

Supplementary Materials for Engineering complex communities by directed evolution

Chang-Yu Chang^{*1,2}, Jean C.C. Vila^{*1,2}, Madeline Bender^{1,2}, Richard Li¹, Madeleine C. Mankowski³, Molly Bassette⁴, Julia Borden⁵, Stefan Golfier⁶, Paul Gerald L. Sanchez⁷, Rachel Waymack⁸, Xinwen Zhu⁹, Juan Diaz-Colunga^{1,2}, Sylvie Estrela^{1,2}, Maria Rebolleda-Gomez^{1,2} & Alvaro Sanchez^{1,2,#}

This PDF includes:

Supplementary Text

Supplementary Methods

Supplementary Figures 1-16

Supplementary Tables 1-5

Supplementary Text

Effect of infant population size on Migrant-pool and Propagule Strategies

The migrant-pool and propagule strategies were inspired by earlier group selection experiments, which were carried out with small populations (e.g. 16 individuals at the start of each batch incubation) of sexually reproducing and genetically diverse animals [1–5]. The combination of small population sizes and sexual recombination ensures a sufficiently large between-population variation, on which group-level selection can act. By contrast, microbial populations are largely clonal and they are much larger (e.g. 10^7 - 10^9 cells/mL is commonplace). For instance, in the original experiment by Swenson et al, the inoculum consisted of 0.06-6 g of soil, which should generally contain no fewer than $\sim 10^6$ and up to $\sim 10^{10}$ bacteria (aside from other microorganisms) [6]. The simulations reported in Fig. 1E-F were inoculated with 10^6 cells, and this number is representative of typical population sizes at the beginning of every batch. Given the large population size and clonal reproduction in our communities, we reasoned that the migrant-pool and propagule strategies may be limited in their ability to generate between-community variation in function, and thus will fail to improve community phenotypes even when communities are stable.

Consistent with this idea, we found that pooling the top-performing communities of a stable metacommunity generally increases the mean F , but decreases F_{\max} (Fig. S5A-B). This is partly due to lower-contributing (low ϕ_i) species coming from low function communities outcompeting the high-contributing (high ϕ_i) species in the top community (Fig. S6). Importantly, pooling the top-performing communities dramatically suppresses between-community variation in the offspring generation (Fig. S5B Inset), rapidly exhausting the ability of artificial selection to act [7]. This caveat has been raised before when the migrant pool was applied to animal populations [5], but it is exacerbated here likely due to the large inocula that are common in microbial community selection experiments, and which we have replicated in our simulations. As for propagule propagation (Fig. S5C-D), when applied to stable communities the community-level heritability (which quantifies the degree to which the function of offspring communities resembles their parents' [8]) is very high, approaching $h^2 \sim 1.0$ in most simulations (Fig. S7). This high heritability explains the strong response of the mean F to propagule selection (Fig. S5D-inset). Unfortunately, given the large population sizes and the fact that our species reproduce asexually, high heritability implies that the best community after selection is very similar to the best community in the parent population, both compositionally and functionally [9]. Propagule strategies can thus be efficient at preserving community function, but when they are combined with a high infant population size do not introduce enough variation to improve it much beyond that point (Fig. S8A-D). Consistent with this idea we find that both of these methods can work when the infant population is much smaller, i.e when they are combined with a harsh bottleneck (Fig. S10).

Supplementary Methods

Assumptions in the Microbial Consumer-Resource Model

ecoprospector uses the Microbial Consumer Resource Model (MiCRM) to simulate microbial communities growing in batch culture. The full form of the MiCRM currently implemented in *ecoprospector* is:

$$\frac{dN_i}{dt} = g_i N_i \sum_{\alpha} [(1 - l_{\alpha}) w_{\alpha} \sigma(c_{i\alpha} R_{\alpha}) - m_i] \quad (\text{Eq. S1})$$

$$\frac{dR_{\alpha}}{dt} = \sum_{j\beta} N_j \sigma(c_{i\alpha} R_{\alpha}) \left[D_{\alpha\beta} \frac{w_{\alpha}}{w_{\beta}} l_{\alpha} \right] - \sum_j N_j \sigma(c_{i\alpha} R_{\alpha}) \quad (\text{Eq. S2})$$

All parameters are described and defined in Table S2 (adapted from [10]). For all simulations in the main text and supplement we assume the following: (1) there is no minimal energy requirement to eliminate starvation-induced depletions in population size ($m=0$), (2) all resources have the same energy content ($w_{\alpha} = 1$). The first assumption is justifiable because we are simulating a batch-culture environment where previous experiments have suggested that cell-death is unlikely to have a significant effect on community assembly [23]. The second assumption is a simplification that has also been adopted in previous modelling work and has been found to have little impact on the ability to reproduce the ecological pattern of microbial communities in natural environments [11].

In addition, in the main text we (1) did not allow cross-feeding ($l=0$), (2) assumed a perfect conversion from energy uptake to growth ($g_i = 1$), and (3) assumed a type III functional response ($\sigma(c_{i\alpha} R_{\alpha}) = \sigma_{III}(R_{\alpha})$). In what follows, we explore the effect of relaxing these and other assumptions on our main results (Fig. S12-S16). Specifically we examine whether our main results hold for:

1. Alternative community functions
2. Alternative ecological scenarios that include cross-feeding
3. Alternative types of functional response
4. Alternative methods for seeding the initial metacommunity
5. Alternative distributions of per-capita contribution to function.

These simulations are described below. Unless otherwise specified simulations were conducted using the same protocol and parameters as in the main text. All simulations in this supplement were repeated 20 times to obtain a statistically sound sample size.

1. Alternative community functions

In the main text we assumed that a species' contribution to the community function F is linearly related to its abundance.

$$F = \sum_i \phi_i N_i \quad (\text{Eq. S3})$$

This results in a smooth structure-function landscape (as shown in Fig. 2A). This is an idealized scenario in which there are no functional interactions and contributions to community function are costless. In Fig. S12 we show the results for four alternative community-level functions. We first consider two predefined functions (one which relaxes the assumption of additivity and one which introduces a cost at the species level). We then consider two biologically motivated functions, where the structure-function landscape is not imposed, but rather emerges from the simulation.

- a) **Non-additive community function:** The effect of species on the community function can be interdependent. For instance, the breakdown of a hard-to-digest nutrient requires complementary metabolism of more than two types of strains [12]. We introduce these types of functional interactions by considering a function that depends solely on the abundance of species pairs:

$$F = \sum_i \sum_j \epsilon_{ij} N_i N_j \quad (\text{Eq. S4})$$

Here ϵ is a square matrix in which the off-diagonals are sampled from a normal distribution $\epsilon_{ij} \sim \text{Normal}(\text{mean} = 0, \text{sd} = 1)$ and the diagonals are set to 0 to avoid self interaction. The structure-function landscape for this function will be extremely rugged, with small changes in the abundance of a single taxa having a large effect on F . We repeated all the simulations using this function (Fig. S12). The only different result we obtained compared to the main simulations was that the synthetic community generated in the simulations for Fig. 4 now performs worse than both the directly evolved community and the no-selection control (Fig. S12E). This can easily be understood when we consider that individual species' function in mono-culture are no longer predictive of the function of the whole community. This suggests that when the function of interest depends on complex functional interactions, bottom-up approaches (or rational design) are likely to be ineffective compared to the ecological directed evolution strategies we have proposed.

- b) **Costly function:** Often a function at the community-level may be costly at the species level (for example we may be selecting for the production of some costly public good [13–15]). We explored the effect of a costly function by considering a case where each species can either divert their total input energy to biomass constitution (g_i) or invest in the community function (ϕ_i). In such a scenario there would be strict negative relationship between the per-capita contribution to function ϕ_i and the species growth

efficiency g_i . We impose this relationship by first sampling per-capita contribution from a uniform distribution $Uniform(min = 0, max = 1)$ and then setting $g_i = 1 - \phi_i$ so that both parameters are bounded between [0,1]. $F = \sum_i \phi_i N_i$ as before. Once again, repeating our simulations using this function gives us the same result as in the main text (with the sole difference being that the synthetic community has a much lower function in Fig. S12E). This can be understood, when we consider that the highest functioning members of the synthetic community will be the slowest growers and so would be outcompeted by the lower functioning synthetic community members. This illustrates a limitation of the bottom-up engineering approach, that is circumvented by the ecological approaches we have proposed. By sampling the community space broadly, we are able to select for taxa that are both high functioning and are good competitors in a community context.

- c) **Target resource consumption:** We have so far considered functions that depend directly on the abundance of each taxa within a community. However often we are interested in communities for the indirect effects they have on the abiotic environment, for example we may be interested in a community that produces a specific metabolic by-product [16,17], or a community that efficiently consumes a specific nutrient (for bioremediation purposes [18,19]). To test whether our protocols could be used to select for a specific environments, we repeated our simulations, this time selecting for communities that minimized the abundance of a randomly chosen target resource (R_{target}). Therefore

$$F = -R_{target} \quad (\text{Eq. S5})$$

For the resource-shift perturbation, we excluded the target resource from the list of spiked-in resources (Fig. S12C). As before, repeating our simulations using this function gives us the same result as in the main text (with the sole difference being that the synthetic community has a much lower function in Fig. S12E). This can be understood by considering that top-down community assembly will result in more efficient niche-packing and greater resource consumption than the bottom-up approach where the high functioning species may be competing for the same subset of resources. Consistent with this explanation when we examine the synthetic communities we find that they typically have significantly lower total abundance than the directly evolved communities (624 ± 98 vs 987 ± 12 , Mean \pm SD, $p < 0.01$; paired t-test, $N=20$)

- d) **Invader resistance:** The final function we consider is the community resistance to the invasion of a highly competitive species. This is a complex function that emerges from interactions between all species and resources. A concrete application would be selecting for a rhizosphere microbiome that suppresses the pathogenic strain undermining the growth of host plants [20]. The uptake rates for the invader were generated by drawing the uptake rates from a gamma distribution with the same mean and variance in total uptake rate as was used for the other species (Methods). All c_{ia} were then multiplied by

10. This means that i) the invader can grow on all resources and ii) on average, the total uptake rate of the invader is much larger than the total uptake rate of a typical community member. This design ensures a high invasion fitness of the invader when introduced to the resident community. We completely excluded the pathogen from all initial metacommunities to avoid counting invader cells that are already present in a resident community. During community phenotyping, we introduced a low amount (i.e., 10 cells) of the invader to the communities (which typically have approximately 10^6 cells after dilution) and grew the co-culture for one passage. At the end of co-culture we counted the number of invaders $n_{invader}$.

$$F = -n_{invader} \quad (\text{Eq. S6})$$

For this function we find that all of our proposed protocols are able to successfully generate higher functioning communities (Fig. S12C and S12D). As with the other non-additive functions, the synthetic community does substantially worse than the additive function presented in the main text (Fig. S12E). In contrast to the other functions we have explored, we do find that previously proposed pooling methods do better than the no selection control (Fig. S12B). We hypothesize that this is due to the well-established relationship between biodiversity and invasion resistance, that has been observed in similar models, as well as in experimental communities [21–23]. Consistent with this hypothesis we find that both the species richness and total abundance of the highest functioning communities generated by a typical pooling algorithm (such as [7]) are significantly higher than the highest functioning communities from the no selection control (26±4.6 vs 16±2.5 for species richness, 999±2 vs 981±15 for total abundance, Mean±SD, p<0.01 in both cases; paired t-test, N=20). This suggests that pooling may be an efficient method for directly evolving communities when the function of interest strongly correlates with the richness and/or biomass of the self-assembled communities.

