.; d'Ovidio, F.; Engelen, S.; Ferrera, I.; Gasol, J. M.; Guidi, L.; Hildebrand, F.;
550 Kokoszka, F.; Lepoivre, C.; Lima-Mendez, G.; Poulain, J.; Poulos, B. T.; Royo-Llonch, M.;
551 Sarmento, H.; Vieira-Silva, S.; Dimier, C.; Picheral, M.; Searson, S.; Kandels-Lewis, S.;
552 Tara Oceans, c.; Bowler, C.; de Vargas, C.; Gorsky, G.; Grimsley, N.; Hingamp, P.; Iudicone,
553 D.; Jaillon, O.; Not, F.; Ogata, H.; Pesant, S.; Speich, S.; Stemmann, L.; Sullivan, M. B.;
554 Weissenbach, J.; Wincker, P.; Karsenti, E.; Raes, J.; Acinas, S. G.; Bork, P., Structure and
555 function of the global ocean microbiome. *Science* **2015,** *348* (6237), 1261359.
556 6. Lozupone, C. A.; Stombaugh, J. I.; Gordon, J. I.; Jansson, J. K.; Knight, R., Diversity,
557 stability and resilience of the human gut microbiota. *Nature* **2012,** *489* (7415), 220-30.
558 7. Bickel, S.; Or, D., Soil bacterial diversity mediated by microscale aqueous-phase processes
559 across biomes. *Nature Communications* **2020,** *11* (1).
560 8. Madsen, E. L., Microorganisms and their roles in fundamental biogeochemical cycles.
561 *Current Opinion in Biotechnology* **2011,** *22* (3), 456-464.
562 9. Hutchinson, G. E., The Paradox of the Plankton. *American Naturalist* **1961,** *95* (882), 137-
563 145.

10. Joseph Fargione, C. S. B., and David Tilman, Community assembly and invasion: An experimental test of neutral versus niche processes. *Proc Natl Acad Sci U S A.* **2003,** *100* (15), 8916–8920.
11. Chesson, P., Mechanisms of maintenance of species diversity. *Annual Reviews* **2000**.
12. Chave, J., Neutral theory and community ecology. *Ecology Letters* **2004,** *7* (3), 241-253.
13. Chase, J. M., Stochastic community assembly causes higher biodiversity in more productive environments. *Science* **2010,** *328* (5984), 1388-1391.
14. Jizhong Zhou, Y. D., Ping Zhang, Kai Xue, Yuting Liang, Joy D. Van Nostrand, Yunfeng Yang, Zhili He, Liyou Wu, David A. Stahl, Terry C. Hazen, James M. Tiedje, and Adam P. Arkin, Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. *Proc Natl Acad Sci U S A.* **2014,** *111* (9), E836-E845.
15. Albright, M. B. N.; Chase, A. B.; Martiny, J. B. H., Experimental Evidence that Stochasticity Contributes to Bacterial Composition and Functioning in a Decomposer Community. *MBio* **2019,** *10* (2).
16. Walters, K. E.; Capocchi, J. K.; Albright, M. B. N.; Hao, Z.; Brodie, E. L.; Martiny, J. B. H., Routes and rates of bacterial dispersal impact surface soil microbiome composition and functioning. *The ISME Journal* **2022**.
17. Kinnunen, M.; Dechesne, A.; Proctor, C.; Hammes, F.; Johnson, D.; Quintela-Baluja, M.; Graham, D.; Daffonchio, D.; Fodelianakis, S.; Hahn, N.; Boon, N.; Smets, B. F., A conceptual framework for invasion in microbial communities. *ISME J* **2016,** *10* (12), 2773-2775.
18. Shmida, A.; Wilson, M. V., Biological Determinants of Species-Diversity. *J Biogeogr* **1985,** *12* (1), 1-20.
19. Vellend, M., Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology* **2010,** *85* (2), 183–206.
20. Custer, G. F.; Bresciani, L.; Dini-Andreote, F., Ecological and Evolutionary Implications of Microbial Dispersal. *Frontiers in Microbiology* **2022,** *13*.
21. 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.; Gonzalez, A., The

metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters* **2004,** *7* (7), 601-613.

22. Albright, M. B. N.; Martiny, J. B. H., Dispersal alters bacterial diversity and composition in a natural community. *The ISME Journal* **2017,** *12* (1), 296-299.

