Supplementary Materials for Engineering complex communities by directed evolution

Chang-Yu Chang*1,2, Jean C.C. Vila*1,2, Madeline Bender1,2, Richard Li1, Madeleine C. Mankowski3, Molly Bassette4, Julia Borden5, Stefan Golfier6, Paul Gerald L. Sanchez7, Rachel Waymack8, Xinwen Zhu9, Juan Diaz-Colunga1,2, Sylvie Estrela1,2, Maria Rebolleda-Gomez1,2 & Alvaro Sanchez1,2,*

This PDF includes:
- Supplementary Discussion
- Supplementary Methods
- Supplementary Figures 1-12
- Supplementary Tables 1-5
SUPPLEMENTARY TEXT

I. SUPPLEMENTARY DISCUSSION

Extended discussion. A growing number of techniques are making it possible to edit the genomes of microbes within microbial communities 1-3. Simple ecological methods to modify microbiome composition also exist, from dilution-to-extinction 4,5, to species engraftment 6. Together, these tools are paving the way to extend directed evolution from its usual sub-organismal domain to the microbiome level. However, working with large, diverse communities of asexual microbes as the unit of (artificial) selection presents unique challenges and opportunities. For instance, the population dynamics within a single batch represent an ecological succession. Recently, this succession has been elegantly conceptualized as a kind of “developmental maturation”, where a community is in an “infant” state at the time of inoculation, and it is an “adult” at the time of harvest and reproduction 7,8. When death rates are high, it is possible for communities to reach a steady state within a single succession, thus reaching generational stability by the time they are an adult 7. When death rates are low, as is the case in our simulations (and our enrichment experiments 9), then multiple serial passages are needed for communities to reach generational stability. This means that adult communities in early generations can still be moving in their ecological structure-function landscape, and should not be subject to selection in that state. At least one artificial selection study found evidence of an unstable succession between transfers in the early selection rounds 10, and this was consistent with recent reports of functional stabilization taking >6 community-generations in sequential enrichment communities 9,11-13.

When we examined the effectiveness of the two main methods of artificial group selection that have been used in the past, the migrant-pool and the propagule methods, we found that both underperform when applied to large microbial ecosystems that are not yet generationally stable (Supplementary Fig. 4D-E). We note, however, that both strategies worked much better when a strong bottleneck is applied and “infant” population sizes are comparable to those used in animal studies (Supplementary Fig. 12). Based on our simulations, we suggest that (i) ensuring that communities are generationally stable before selection is applied, and (ii) systematically exploring the effect of bottleneck size on between-community variation, will both enhance the effectiveness of both strategies.

Limitations of our study. Our study has important limitations, as the space of all possible ecological scenarios and methods of artificial community-selection is enormous and we have barely scratched its surface. For instance, we have limited ourselves to rank-based selection, as opposed to assigning communities a reproductive success based on their function, i.e. a “fitness”. Our simulations lack a host, and focus on community-level functions, whereas many of the microbiome selection experiments to date have applied artificial selection based on host traits, rather than direct community-level phenotypes 14-17. Additional work will be needed to extend the ecoprospector package to indirect selection on host traits 15. In our simulations, species interact exclusively via competition for substitutable resources. This excludes the important
case of inhibitory and other chemically mediated interactions\textsuperscript{18}. The MiCRM can accommodate these interactions organically, and it should be straightforward to implement them in future iterations of this work.

We have also not exhaustively considered several practical aspects of artificial community-level selection, which have been studied recently and are known to be important for the success of this approach \textsuperscript{8,19}. Perhaps the most relevant of these would be the role of non-heritable variation, which may arise from various sources: from pipetting during transfers to day-to-day environmental fluctuations. These non-heritable sources of between-community variation will all reduce heritability, working against ecosystem-level selection \textsuperscript{19}. Along the same lines, our analysis and its discussion has centered in an ecological regime where within-community population dynamics is dominated by selection, with little ecological drift. Since many microbiome selection experiments take place in open ecosystems \textsuperscript{14,15,20}, it is very likely that one would find species entering and leaving the communities, transiently invading without fixing \textsuperscript{21–23}. How these transient species should affect ecosystem-level heritability remains poorly understood.