2. Alternative ecological scenarios.

In the main text, we have assumed that microbial consumers compete for a wide range of nutrients that are externally supplied. To confirm that our results do not depend on choosing a pure competition-model we relaxed this assumption by allowing cross-feeding among strains (i.e., facilitation) a feature which has already been incorporated into the microbial consumer resource model [10]. For simplicity we set the degree of metabolite leakage l to 0.5 so that half of the community biomass is generated from the supplied resource and half from metabolic-by-products. The secreted metabolite composition is determined by the stoichiometric matrix D . For simplicity, we sample D using the default approach in the *community-simulator* package (i.e each column in D is sampled from a Dirichlet distribution with concentration parameters $d_{\alpha\beta} = \frac{1}{sM}$). We first consider the original ‘rich media’ described in the main text. We then consider a minimal media environment with only a single resource as in [24]. Both conditions have the same number of resources (M) and total amount of supplied

resources ((i.e., $R_{tot} = 1000$). We show that in Fig. S13, that neither the presence of cross-feeding nor the single-resource environment significantly changed our results. The lack of significance in Fig S13D is due to the relatively lower sample size (20) and disappears when we increase the number of simulations to 100 (as in the main text).

3. Alternative functional responses

The MiCRM assumes that all resource utilization reactions are independent and that the import rate of a particular resource depends on its concentration. *Community simulator* allow for three types of relationship between resource availability and uptake rate (or functional responses):

a linear function (Type-I) where

$$\sigma_I(R_\alpha) = c_{i\alpha} R_\alpha \quad (\text{Eq. S7})$$

a Monad function (Type-II),

$$\sigma_{II}(R_\alpha) = \frac{c_{i\alpha} R_\alpha}{1 + \frac{c_{i\alpha} R_\alpha}{\sigma_{max}}} \quad (\text{Eq. S8})$$

and a Hill function (Type-III)

$$\sigma_{III}(R_\alpha) = \frac{(c_{i\alpha} R_\alpha)^n}{1 + \frac{(c_{i\alpha} R_\alpha)^n}{\sigma_{max}}} \quad (\text{Eq. S9})$$

where n is the Hill coefficient for functional response and σ_{max} is the maximum input flux (mass/time). Throughout the main text and supplements we arbitrarily choose a Type-III functional response using the default parameters in *community-simulator* $\sigma_{max}=1$ and $n = 2$. In Fig. S14, we show that using different functional responses does not qualitatively change our results.

4. Alternative initial metacommunity sampling methods

In total each metacommunity is seeded from a universe of $H=2100$ species. The *ecoprospector* package implements three kinds of sampling methods:

(i) Each community is generated by sampling $n_{inoc} = 10^6$ cells from a different regional pool where the species abundance of all H species in each pool follows a power-law distribution .

$$P_{power}(x) = ax^{-a-1} \quad (\text{Eq. S10})$$

(ii) Each community is generated by sampling $n_{inoc} = 10^6$ cells from a different regional pool where the species abundance of all H species in each pool follows a log-normal distribution

$$P_{log-normal}(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} \quad (\text{Eq. S11})$$

(iii) Each community is generated by sampling a fixed number of species Z out of H and initializing them at uniform abundance such that the total number of cells is $n_{inoc} = 10^6$

For (i) we set $a = 0.01$, for (ii) we set $\mu = 8$ and $\sigma = 8$ and for (iii) we set $Z = 225$. These values were chosen so that all three methods started with a comparable initial species richness, which is also in line with previous empirical data. We show this in Fig. S11 where we plot the rarefaction curves for 11 communities generated using these methods as well as the rarefaction curves for 11 soil communities sequenced in [24]. In addition, we show in Fig. S15, that none of these metacommunity sampling methods qualitatively changed our result.

5. Alternative distribution of per-capita species contribution

In the main text we assumed that a species per-capita contribution to the community function $\phi_i \sim \text{Normal}(\text{mean} = 0, \text{sd} = 1)$. This assumes that i) all species contribute, ii) many species have negative contributions, and iii) rarely are any species strong contributors. We relaxed these three assumptions by considering three alternative distributions from which to sample ϕ_i

1. Most contributions are positive but strong contributors remain rare

$$\phi_i \sim \text{Normal}(\text{mean} = 1, \text{sd} = 1)$$

2. All contributors are positive and many species contribute strongly

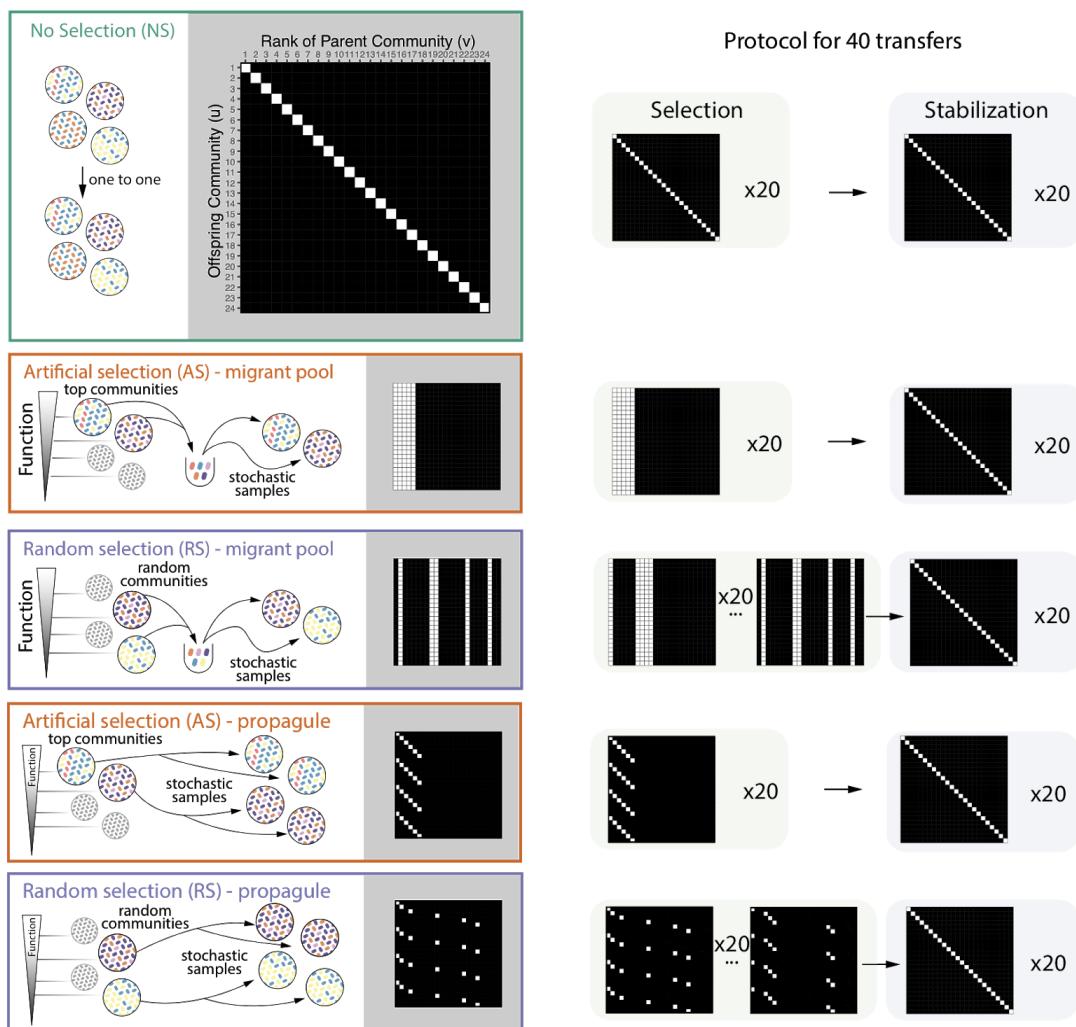
$$\phi_i \sim \text{Uniform}(\text{min} = 0, \text{max} = 1)$$

3. Only 20% species contribute to community function:

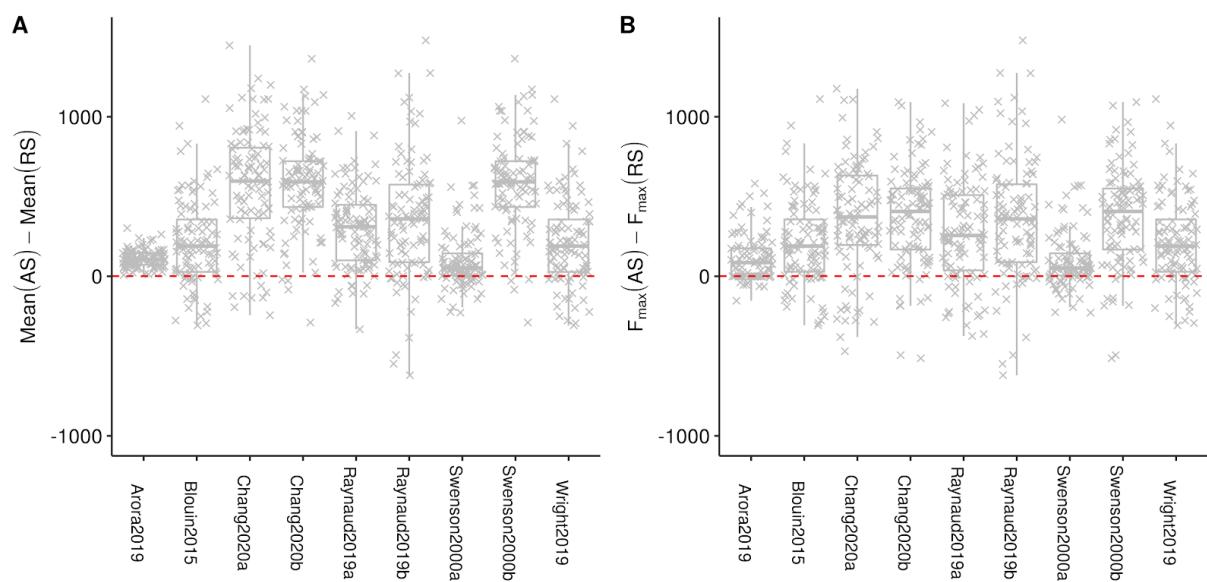
$$\phi_i \sim \text{Normal}(\text{mean} = 0, \text{sd} = 1) \times \text{Bernoulli}(0.2)$$

In Fig. S16, we show that our main results hold for all 4 scenarios we have considered. Interestingly we do find that when many species contribute strongly to community function the synthetic community does much worse than before (1,2), whereas when only a small fraction of species contribute to community function the synthetic community does much better (3). This suggests that when a function is highly idiosyncratic and only a small fraction of species can perform it a bottom-up approach may be more justifiable. In contrast when a function is cosmopolitan and shows broad taxonomic distribution it becomes worthwhile to use directed evolution to select for communities with ecological interactions that increase function.