23. Richter-Heitmann, T.; Hofner, B.; Krah, F.-S.; Sikorski, J.; Wüst, P. K.; Bunk, B.; Huang, S.; Regan, K. M.; Berner, D.; Boeddinghaus, R. S.; Marhan, S.; Prati, D.; Kandeler, E.; Overmann, J.; Friedrich, M. W., Stochastic Dispersal Rather Than Deterministic Selection Explains the Spatio-Temporal Distribution of Soil Bacteria in a Temperate Grassland. *Frontiers in Microbiology* **2020,** *11*.

24. Meroz, N.; Tovi, N.; Sorokin, Y.; Friedman, J., Community composition of microbial microcosms follows simple assembly rules at evolutionary timescales. *Nature Communications* **2021,** *12* (1).

25. Gore, J., Simple organizing principles in microbial communities. *Curr Opin Microbiol* **2018,** *45*, 195-202.

26. Fodelianakis, S.; Lorz, A.; Valenzuela-Cuevas, A.; Barozzi, A.; Booth, J. M.; Daffonchio, D., Dispersal homogenizes communities via immigration even at low rates in a simplified synthetic bacterial metacommunity. *Nat Commun* **2019,** *10* (1), 1314.

27. Abreu, C. I.; Friedman, J.; Andersen Woltz, V. L.; Gore, J., Mortality causes universal changes in microbial community composition. *Nat Commun* **2019,** *10* (1), 2120.

28. Vila, J. C. C.; Jones, M. L.; Patel, M.; Bell, T.; Rosindell, J., Uncovering the rules of microbial community invasions. *Nat Ecol Evol* **2019,** *3* (8), 1162-1171.

29. Friedman, J.; Higgins, L. M.; Gore, J., Community structure follows simple assembly rules in microbial microcosms. *Nat Ecol Evol* **2017,** *1* (5), 109.

30. Marino, S.; Baxter, N. T.; Huffnagle, G. B.; Petrosino, J. F.; Schloss, P. D., Mathematical modeling of primary succession of murine intestinal microbiota. *Proc Natl Acad Sci U S A* **2014,** *111* (1), 439-44.

31. Evans, S.; Martiny, J. B.; Allison, S. D., Effects of dispersal and selection on stochastic assembly in microbial communities. *ISME J* **2017,** *11* (1), 176-185.

32. Diaz-Colunga, J.; Lu, N.; Sanchez-Gorostiaga, A.; Chang, C.-Y.; Cai, H. S.; Goldford, J. E.;

Tikhonov, M.; Sánchez, Á., Top-down and bottom-up cohesiveness in microbial community coalescence. *Proceedings of the National Academy of Sciences* **2022,** *119* (6).

33. Liu, X.; Wang, M.; Nie, Y.; Wu, X.-L., Isolation Chip Increases Culturable Bacterial Diversity and Reduces Cultivation Bias. *Current Microbiology* **2021,** *78* (5), 2025-2032.

34. Nemergut, D. R.; Schmidt, S. K.; Fukami, T.; O'Neill, S. P.; Bilinski, T. M.; Stanish, L. F.; Knelman, J. E.; Darcy, J. L.; Lynch, R. C.; Wickey, P.; Ferrenberg, S., Patterns and processes of microbial community assembly. *Microbiol Mol Biol Rev* **2013,** *77* (3), 342-56.

35. Orrock, J. L.; Watling, J. I., Local community size mediates ecological drift and competition in metacommunities. *Proc Biol Sci* **2010,** *277* (1691), 2185-91.

36. Orrock, J. L.; Fletcher, R. J., Changes in community size affect the outcome of competition. *American Naturalist* **2005,** *166* (1), 107-111.

37. Chase, J. M.; Myers, J. A., Disentangling the importance of ecological niches from stochastic processes across scales. *Philos Trans R Soc Lond B Biol Sci* **2011,** *366* (1576), 2351-63.

38. Petro, C.; Starnawski, P.; Schramm, A.; Kjeldsen, K. U., Microbial community assembly in marine sediments. *Aquatic Microbial Ecology* **2017,** *79* (3), 177-195.

39. Stegen, J. C.; Lin, X.; Fredrickson, J. K.; Chen, X.; Kennedy, D. W.; Murray, C. J.; Rockhold, M. L.; Konopka, A., Quantifying community assembly processes and identifying features that impose them. *ISME J* **2013,** *7* (11), 2069-79.