Our models contain a single dynamical attractor. Theory has suggested that complex communities can sometimes display more complex dynamics, such as having multiple equilibria or converging to a non-equilibrium state with cyclical or chaotic dynamics \textsuperscript{24–26}. If communities can converge to multiple equilibria, we expect this to make it easier to navigate a structure-function landscape as this would increase the number of accessible stable points. In contrast chaotic or cyclical dynamics may reduce ecosystem-level heritability rendering any attempt to engineer a stable function futile. What types of dynamics are exhibited by complex communities is an outstanding empirical question that is beyond the scope of this study, though we note that under closed conditions enrichment communities do seem to converge to one or more (generationally) stable equilibria \textsuperscript{9,13}. We also have not exhaustively explored the various methods of constructing synthetic communities. While in our simulation we limited ourselves to simply assembling those taxa with high $\phi_r$ one could also attempt to engineer stability, for example by choosing taxa with non-overlapping uptake rates, as strains with lower niche overlap are more likely to coexist \textsuperscript{27}. Although the above (and many other) factors have not been considered here, and therefore we cannot claim that our qualitative findings are general for all functions and ecological conditions, the ecoprospector package we have developed is flexible and we are confident that it can accommodate all of those additional scenarios and many others.

Previous work has often framed the challenges of community level selection as arising from a conflict between the competitiveness of individuals or species and the function of interest at the community level \textsuperscript{8,19}. The ecological approach that we have outlined here circumvents this conflict by relying on standing genetic variation and a diversity of ecological strategies. For this reason even incorporating a strict physiological trade-off between growth efficiency and per capita contribution to function does not qualitatively change our results (Supplementary Materials). For instance, we find that species may offset the diminished growth-yield per resource molecule (which is the result of the cost we impose) when they also have a higher
resource uptake rate. For a diverse enough pool of species, it is not inconceivable that one may encounter a diversity of metabolic strategies where the costs will not be equally affecting all.

**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. 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). 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 $F_{\text{max}}$ (Supplementary Fig. 5A-B). This is partly due to lower-contributing (low $\phi_i$) species coming from low function communities outcompeting the high-contributing (high $\phi_i$) species in the top community (Supplementary Fig. 6). Importantly, pooling the top-performing communities dramatically suppresses between-community variation in the offspring generation (Supplementary Fig. 5B Inset), rapidly exhausting the ability of artificial selection to act. This caveat has been raised before when the migrant pool was applied to animal populations, 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 (Supplementary Fig. 5C-D), when applied to stable communities the community-level heritability (which quantifies the degree to which the function of offspring communities resembles their parents) is very high, approaching $h^2\sim 1.0$ in most simulations (Supplementary Fig. 7). This high heritability explains the strong response of the mean $F$ to propagule selection (Supplementary Fig. 5D-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. 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 (Supplementary Fig. 8). 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 (Supplementary Fig. 12).
II. SUPPLEMENTARY METHODS

Comparing the effectiveness of all previously used artificial community-level selection protocols using ecoprospector. To systematically evaluate all of the previously used experimental protocols for artificial microbiome selection, we have adapted them into a standardized format that can be simulated using ecoprospector. These protocols, listed in Supplementary Table 1, were originally designed for microbial communities that had assembled in environments as diverse as the rhizosphere, animal guts, or water treatment plants. Thus, they vary considerably in design. Differences include, the phenotype under selection, the incubation time, dilution factor, number of artificial selection lines, the number of communities per line, and the number of community generations and the controls that were carried out (Supplementary Table 1).

To standardize the size of all artificial selection lines we consider a metacommunity of 96 communities because: i) this is close to the largest number of communities previously considered in a single artificial selection experiment \(^{36}\), and ii) 96 well plates are widely used in high-throughput microbial ecology and evolution experiments. To assess these protocols independently of system-specific details, we consider only the selection method used, i.e. at the end of each community generation what fraction of communities are selected and how are they transferred, as well as which generation selection starts. This means that our adaptation of each protocol differs solely in the specific \(96 \times 96\) \(S\) matrix. We illustrate the \(S\) matrices used for all protocols in Supplementary Table 1.