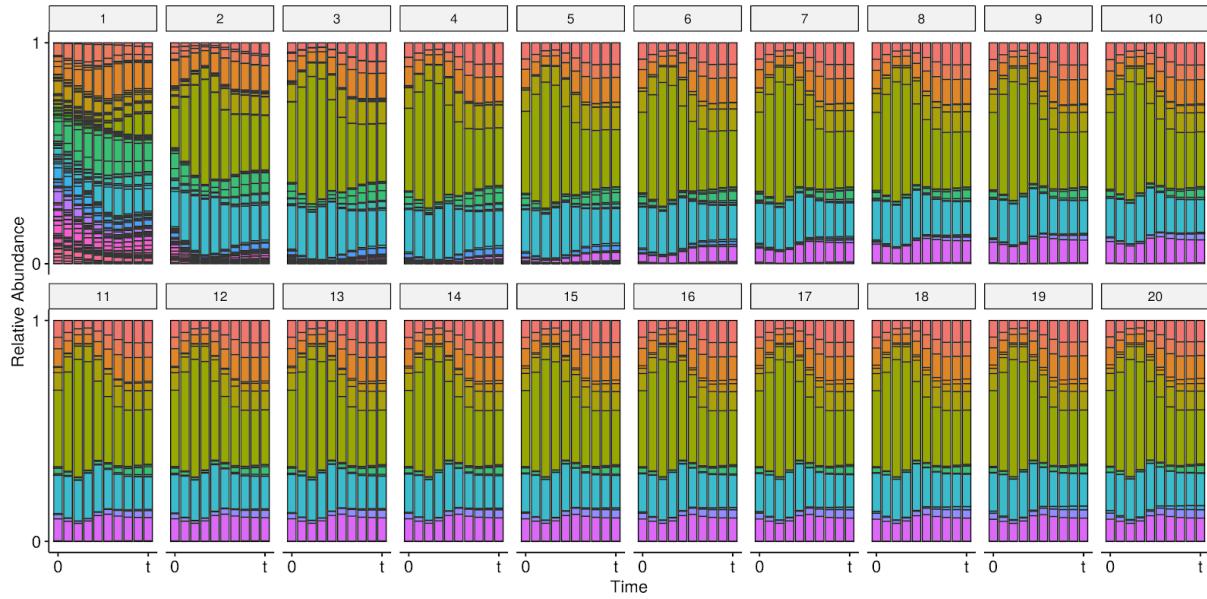
Supplementary Figures



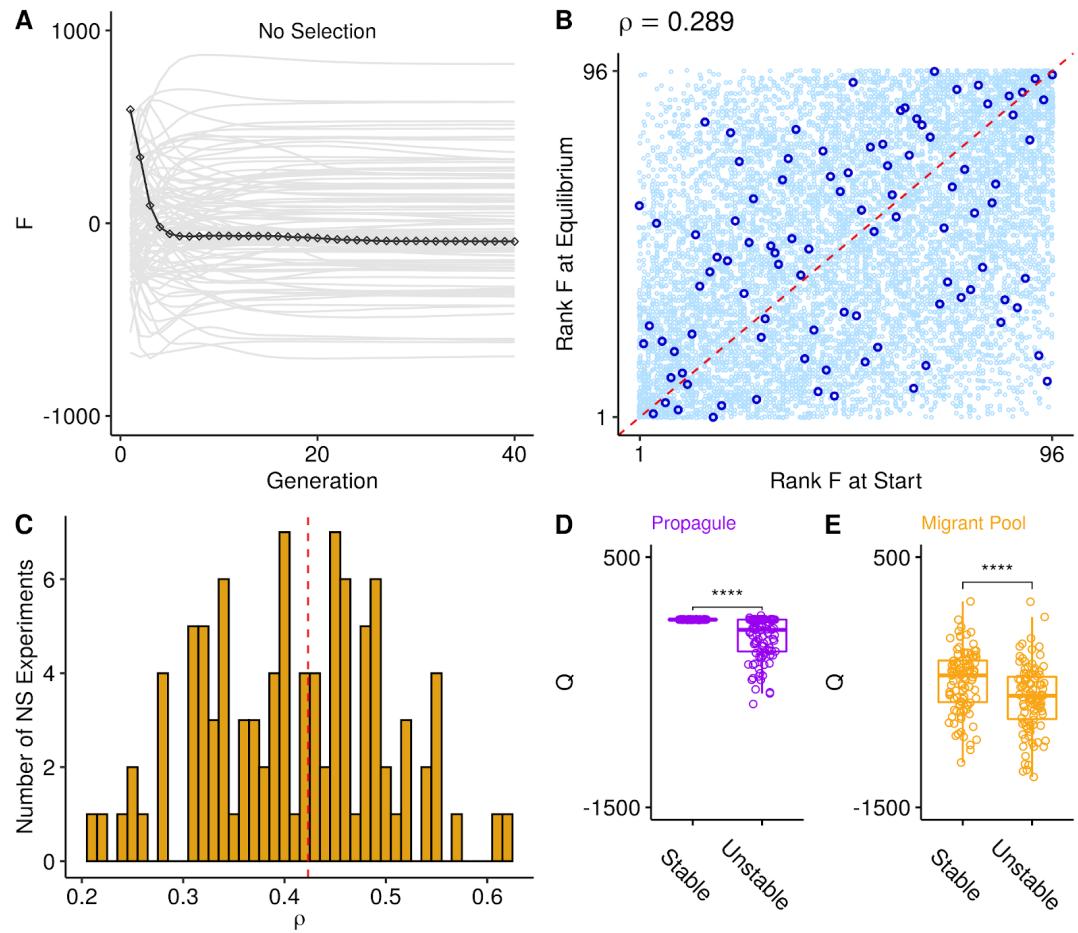
Supplementary Figure 1. Selection matrices. Example for protocols on a metacommunity of 24 communities. A selection scheme can be encoded as a selection matrix S whose element S_{uv} represents the fraction of parent community at rank v that is transferred to offspring at position u . The value of non-zero S_{uv} (white blocks) is set to the dilution factor 10^3 , and the rest values are zero (black blocks) which indicate no transfer between the parent and offspring communities. The parental communities are ordered according to their functions such that the top performing community will locate at the leftmost column in S whereas the community with lowest function will be on the rightmost. An identity matrix (topleft) represents a one-to-one transfer of 24 communities without selection. The two widely used community selection strategies in the microbiome selection studies: propagule and migrant-pool approaches and their random-selection controls can be represented as selection matrices. In these cases, the fraction of parental communities being selected to seed the offspring community is $q = 0.25$ (6 communities).



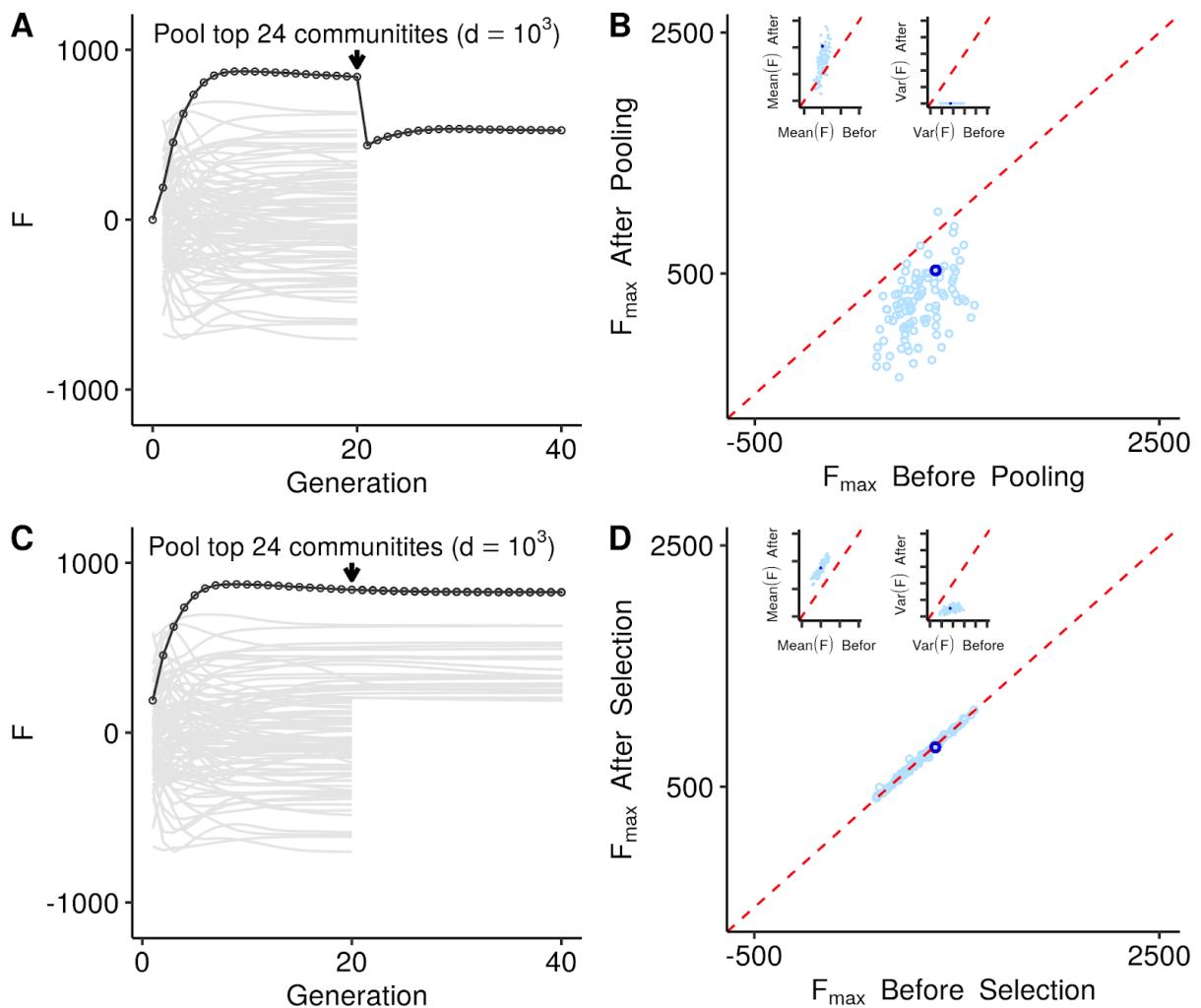
Supplementary Figure 2. Mean and maximum function of artificial selection (AS) line relative to the random selection line (RS). Difference in (A) mean function and (B) F_{\max} between the AS and RS lines. Only experimental protocols that have described RS are shown (Table S1). All differences are statistically significant (Welch's t-test, $P < 0.01$, $N = 100$).