40. Stegen, J. C.; Lin, X.; Fredrickson, J. K.; Konopka, A. E., Estimating and mapping ecological processes influencing microbial community assembly. *Front Microbiol* **2015,** *6*, 370.

41. Sloan, W. T.; Lunn, M.; Woodcock, S.; Head, I. M.; Nee, S.; Curtis, T. P., Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental Microbiology* **2006,** *8* (4), 732-740.

42. Hu, J.; Amor, D. R.; Barbier, M.; Bunin, G.; Gore, J., Emergent phases of ecological diversity and dynamics mapped in microcosms. **2021**.

43. Gokhale, S.; Conwill, A.; Ranjan, T.; Gore, J., Migration alters oscillatory dynamics and promotes survival in connected bacterial populations. *Nat Commun* **2018,** *9* (1), 5273.

44. Jiliang Hu, D. R. A., Matthieu Barbier, Guy Bunin, Jeff Gore, Emergent phases of ecological diversity and dynamics mapped in microcosms. *Science* **2022,** *378*, 85-89.
45. Zhou, D.; Feng, H.; Schuelke, T.; De Santiago, A.; Zhang, Q.; Zhang, J.; Luo, C.; Wei, L., Rhizosphere Microbiomes from Root Knot Nematode Non-infested Plants Suppress Nematode Infection. *Microb Ecol* **2019,** *78* (2), 470-481.
46. J Gregory Caporaso, J. K., Jesse Stombaugh, Kyle Bittinger, Frederic D Bushman, Elizabeth K Costello, Noah Fierer, Antonio Gonzalez Peña, Julia K Goodrich, Jeffrey I Gordon, Gavin A Huttley, Scott T Kelley, Dan Knights, Jeremy E Koenig, Ruth E Ley, Catherine A Lozupone, Daniel McDonald, Brian D Muegge, Meg Pirrung, Jens Reeder, Joel R Sevinsky, Peter J Turnbaugh, William A Walters, Jeremy Widmann, Tanya Yatsunenko, Jesse Zaneveld & Rob Knight, QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **2010,** *7* (5), 335-336.
47. Bolyen, E.; Rideout, J. R.; Dillon, M. R.; Bokulich, N. A.; Abnet, C. C.; Al-Ghalith, G. A.; Alexander, H.; Alm, E. J.; Arumugam, M.; Asnicar, F.; Bai, Y.; Bisanz, J. E.; Bittinger, K.; Brejnrod, A.; Brislawn, C. J.; Brown, C. T.; Callahan, B. J.; Caraballo-Rodriguez, A. M.; Chase, J.; Cope, E. K.; Da Silva, R.; Diener, C.; Dorrestein, P. C.; Douglas, G. M.; Durall, D. M.; Duvallet, C.; Edwardson, C. F.; Ernst, M.; Estaki, M.; Fouquier, J.; Gauglitz, J. M.; Gibbons, S. M.; Gibson, D. L.; Gonzalez, A.; Gorlick, K.; Guo, J.; Hillmann, B.; Holmes, S.; Holste, H.; Huttenhower, C.; Huttley, G. A.; Janssen, S.; Jarmusch, A. K.; Jiang, L.; Kaehler, B. D.; Kang, K. B.; Keefe, C. R.; Keim, P.; Kelley, S. T.; Knights, D.; Koester, I.; Kosciolek, T.; Kreps, J.; Langille, M. G. I.; Lee, J.; Ley, R.; Liu, Y. X.; Loftfield, E.; Lozupone, C.; Maher, M.; Marotz, C.; Martin, B. D.; McDonald, D.; McIver, L. J.; Melnik, A. V.; Metcalf, J. L.; Morgan, S. C.; Morton, J. T.; Naimey, A. T.; Navas-Molina, J. A.; Nothias, L. F.; Orchanian, S. B.; Pearson, T.; Peoples, S. L.; Petras, D.; Preuss, M. L.; Pruesse, E.; Rasmussen, L. B.; Rivers, A.; Robeson, M. S., 2nd; Rosenthal, P.; Segata, N.; Shaffer, M.; Shiffer, A.; Sinha, R.; Song, S. J.; Spear, J. R.; Swafford, A. D.; Thompson, L. R.; Torres, P. J.; Trinh, P.; Tripathi, A.; Turnbaugh, P. J.; Ul-Hasan, S.; van der Hooft, J. J. J.; Vargas, F.; Vazquez-Baeza, Y.; Vogtmann, E.; von Hippel, M.; Walters, W.; Wan, Y.; Wang, M.; Warren, J.; Weber, K. C.; Williamson, C. H. D.; Willis, A. D.; Xu, Z. Z.; Zaneveld, J. R.;

680 Zhang, Y.; Zhu, Q.; Knight, R.; Caporaso, J. G., Reproducible, interactive, scalable and
681 extensible microbiome data science using QIIME 2. *Nat Biotechnol* **2019,** *37* (8), 852-857.