The microbiome selection experiments we have examined used either a propagule strategy or migrant-pool strategy. For three of five studies that employ propagule strategies (Swenson2000b\(^{37}\), Chang 2020a \(^{36}\), Chang 2020b \(^{36}\)) this simply involved selecting a fraction \(\varrho\) of communities from the parent generation, and seeding \(1/\varrho\) offspring communities from each of these selected communities. In the \(S\) matrix this is encoded by each of the top \(\varrho \times 96\) parent communities being transferred to (different) \(1/\varrho\) offspring communities. In cases where \(\varrho \times 96\) is not a whole number, it is rounded up to the nearest integer and all lower function offspring communities are discarded (to ensure metacommunity size remains at 96). For example, if \(\varrho\) is 0.1, 10 communities will be selected. The 6 top communities will seed 10 offspring communities and the remaining 4 will seed 9 offspring communities. For six of the seven studies that employ migrant-pool strategies this involves selecting a fraction \(\varrho\) of communities from the parent generation, mixing them together and seeding all communities in the offspring generation from this pool. In the \(S\) matrix this is encoded by each of the top \(\varrho \times 96\) communities being transferred to every offspring community. In cases where \(\varrho \times 96\) is not a whole number, it is rounded up to the nearest integer. Within each strategy, the various experiment protocols differed in the fraction \(\varrho\) (ranging from 0.1-0.33).

In two studies, both strategies are carried out with a slight modification in how communities are used to seed the next generation. Specifically Raynaud et al and Arora et al performed artificial
community level selection using multiple sub-lines rather than one single line \cite{17,38}. In Raynaud et al, both of the experiments had three sub-lines within an artificial selection line. In the experiment that used a propagule strategy, the top community from each sub-line of the 10 communities is selected and used to seed a sub-line of the next generation. In the experiment using the migrant-pool strategy, the top community from each of the three sub-lines are mixed into a pool that is used to seed all new communities. This is reflected in the selection matrix dividing the 96 communities into three sub-lines of 32 communities each (with only the top member of each sub-line being selected). Arora et al used a similar propagule strategy \cite{17}, and also used multiple parallel sub-lines this time with each containing three communities. This selection scheme is adapted in our simulation by grouping the communities into 32 sub-lines of three communities.

A number of studies implemented a control strategy involving random selection at each generation \cite{10,14,16,17,34,36,38}. For these studies, we also tested the random selection controls by randomizing the rank of all communities, or of communities within a sub-line where applicable. We illustrate one example of a random control S matrix for each protocol in Supplementary Table 1.

Our simulations are seeded with an initial inoculum size \( n_{\text{inoc}} = 10^6 \) cells. The initial inoculum size was chosen to be comparable with the pioneering work on microbiome selection which includes treatments with two inoculum sizes 0.06g and 6g of soil, which may contain anywhere between \( 10^6 - 10^{10} \) cells \cite{14}. This initial inoculum size gives us communities at the start of our simulations with 225±12 (Mean±SD) species, which is also comparable to previous work (i.e 110-1290 ESVs in Goldford et al 2018). The metacommunity is simulated for 40 community generations with a fixed incubation time \( (t=1) \) and dilution factor \( (d=10^3 \times) \). This number of generations is equal to the largest number of community generations previously reported \cite{14}. The incubation time and dilution factor are set to ensure that all resources are fully depleted at the end of each incubation and the community reaches a stationary phase. The dilution factor is comparable to the largest dilution factor that has been applied to the previous studies (i.e., 1417-fold; most dilution factors reported are below 125-fold). The population size during stationary phase as well as the number of generation per incubation are of the same order of magnitude as those in experimental microbial populations (~10 bacterial generations per incubation and a final stationary phase population size of \( 10^9 \)). We note that whilst we have kept these parameters constant, they in fact varied substantially in prior community selection studies depending on the empirical systems (Supplementary Table 1). One protocol optimized incubation time between batches \cite{10}. While we did not capture this “variable t” feature, it is straightforward to do in ecoprospector.

Unlike previous studies we divide our experiments into two phases each lasting 20 generations. In the first ‘selection’ phase, the protocol-specific selection matrix \( S \) is applied at the end of each growth cycle. In the second ‘stabilization phase’, communities were passaged without selection, i.e., \( S=(1/d)I \), where \( I \) is the identity matrix and \( d \) the dilution factor, as above. Each protocol can
thus be expressed in a sequence of 40 selection matrices which contains 20 consecutive $S$ matrices and 20 consecutive $(1/d)I$ matrices (Supplementary Figure 1). The function and composition of all communities at the end of each generation (i.e., during stationary phase) is recorded. We have based our protocols on a fairly long period of stabilization (20 community generations) as we found that this guaranteed that communities had achieved generational stability across protocols. Experimenters may wish to optimize the number of generations and time between passaging to ensure rapid equilibration for the particular systems they work on.