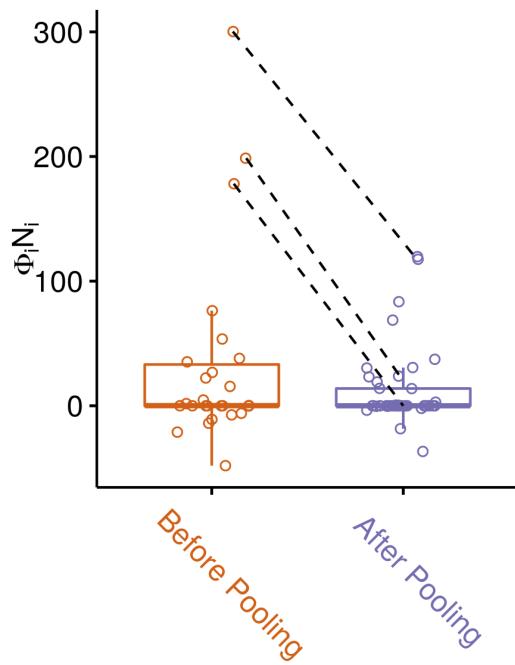
Supplementary Figure 3. Internal dynamics of a community within a generation. To illustrate the concept of generational stability we plot the within-batch community dynamics of a single community in the no-selection line over 20 generations. Each vertical bar represents one of 10 time points within a growth cycle, and colors represent taxa. The initial inoculum has 236 taxa, most of which rapidly go extinct in the first five growth cycles. Despite the temporal dynamics changes within each growth cycle, after ~10 generations community composition converges to a dynamic equilibrium reflected in a repeatable ecological succession in consecutive generations. **25 taxa survive after 20 transfers.**



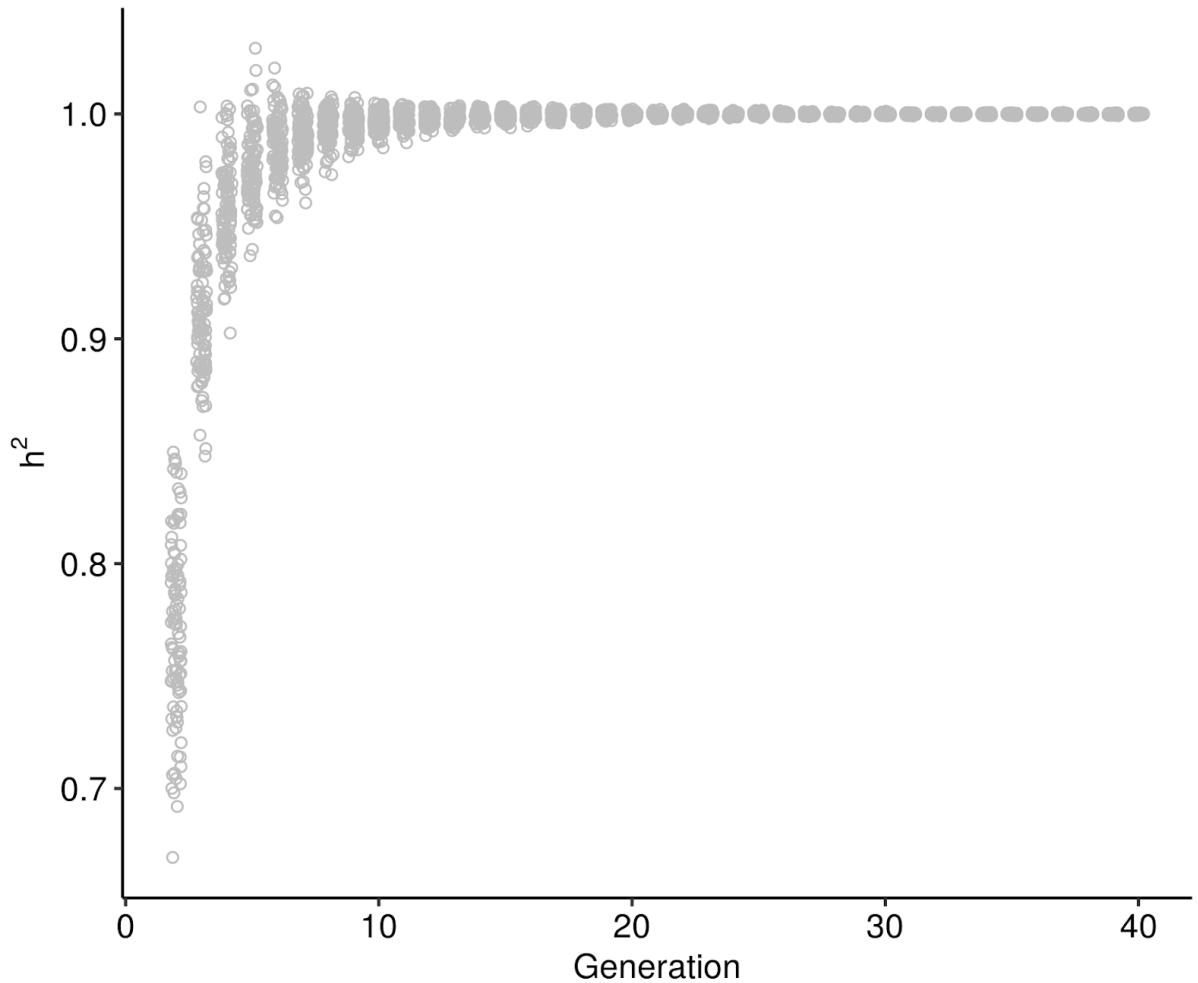
Supplementary Figure 4. Selecting for transient communities is inefficient. (A) Functions of 96 communities grown for 40 generations without selection (NS). The highlighted line shows the community with the highest function at the end of the first generation (B) Community rank function at the start of an experiment is a poor predictor for the rank function at equilibrium. Dark blue points represent 96 communities shown in panel A, and the light blue points denote the communities in the other 99 replicate NS lines. We calculated Spearman's ρ between Rank F at Start and Rank F at equilibrium for this dark blue points (C) Distribution of Spearman's ρ for all 100 NS lines. Dotted red line corresponds to Spearman's ρ for the 96 communities shown in panel A. (D-E) To confirm that selecting for transient communities reduces the effectiveness of both migrant pool and propagule selection methods, we compare an experiment where we apply 20 generations of artificial community level selection to newly inoculated metacommunity with an experiment where we apply 1 single round of artificial community level selection on a metacommunity that has already been stabilized for 19 generations. After this the metacommunity is grown for another 20 generations without selection so that the communities reach equilibrium. In panel (D) we use the propagule method and select 25% of communities after each generation. In panel (E) we use the migrant pool method and also select 25% of communities after each generation. Each of these experiments is repeated 100 times and their effectiveness compared to the NS control is quantified by $Q=F_{\max}[\text{AS}]-F_{\max}[\text{NS}]$. For both the propagule (D) and migrant pool (E) methods we find that a single round of selection on a set of stable communities does better than 20 rounds of selection starting on unstable communities. Brackets represent paired t-tests ($N=100$ for each test). ****: $p<0.0001$.



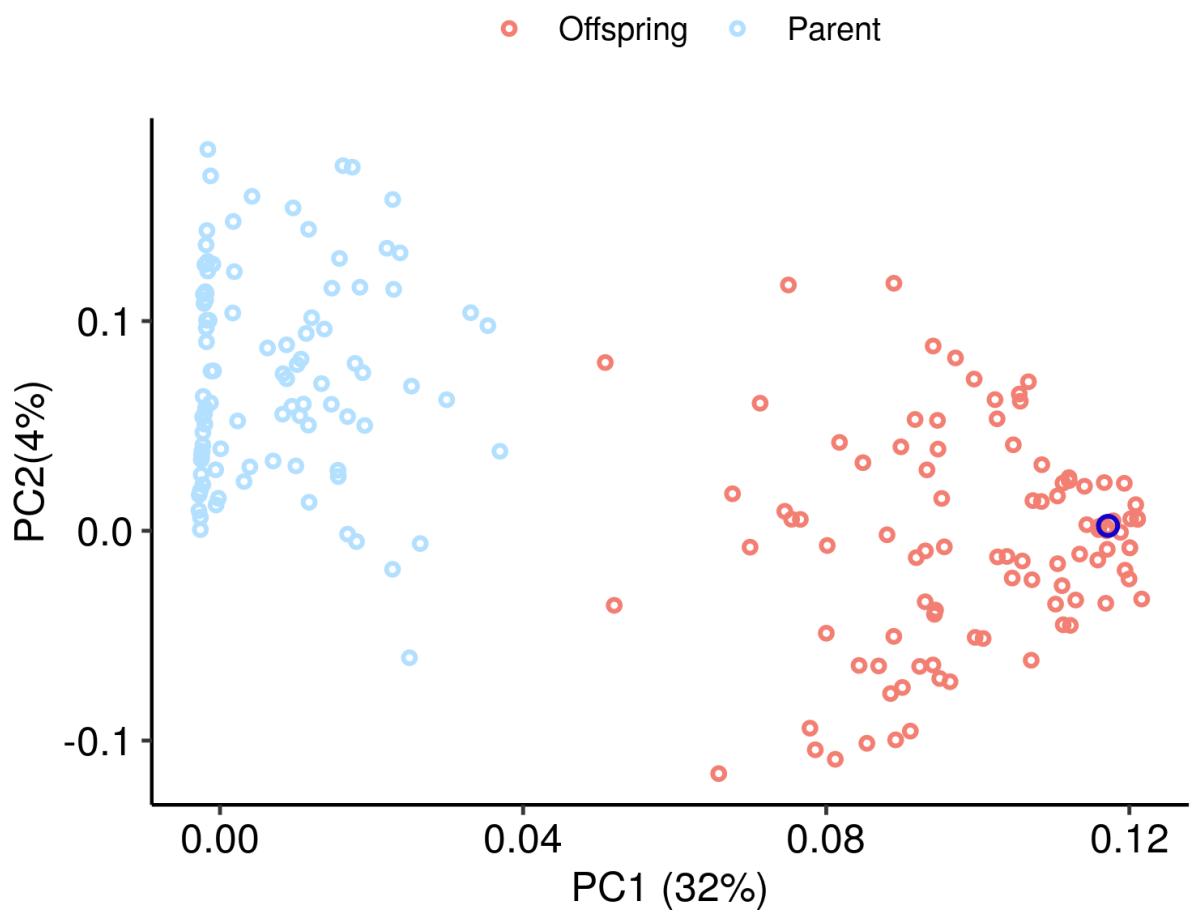
Supplementary Figure 5 Migrant pool and propagule methods fail due to high infant community population size. (A) The community with the highest function at transfer 20 (solid dark line, circular points) may drop when pooled with lower functioning communities (grey lines). (B) Pooling the functionally distinct communities into a single inocula usually results in higher mean function (left inset) but results in lower maximum function and reduced functional variation (right inset). Dark blue point represents the one experiment shown in panel A. (C) Selecting the top 25% of communities at transfer 20 using the propagule method with a modest dilution factor preserves the function of the top community at (solid dark line, circular points). This is due to high heritability of community function when communities are at equilibrium (Fig S7). (B) High heritability means that the propagule strategy consistently results in higher mean function (left inset). However it also means that the propagule strategy is unable to generate new functional variation (right inset) and so we see minimal change in the maximum function before and after selection. Dark blue points represent the one experiment shown in panel (B).