682

683

Supporting Information

Full title: High immigration rates critical for establishing emigration-driven diversity in microbial communities

Xiaoli Chen[1, 2], Miaoxiao Wang[1], Laipeng Luo[1], Liyun An[3], Yuan Fang[4], Ting Huang, Yong Nie[1#] and Xiao-Lei Wu[1, 2, 5#]

[1] College of Engineering, Peking University, Beijing 100871, China

[2] Institute of Ocean Research, Peking University, Beijing 100871, China

[3] College of Architecture and Environment, Sichuan University, Chengdu, China

[4] School of Resource and Environmental Engineering, Hefei University of Technology, Hefei 230000, China

[5] Institute of Ecology, Peking University, Beijing 100871, China

[#]Corresponding author: Research Scientist, College of Engineering, Peking University.

Tel: +86 10-62759047; Fax: +86 10-62759047; E-mail: nieyong@pku.edu.cn

[#]Corresponding author: Professor, College of Engineering, Peking University.

Tel: +86 10-62759047; Fax: +86 10-62759047; E-mail: xiaolei_wu@pku.edu.cn

S1 Construction of community structure detection method

S1.1 Validation of DNA extraction method

The 20-strain collection is phylogenetically diverse and spans 20 species across 12 genera and eight families. Since the selected 20 strains are from different genera (Fig. S1), there is a preference for DNA extraction. The cloning library experiment was performed to preliminarily judge the difference in DNA extraction efficiency of the 20 strains. After 20 strains were mixed in the same proportion, total DNA was extracted, and the V3V4 region sequence of 16S rRNA was then amplified. The amplified sequence was constructed into 4 parallel clone libraries. Clone library 1 contained species Ag, Ac, Av, Cc Sf, and Lm. Clone library 2 was consisted of species Ba, Av, Ca, Ka, Fb, and Pk. Species Bp, Av, Fg, Kr, Pa, and Ki were included in clone library 3. Clone library 4 was composed of species Ao, At, Av, Lf, and Pg. A total of 60 single clones were picked and sequenced to determine the actual number of clones of different strains in each clone library. The results showed that different strains differed in the number of clones in the clone library, which proved that there were indeed differences among 20 strains in the DNA extraction efficiency (Figure S2).

S1.2 DNA extraction efficiency correction

To accurately determine the DNA extraction efficiency of 20 strains for later community composition correction, reference strain *Agrobacterium tumefaciens* GV 3101 was introduced. According to the preliminary differences in the DNA extraction efficiency of the 20 strains in the clone library results (Figure S2), the 20 strains were divided into two groups (Table S1), the group with higher DNA extraction efficiency and the group with lower DNA extraction efficiency, respectively. Each strain was mixed with *A. tumefaciens* GV 3101 according to the

724 relative proportion gradient in Table S1. The DNA of the mixtures was extracted. V3-V4
725 sequence region was then amplified to obtain the relative abundance of *A. tumefaciens* GV 3101
726 and the target strain at each concentration. A linear fit between the ratio obtained by sequencing
727 and the initial mixing ratio was performed. The slope of the fit was defined as the extraction
728 correction coefficient of the target strain (Table S2).
729

S2 Supporting Figures

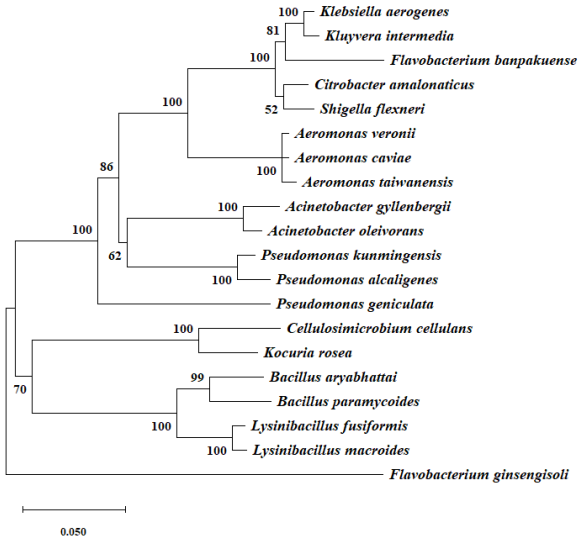

Figure S1 Phylogenetic tree of the 20 bacterial species used in this study. The tree was constructed based on the 16S rRNA genes and the branch lengths indicated the number of substitutions per base pair. Scores on nodes indicate the posterior probability.