We compare the effectiveness of each protocol by applying it (and where applicable its corresponding random selection control) to an identical set of starting communities. We also compare all protocols to a ‘no-selection’ control where all communities are passaged without selection for all generations, i.e $S = (1/d)I$ for 40 generations. We repeated all protocols 100 times with 100 different random seeds to obtain a statistically sound sample size.

**Quantifying the ecological resistance of directly evolved, synthetic and no selection communities.** In Figures 3B-D, We compare the function and resistance to ecological perturbations of i) a community obtained from directed evolution (DE) ii) a community constructed ‘synthetically’ and iii) the highest function community obtained from a no-selection line ($S = I/d$ for 40 transfers). The DE community was obtained by iteratively applying directed evolution using both bottlenecks and migration (as shown in Figure 3D and described in the previous section).

To construct the ‘synthetic community’ we first take the DE community and count the number of coexisting taxa ($n$). We then combine the $n$ top species in the species pool (i.e the $n$ species with the highest $\phi_i$) into a single community. These species are introduced at equal abundance with the total abundance of the synthetic community ($\Sigma N_i$) being equal to the total abundance of the directly evolved community. 3 metacommunities are simulated, one consisting of 96 copies of the DE community, one made up of 96 copies of the synthetic community and one made up of 96 copies of NS community. All three metacommunities are allowed to equilibrate for 20 generations (Fig 4B). The maximum function at this point $F_{\text{max}}$ is recorded. At generation 20 the 3 metacommunities are each subject to an identical round of directed evolution using migration from different regional species ($n_{\text{mig}} = 10^2$; see section on Directed Evolution). The metacommunity is then grown for another 20 generations without selection ($S = I/d$).

$F^*$ denotes the function of the new stable community that forms after the perturbation. We quantify the ecological resistance ($R$) as the deviation of community function after a pulse or press perturbation, as defined in $^{39}$.

$$R = 1 - \frac{2[F_{\text{max}} - F^*]}{|F_{\text{max}} + F_{\text{max}} - F^*|}, \quad \text{(Eq. 13)}$$

The resistance $R[{-1,1}]$, where $R=1$ if the community function has not changed after the perturbation. Note that the resistance defined here does not reflect the sign of functional
changes (e.g., increased or decreased function due to perturbation). In Figure 4C and 4D we show the $F^*$ and $R$ of the 95 compositional variants for each of the three communities considered.

We can calculate the overall resistance of a single community as the Mean($R$) and Mean($F^*$) across all of the 95 compositional variants (i.e. ignoring the unperturbed community). To obtain a statistically sound sample size we repeat the entire procedure 100 times (i.e. 100 different starting DE communities and 100 corresponding Synthetic communities and NS communities (Fig 4E-4F). In Supplementary Figure 9 we repeat this whole procedure using different types of perturbations other than migration (specifically bottleneck, resource shift, and species knock out). For these simulations we use $d_{bot}=10^4$ and $\delta = 1$.

**Iterative directed evolution.** To test the effects of iteratively applying directed evolution, we designed an extended protocol with over 460 community generations that includes 20 consecutive rounds of directed evolution (Figure 3). The protocol starts by seeding the initial metacommunity of 96 communities as before. We grew the communities for 30 generations without selection ($S=(1/d)I$ with $d=10^5 \times$). At generation 30 we performed a single round of perturbations (using one or more of the approaches described in the previous section). To reiterate, the top community was selected and passaged (with $d=10^5 \times$) into 96 fresh habitats and 95 of these copies were then perturbed to generate compositional variants. The offspring metacommunity was then stabilized for 20 generations without selection after which another round of perturbations was performed. We repeated this sequence of stabilization followed by perturbations until a total of 20 rounds of directed evolution was completed. After the final round of directed evolution (at generation 410) the metacommunity was grown for 50 generations without selection to ensure it reaches equilibrium.

We simulate this extended protocol using one of three different types of perturbation iteratively.