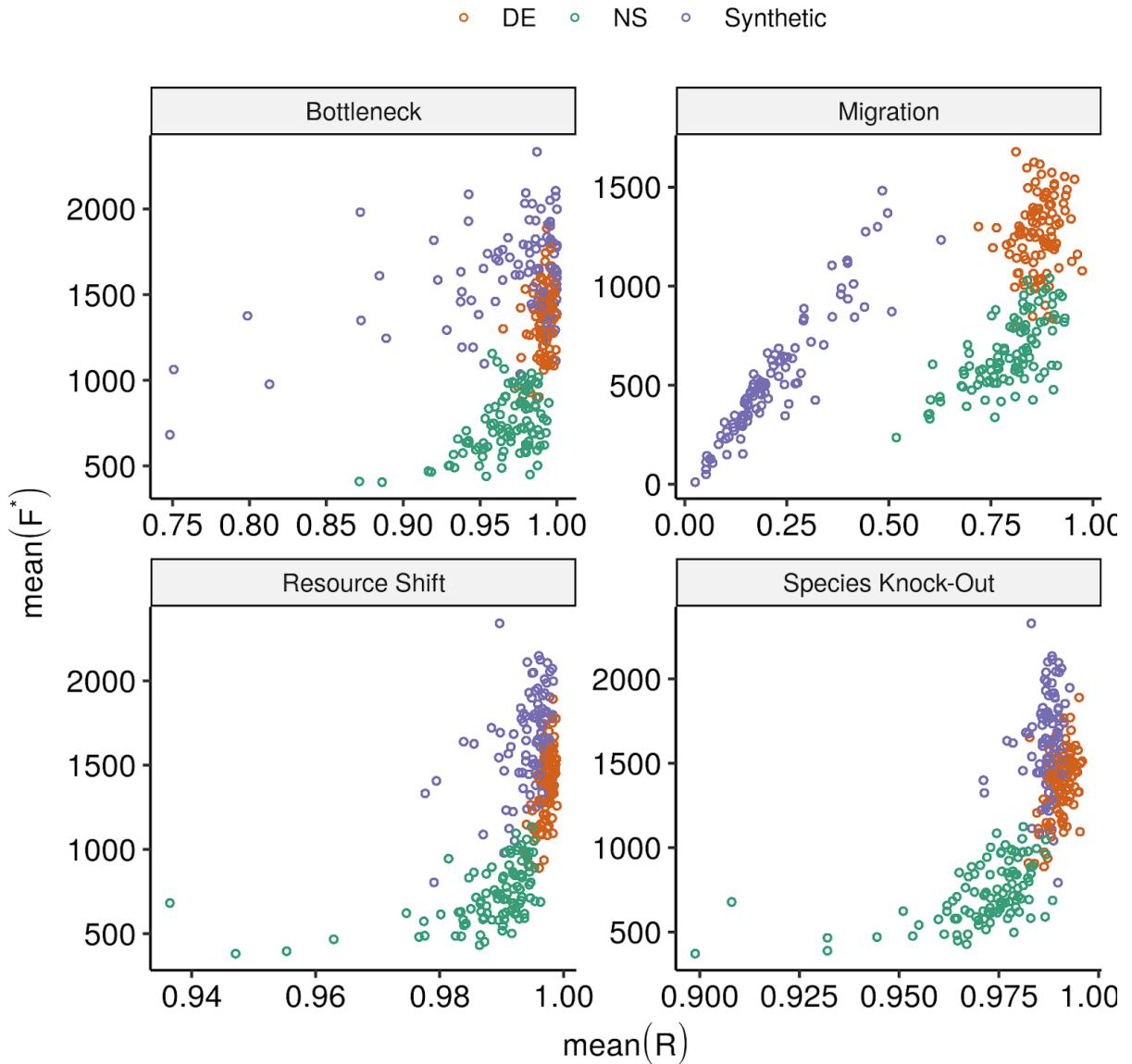
Supplementary Figure 6. Per-species contribution to function before and after pooling. We compare the distribution of per-species contribution to function ($\Phi_i N_i$) for the top one of 96 communities in Figure S5A at the end of generation 20 (before pooling) and generation 40 (after pooling). Each point represents a single species in each community. The drop in community function ($F = \sum_i \Phi_i N_i$) from 841 to 526 shown in Fig S5A is largely due to a substantive drop in abundance of the three highest performing taxa (dashed black lines) as a result of competition with migrants introduced from lower functioning communities .



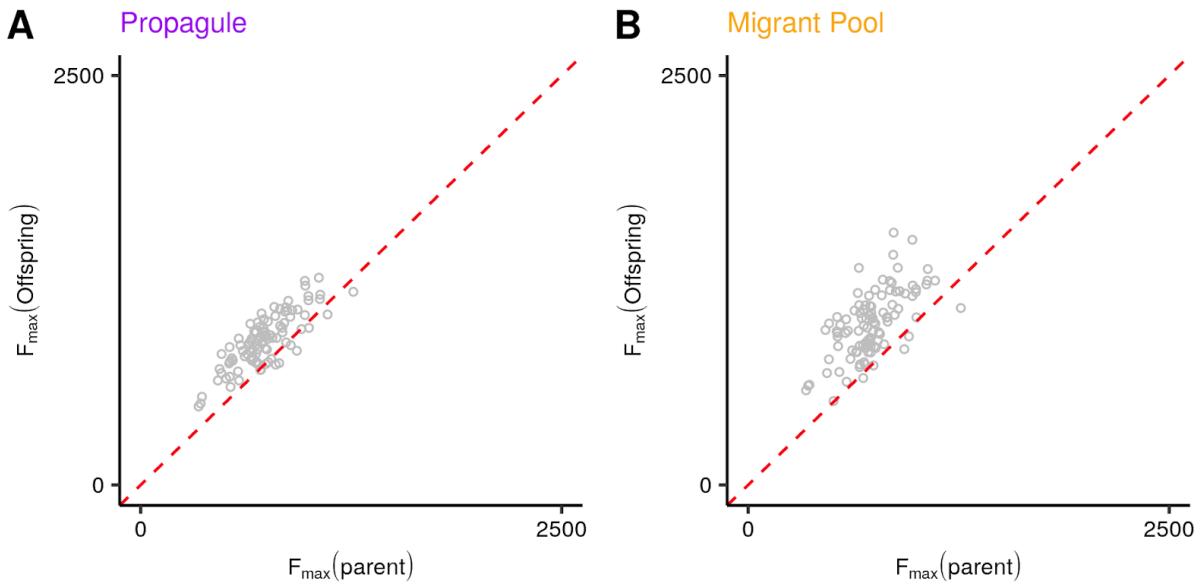
Supplementary Figure 7. Heritability as a function of time in 100 no-selection lines. Each point is the heritability in community function calculated using 96 parental and offspring communities. Heritability is estimated by the slope of linear regression between parental and offspring community function.



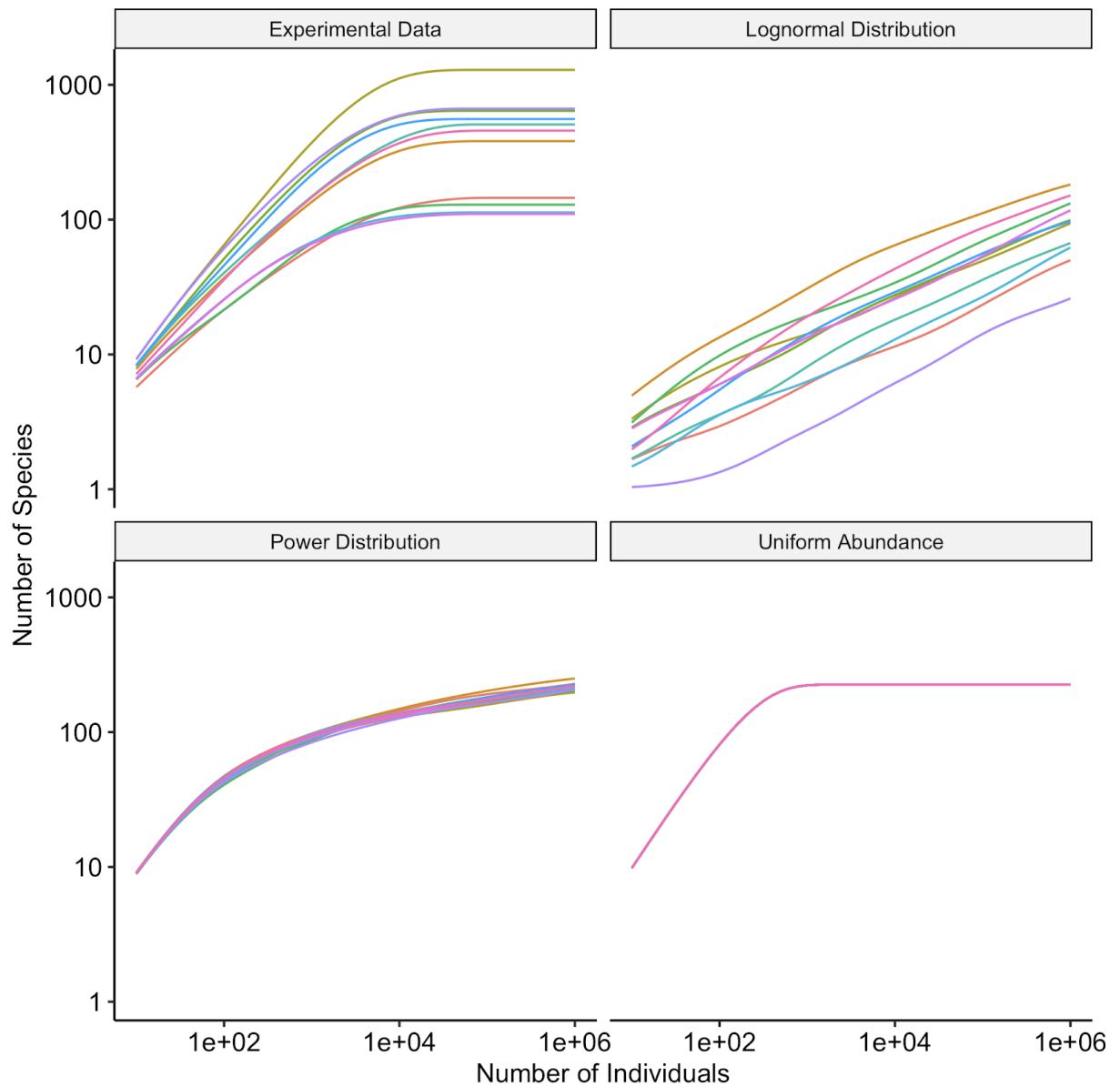
Supplementary Figure 8 Compositional variants generated using dilution shocks. Principal component analysis of the species relative abundance for the communities shown in Fig 2B. Light blue circles correspond to the 96 communities at the end of generation 20 (parent). Red circles correspond to the 96 communities at the generation 40 (offspring). The dark blue circle corresponds to the highest functioning community at generation 20 (i.e the one that is used to seed the offspring generation).