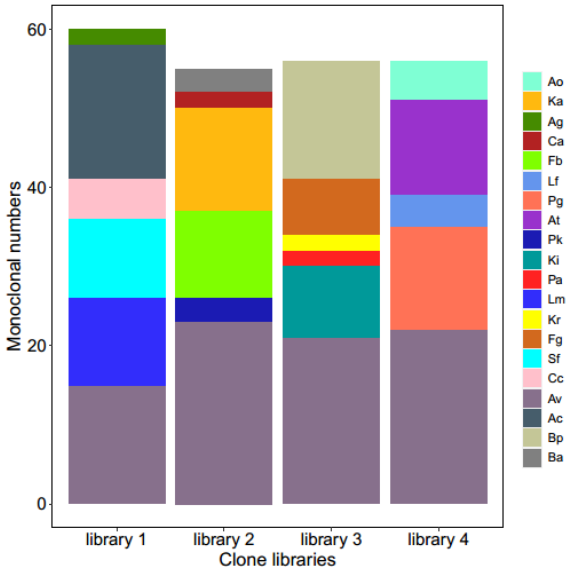

Figure S2 Statistics of monoclonal numbers in the clonal library. After 20 strains were mixed in the same proportion, total DNA was extracted, and then the sequence of the V3V4 region of 16S rRNA was amplified. The clone libraries were constructed using amplified sequences. About 60 single clones were picked from each clone library and sequenced to obtain the actual number of clones of different strains in the clone library. The cloning library experiments preliminarily showed the difference in DNA extraction efficiency of 20 strains.

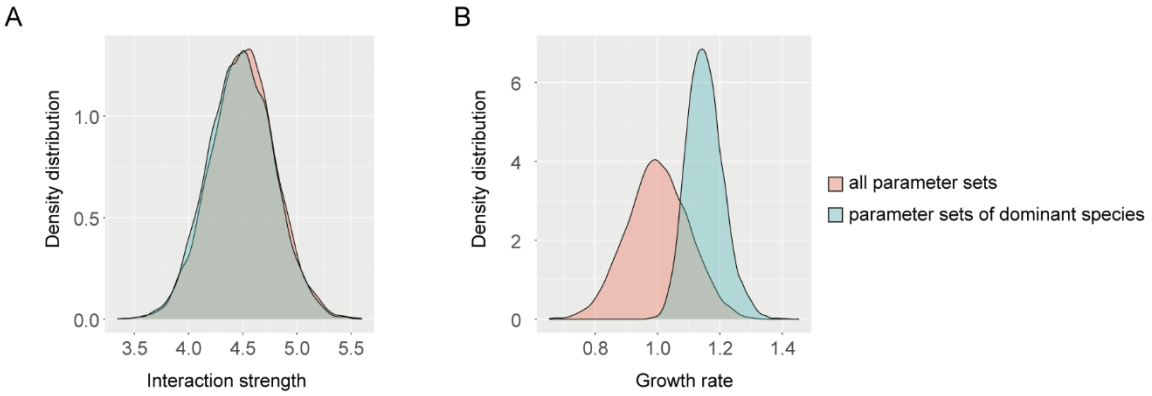

Figure S3 Parameter settings of interaction strength (A) and growth rate (B) in a total of 10 species and dominant species, respectively. Here, the dominant bacteria had no advantage in interaction strength but had a growth rate advantage. The interaction strength of dominant bacteria was obtained by calculating the sum of the interaction strength of other species on advantaged bacteria in a community.