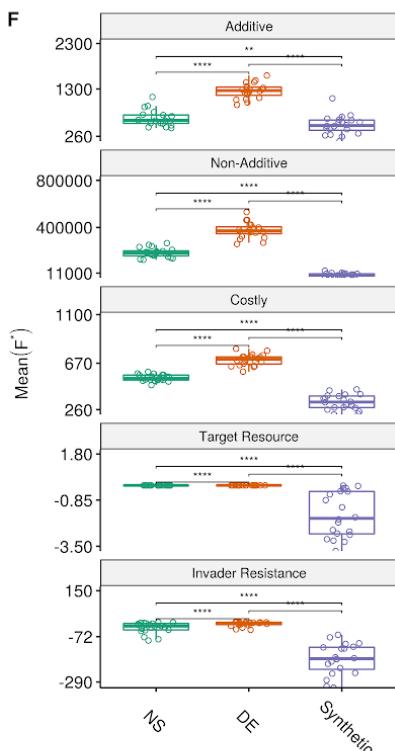
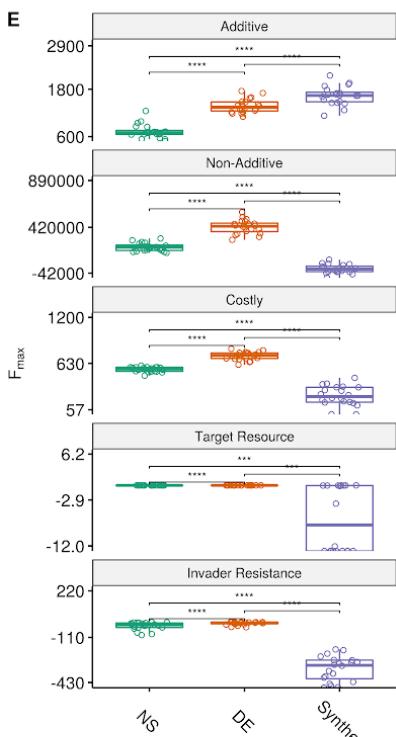
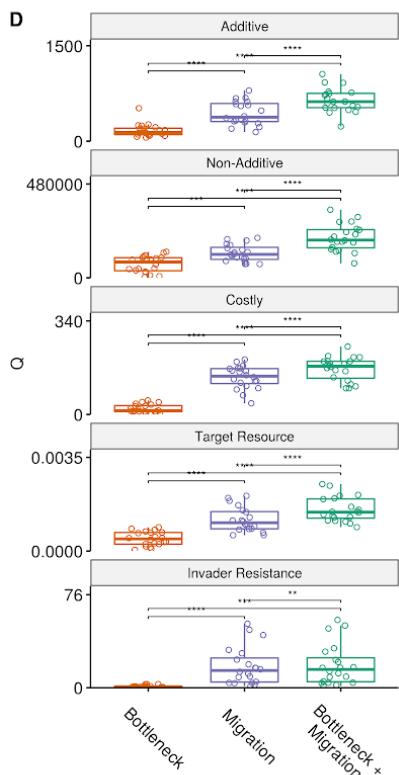
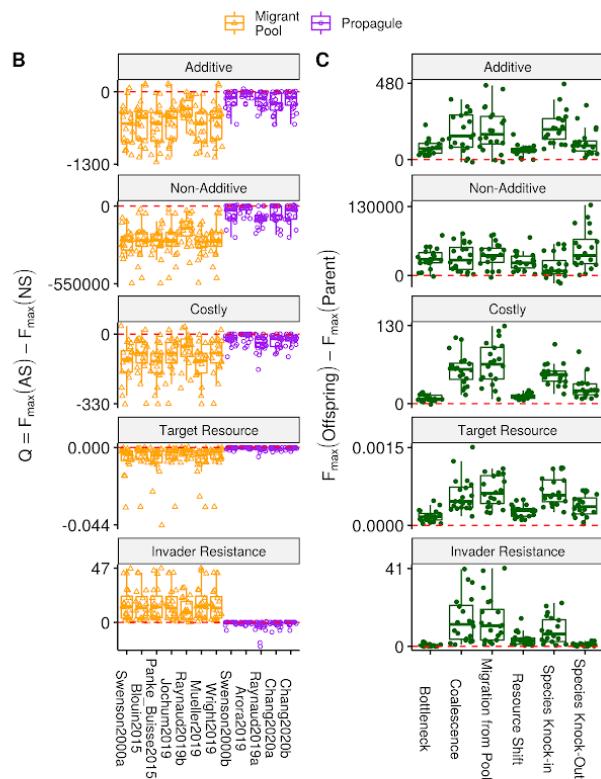
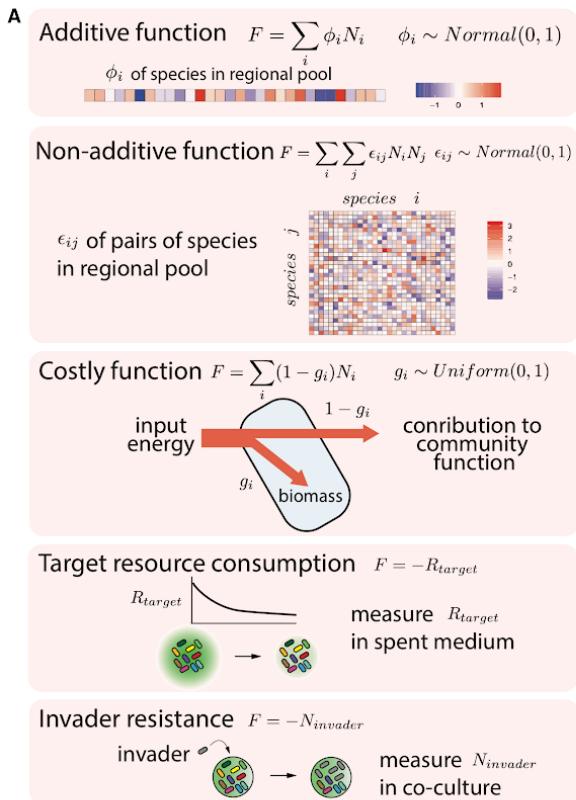
Supplementary Figure 9. Community resistance to various types of ecological perturbation Each subplot shows the Mean(R) vs Mean(F^*) for 100 independent experiments where we subjected the three types of communities described in Figure 4A to 95 replicates of single type of perturbations. The top right panel is the same as Figure 4F where the ecological perturbation examined was migration ($n_{mig}=10^2$). We repeat this experiment for bottlenecks ($d_{bot}=10^4$) (top left), resource shifts ($\delta = 1$) (bottom left), or species knock-outs (bottom right).



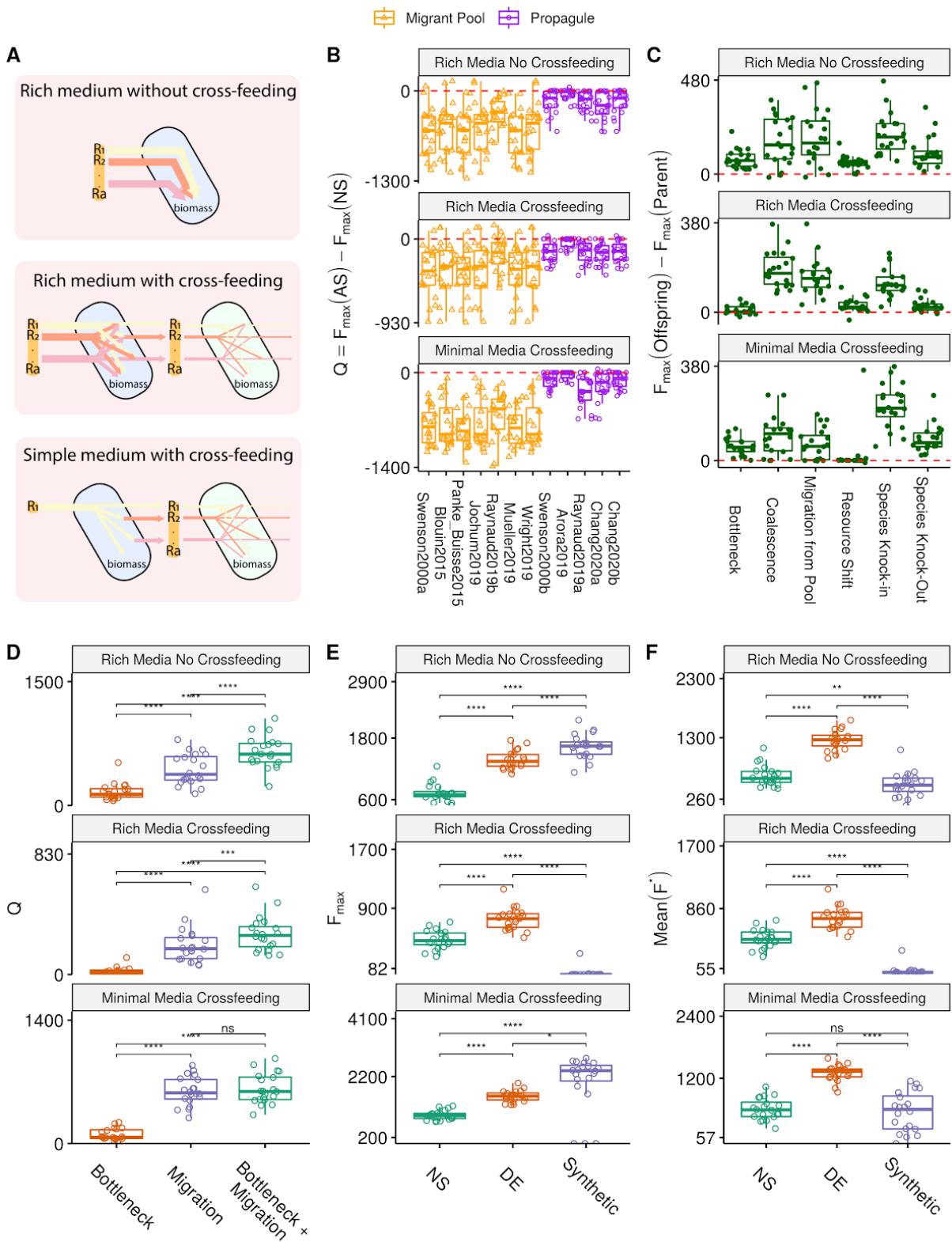
Supplementary Figure 10. Propagule and migrant-pool approaches can improve maximum community function when a harsh bottleneck is applied. In an experiment, a metacommunity of 96 communities is passaged for 20 generations without selection. At generation 20, 24 communities are selected and passaged according to either a propagule selection strategy (A) or a migrant pool selection strategy (B). Immediately after selection a harsh dilution shock is applied to all communities. For (A) we apply a 10^5 bottleneck whereas for B we apply a 2×10^6 bottleneck which means we end up with an average of ~ 10 cells in panel A and ~ 12 cells in panel B.. The communities are subject to another 20 serial transfers.