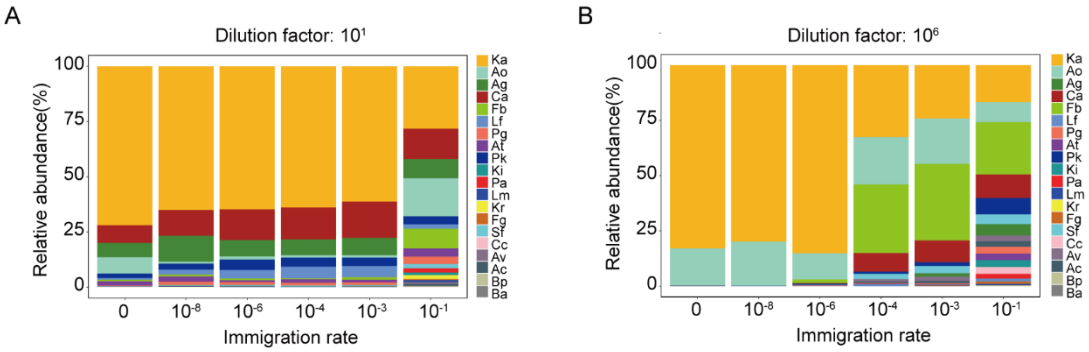

Figure S4 Effects of immigration rates on the diversity of bacterial communities. The community composition with dilution factors of 10^1 (A) and 10^6 (B), respectively.

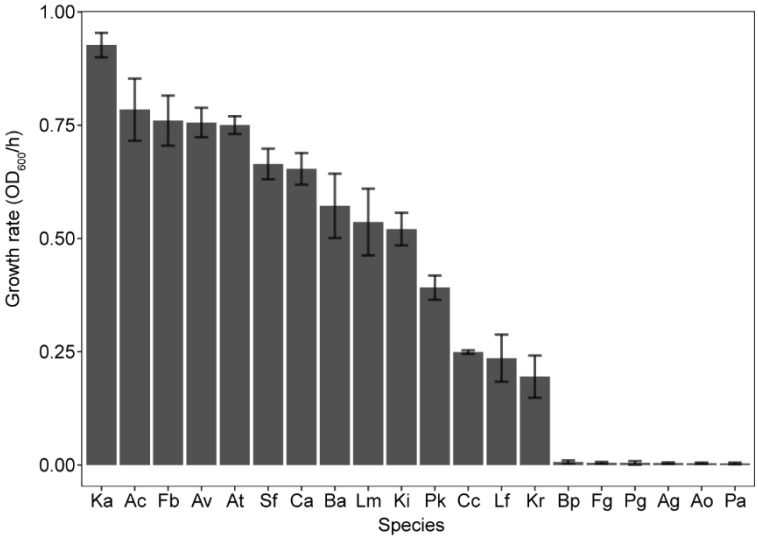

Figure S5 Experimentally measured growth rates of 20 strains.

S3 Supplementary Tables

Table S1 Strain grouping in the determination of DNA extraction efficiency

Group	**Species**	**Strain tested/ *A. tumefaciens* GV3101**
Species with higher extraction efficiency	Ag, Ao, Ba, Cc, Ca, Kr, Lf, Pk, Pa	1:10, 1:5, 1:2, 1:1, 2:1, 5:1
Species with lower extraction efficiency	Bp, Ac, At, Av, Ka, Sf, Fg, Fb, Lm, Pg, Ki	1:1, 2:1, 5:1, 10:1, 15:1, 30:1

Table S2 Correction formulas for DNA extraction efficiencies of 20 species.

No.	Species	Formula	R^2
1	Ag	y=1.73+1.92x	0.73
2	Ao	y=-2.02+2.66x	0.96
3	Ba	y=-4.78+42.5x	0.96
4	Bp	y=-0.3+2.36x	0.99
5	Ac	y=-0.485+1.25x	0.91
6	At	y=-0.621+1.34x	0.88
7	Av	y=-0.446+0.839x	0.96
8	Cc	y=-0.811+63.1x	0.96
9	Ca	y=-1.1+1.96x	0.94
10	Ka	y=-0.48+1.64x	0.93
11	Sf	y=-0.39+1.19x	0.97
12	Fg	y=-0.373+1.51x	0.98
13	Fb	y=-0.666+1.81x	0.95
14	Kr	y=-0.208+48.4x	0.96
15	Lf	y=-1.66+7.04x	0.98
16	Lm	y=-0.479+1.46x	0.95
17	Pk	y=0.429+1.05x	0.96
18	Pg	y=-0.707+1.67x	0.95
19	Pa	y=-2.32+1.5x	0.95
20	Ki	y=-0.315+1.51x	0.94