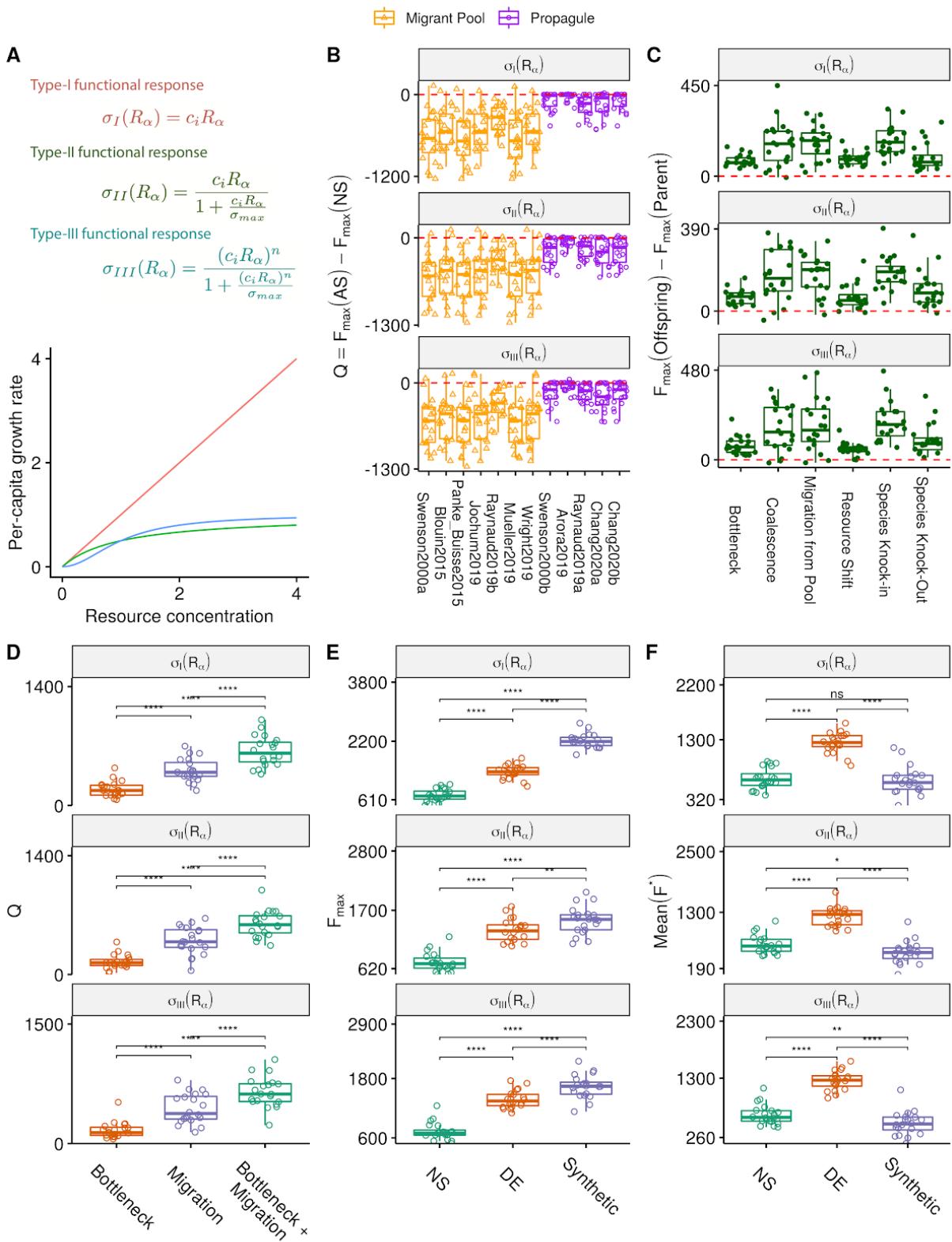
Supplementary Figure 11. Rarefaction curves of metacommunity sampling approaches. An empirical rarefaction curve [24] compared with those generated by the 3 different metacommunity sampling methods in our simulations. In the main text we use the power distribution, whereas the in Fig S15 we show results when using either a lognormal distribution or assuming uniform initial abundances



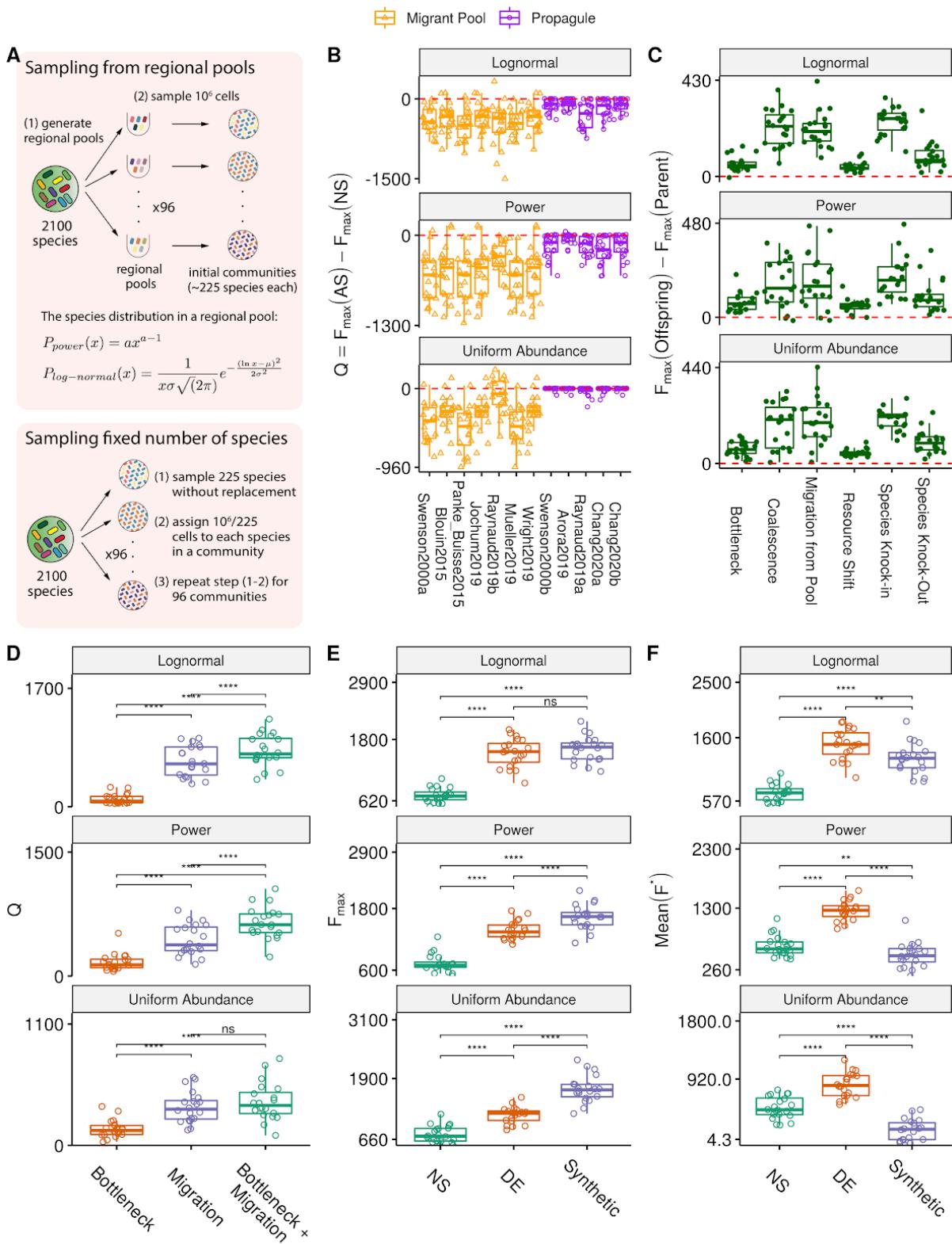
Supplementary Figure 12. Non-additive function, costly function, and two empirically motivated functions. (A) Illustration of the different types of community function we have considered. In addition to the additive function used in the main text we have simulated four other community functions: a non-additive pairwise function, a costly function, a function that maximises the consumption of a target resource, and a function that maximizes resistance to an invader. Panels B-F reproduce the main results reported in Figures 1-4. (B) Difference in F_{\max} between the artificial selection line (AS) and no-selection line (NS) for all previously published protocols, corresponding to Fig. 1F. (C) Difference in F_{\max} between parent (before directed evolution) and offspring (after directed evolution) for the 6 types of perturbation considered in figure 2, this plot aggregates the results shown in Fig. 2D-I. (D) Reproduction of Fig. 3E, to show that iteratively combining migrations and bottlenecks does better than either alone. Q is obtained from each of the three iterative protocols at generation 460 (E) Reproduction of Fig. 4E, where we compare F_{\max} of the no-selection (NS), directed evolution (DE), and synthetic communities; (F) Mean function (F^*) of the DE, NS and Synthetic communities following an ecological perturbation (migration). This corresponds to the y-axis of Fig. 4F.



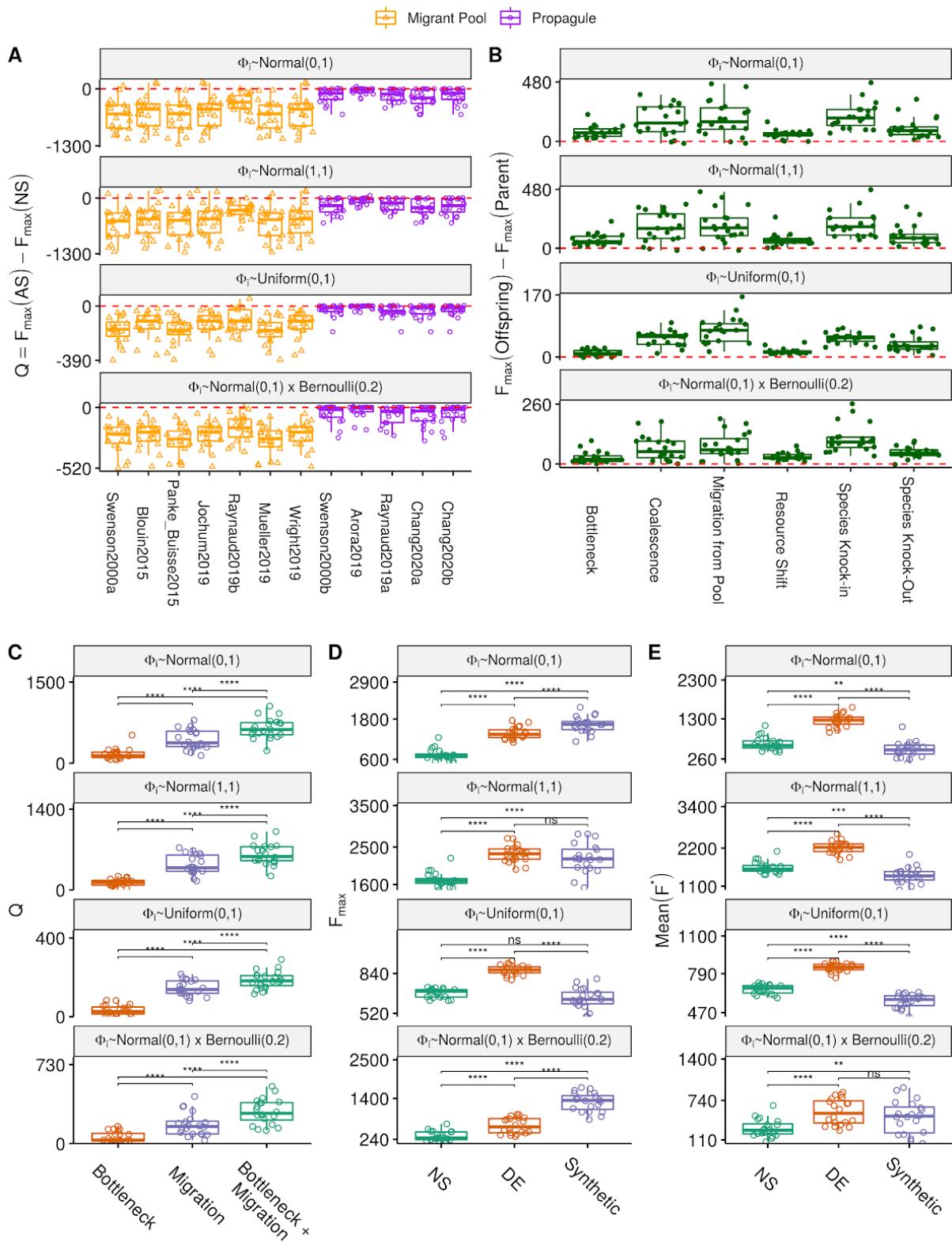
Supplementary Figure 13. Alternative ecological scenarios with metabolic cross-feeding. Besides the rich medium without cross-feeding shown in the main text, we have included two other ecological scenarios: i) rich medium with cross-feeding and ii) simple minimal medium with cross-feeding. The layout of (B-F) follows Fig. S12B-F, reproducing the main results from Fig. 1-4.



Supplementary Figure 14. Functional responses. The resource import rate depends on its concentration in the environments, which can take a linear (type I), Monod (type II), or Hill (type III) form. A Type-III functional response is used in the simulation presented in the main text. The layout of (B-F) follows Fig. S12B-F, reproducing the main results from Fig. 1-4.



Supplementary Figure 15. Alternative Metacommunity sampling approaches. We simulate three metacommunity sampling approaches: i) Each community is seeded with 10^6 cells drawn from a different regional pool, where the species abundances in each regional pool are drawn from a power-law distribution with $a = 0.01$, ii) Each community is seeded with 10^6 cells drawn from a different regional, where the species abundances in each regional pool are drawn from a log-normal distribution with mean $\mu = 8$ and standard deviation $\sigma = 8$, iii) Each community is seeded with a randomly chosen set of 225 species and they are all set to have the same initial abundance. The simulation in the main text adopts the power-law distribution approach. The layout of (B-F) follows Fig. S12B-F, reproducing the main results from Fig. 1-4.



Supplementary Figure 16. Different distributions of per capita species contribution to additive community function. Per capita species contribution drawn from i) normal distribution centered around 0 with standard deviation $sd=1$, ii) normal distribution with $mean=11$ and $sd=1$, iii) uniform distribution ranged from $min=0$ to $max=1$, iiiii) a sparse additive function where 20% of the species contribute to community function. In the main text, per capita species contribution uses normal distribution with $mean=0$ and $sd=1$. The layout of (B-F) follows Fig. S12B-F, reproducing the main results from Fig. 1-4.

Supplementary Table 1. Experimental protocols on artificial community selection. Seven protocols fall into the category of migrant-pool strategy in which a selected set of communities is pooled into a single inoculum to seed the next generation of communities. The other five protocols use the propagule strategy where a selected set of communities are propagated asexually to generate the offspring. Three protocols (Raynaud2019a, Raynaud2019b, and Arora2019) have sublines, which are represented as blocks separated by the red lines. The selection schemes of 12 experimental protocols are converted into selection matrices, which was used to simulate the protocols *in silico* using *ecoprospector*.

Strategy	Protocol	Community source	Targeted function	Random selection control	Number of generations	Number of selection lines	Number of community per lines	Number of communities selected each generation	Percentage of selected communities	Dilution factor	Selection matrix ^a	Random selection matrix ^b
migrant pool	Swenson2000a	plant-associated soil	soil microbiome, plant host biomass	yes	16	2	15	3	0.20	14 and 1417		
	Blouin2015	water treatment plant	lowest CO2 emission	yes	20	6	30	3	0.10	16		
	Panke-Buisse2015	plant-associated soil microbiome	early or late flowering time in the plant host	not available	10	1	14	4	0.28	unclear		
	Jochum2019	soil microbiome associated to drought-persistent grass	delayed onset of drought stress symptom in wheat seedlings establishment	not available	6	1	50	5	0.10	10		
	Raynaud2019b	topsoil	biomass estimated by OD	yes	14	1	30	3	0.10	20		
	Mueller2019	rhizospheres	biomass of plant host with salt-stress tolerance	not available	9	5	8	2	0.25	unclear		
	Wright2019	bulk marine debris	highest chitinase activity	yes	7	1	30	3	0.10	100		
propagule	Swenson2000b	aquatic microbiome from a pond	the highest or the lowest water pH	yes	40	1	24	6	0.25	6		
	Arora2019	fruitfly gut	shortest host eclosion time	yes	4	10	3	1	0.33	unclear		
	Raynaud2019a	topsoil	biomass estimated by OD	yes	14	3	10	1	0.10	20		
	Chang2020a	synthetic communities with four known strains	amylolytic activities	yes	17	1	24	4	0.16	10		
	Chang2020b	soil and leaves	cross-feeding potential	yes	7	1	92	23	0.25	125		

^aFor illustration convenience, the selection matrices shown here are designed for 24 communities rather than 96 that are used otherwise in the main text.

^bIn our simulation, a new random selection matrix for a protocol is drawn every time it needs to transfer from parents to offsprings, so they differ from generation to generation.

^cThe red lines indicate the division of multiple parallel sublines.

Supplementary Table 2. Parameters for Microbial Consumer-Resource Model. Adapted from Marsland2020 Table 1.

Parameter	Description and units	Value
N_i	population density of species i (individuals/volume)	a
R_a	Concentration of resource a (mass/volume)	a
C_{ia}	Uptake rate per unit concentration of resource a by species i (volume/time)	b
$D_{\alpha\beta}$	Fraction of byproducts from resource β converted to α (unitless)	bc
g_i	Conversion factor from energy uptake to growth rate (1/energy)	1
w_a	Energy content of resource a (energy/mass)	1
l_a	Leakage fraction for resource a (unitless)	0
m_i	Minimal energy uptake for maintenance of species i (energy/time)	0
σ_a	Functional response of utilization on resource a	d

^aValues change with consumer-resource dynamics.

^bValues are assigned randomly to each species during simulation setup.

^cThe values in $D_{\alpha\beta}$ do not matter if l_a is 0.

^dDepending on type of functional response chosen.

Supplementary Table 3. Parameters for MiCRM. Most parameters are adapted from Marsland2020 Table 2 except for a , scale, n_{inoc} and α .

Parameter	Description and units	Value
M	Number of resources	90
T	Number of resource classes	1
H	Number of microbial species in global pool	2100
R_{tot}	Total resource abundance	1000
S_f	Number of specialist families	1
u_c	Mean sum over a row of the preference matrix c_{ia}	10
σ_c	Standard deviation of sum over a row of the preference matrix c_{ia}	3
c_0	Low consumption level for Binary c_{ia}	0
c_1	High consumption level for Binary c_{ia}	1
q	Fraction of consumption capacity allocated to preferred resource class	0 ^a
s	Sparsity of metabolic matrix	0.2 ^b
f_w	Fraction of secreted byproducts allocated to waste resource class	0.45 ^a
f_s	Fraction of secreted byproducts allocated to the same resource class	0.45 ^a
a	Exponent parameter in power-law distribution that determines the species abundance in regional pool	0.01
ψ	Number of cells when $N_i = 1$	1e+06
n_{inoc}	Number of cells in the initial inoculum	1e+06
α	Relative functional contribution of species interaction to the additive case	1

^aThese values do not matter if S_f is 1

^bThis value does not matter if I_a is 0

Supplementary Table 4. Protocol-specific parameters. These parameters are used in the protocol used to systematically evaluate the selection matrices from empirical studies (Figure 1E-F; Table S1).

Parameter	Description and units	Value
d	Dilution factor in the batch culture	0.001
t	Incubation time	1
n _{wells}	Number of wells; number of metacommunities	96
T _{tot}	Number of total transfers (generations)	40
T _{selc}	Number of selection transfers (generations)	20

Supplementary Table 5. Parameters for directed selection in Figure 2D-2I.

Parameter	Description and units	Value
θ	The percentile determining the high-performing species in the species pool used to knock in	0.95
d_{bot}	Bottleneck size	1e+05
n_{mig}	Number of cells in the migrant community	1e+06
f_{coa}	Mixing ratio of coalescence; biomass of immigrant community relative to that of a perturbed community copy	0.5
δ	Tunes the magnitude of resource perturbation. The fraction from depleting a resource and move the same amount to another	1

References

1. Wade MJ. Group selections among laboratory populations of *Tribolium*. *Proc Natl Acad Sci U S A.* 1976;73: 4604–4607.
2. Wade MJ. An experimental study of group selection. *Evolution.* 1977;31: 134–153.
3. Wade MJ. A Critical Review of the Models of Group Selection. *Q Rev Biol.* 1978;53: 101–114.
4. Goodnight CJ. Experimental Studies of Community Evolution I: The Response to Selection at the Community Level. *Evolution.* 1990;44: 1614–1624.
5. Goodnight CJ. Evolution in metacommunities. *Philos Trans R Soc Lond B Biol Sci.* 2011;366: 1401–1409.
6. Torsvik V, Øvreås L. Microbial diversity and function in soil: from genes to ecosystems. *Curr Opin Microbiol.* 2002;5: 240–245.
7. Blouin M, Karimi B, Mathieu J, Lerch TZ. Levels and limits in artificial selection of communities. *Ecol Lett.* 2015;18: 1040–1048.
8. Goodnight CJ. Heritability at the ecosystem level. *Proceedings of the National Academy of Sciences of the United States of America.* 2000. pp. 9365–9366.
9. Chang C-Y, Osborne ML, Bajic D, Sanchez A. Artificially selecting microbial communities using propagule strategies. *bioRxiv.* 2020. p. 2020.05.01.066282. doi:10.1101/2020.05.01.066282
10. Marsland R, Cui W, Goldford J, Mehta P. The Community Simulator: A Python package for microbial ecology. *PLoS One.* 2020;15: e0230430.
11. Marsland R 3rd, Cui W, Mehta P. A minimal model for microbial biodiversity can reproduce experimentally observed ecological patterns. *Sci Rep.* 2020;10: 3308.
12. Enke TN, Datta MS, Schwartzman J, Cermak N, Schmitz D, Barrere J, et al. Modular Assembly of Polysaccharide-Degrading Marine Microbial Communities. *Curr Biol.* 2019;29: 1528–1535.e6.
13. Cordero OX, Ventouras L-A, DeLong EF, Polz MF. Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations. *Proc Natl Acad Sci U S A.* 2012;109: 20059–20064.
14. Griffin AS, West SA, Buckling A. Cooperation and competition in pathogenic bacteria. *Nature.* 2004;430: 1024–1027.
15. West SA, Buckling A. Cooperation, virulence and siderophore production in bacterial parasites. *Proc Biol Sci.* 2003;270: 37–44.
16. Sierociński P, Milferstedt K, Bayer F, Großkopf T, Alston M, Bastkowski S, et al. A Single Community Dominates Structure and Function of a Mixture of Multiple Methanogenic

- Communities. Curr Biol. 2017;27: 3390–3395.e4.
- 17. Embree M, Liu JK, Al-Bassam MM, Zengler K. Networks of energetic and metabolic interactions define dynamics in microbial communities. Proceedings of the National Academy of Sciences. 2015. doi:10.1073/pnas.1506034112
 - 18. Piccardi P, Vessman B, Mitri S. Toxicity drives facilitation between 4 bacterial species. Proceedings of the National Academy of Sciences. 2019;116: 15979–15984.
 - 19. van der Gast CJ, Knowles CJ, Starkey M, Thompson IP. Selection of microbial consortia for treating metal-working fluids. J Ind Microbiol Biotechnol. 2002;29: 20–27.
 - 20. Yin C, Vargas JMC, Schlatter DC, Hagerty CH, Hulbert SH, Paulitz T. Rhizosphere Community Selection Reveals Bacteria Associated With Reduced Root Disease. In Review. 2020. doi:10.21203/rs.3.rs-64051/v1
 - 21. Tilman D. Resource Competition and Community Structure. Princeton University Press; 1982.
 - 22. Acosta F, Zamor RM, Najar FZ, Roe BA, Hambright KD. Dynamics of an experimental microbial invasion. Proc Natl Acad Sci U S A. 2015;112: 11594–11599.
 - 23. Tilman D. Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. Proc Natl Acad Sci U S A. 2004. doi:10.1073/PNAS.87.24.9610
 - 24. Goldford JE, Lu N, Bajić D, Estrela S, Tikhonov M, Sanchez-Gorostiaga A, et al. Emergent simplicity in microbial community assembly. Science. 2018;361: 469–474.