- **Bottlenecks:** After the top community has been selected and replicated using a standard dilution factor ($d$) 95 of 96 communities are subject to an additional bottleneck ($d_{bot}=10^4$). An average of N=95±9.7 (Mean±SD) cells remain in the community after each round of bottlenecking of this magnitude (Fig. 3B).
- **Migrations:** After the top community has been selected and replicated using a standard dilution factor ($d$) 95 out of 96 communities are subject to a round of migration ($n_{mig}=10^2$). Note that this amplitude is smaller than the single round of directed evolution using migration shown in Figure 2G. We choose this migration factor so that the number of cells introduced via migration was comparable to the number of cells left over by the bottlenecking (Fig. 3C).
- **Bottlenecks + Migrations:** After the top community has been selected and replicated using a standard dilution factor ($d$) 95 of 96 communities are first subject to a round of bottleneck ($d_{bot}=10^4$). These communities and then subject to a round of migration ($n_{mig}=10^2$).
To obtain a statistically sound sample size each of these approaches was repeated 100 times with 100 different random seeds (Fig 3E). Each iterative directed evolution experiment is compared to an equivalent NS line started with the same metacommunity. The metacommunity was propagated for 40 generations without selection (S=(1/d)l with d=10^3×).

Assumptions in the Microbial Consumer-Resource Model. Our ecoprospector package 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:

\[ \frac{dN_i}{dt} = g_i N_i \sum_a [(1 - l_a) w_a \sigma(c_{ia} R_a) - m_i] \]  
(Eq. S1)

\[ \frac{dR_a}{dt} = \sum_j N_j \sigma(c_{ia} R_a) \left[ D_{aj} \frac{w_j}{w_a} l_a \right] - \sum_j N_j \sigma(c_{ia} R_a) \]  
(Eq. S2)

All parameters are described and defined in Supplementary Table 2 (adapted from 20). 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_a = 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 21.

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 (\( \sigma(c_{ia} R_a) = \sigma_{III}(R_a) \)). In what follows, we explore the effect of relaxing these and other assumptions on our main results (Extended Data Fig. 1-5). 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 Extended Data Fig. 1 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 \(^{42}\). We introduce these types of functional interactions by considering a function that depends solely on the abundance of species pairs:

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

Here $\varepsilon$ is a square matrix in which the off-diagonals are sampled from a normal distribution $\varepsilon_{ij} \sim \text{Normal} (\text{mean} = 0, \text{sd} = 1)$ 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 (Extended Data Fig. 1). 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 (Extended Data Fig. 1E). 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 \(^{43-45}\)). 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 ($\phi_i$). In such a scenario there would be strict negative relationship between the per-capita contribution to function $\phi_i$ and the species growth efficiency $g_i$. We impose this relationship by first sampling per-capita contribution from a uniform distribution $\text{Uniform} (\text{min} = 0, \text{max} = 1)$ 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 Extended Data Fig. 1E). 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, or a community that efficiently consumes a specific nutrient (for bioremediation purposes). 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 ($R_{\text{target}}$). Therefore

$$F = -R_{\text{target}}$$  \hspace{1cm} (Eq. S5)

For the resource-shift perturbation, we excluded the target resource from the list of spiked-in resources (Extended Data Fig. 1C). 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 Extended Data Fig. 1E). 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±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. 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 $n_{\text{invader}}$:

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

For this function we find that all of our proposed protocols are able to successfully generate higher functioning communities (Extended Data Fig. 1C and Extended Data Fig. 1D). As with the other non-additive functions, the synthetic community does substantially worse than the additive function presented in the main text (Extended Data Fig. 1E). In contrast to the other functions we have explored, we do find that previously proposed pooling methods do better than the no selection control (Extended Data Fig. 1B). 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. 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 $51-53$) are significantly higher than the highest functioning communities from the no selection control ($26\pm4.6$ vs $16\pm2.5$ for species richness, $999\pm2$ vs $981\pm15$ 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. 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 sampled 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 $9$. Both conditions have the same number of resources (M) and total amount of supplied resources
We show in Extended Data Fig. 2, that neither the presence of cross-feeding nor the single-resource environment significantly changed our results. The lack of significance in Extended Data Fig. 2D 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
  \[ \sigma_I(R_a) = c_\text{lin} R_a \]  
  (Eq. S7)
- A Monad function (Type-II),
  \[ \sigma_{II}(R_a) = \frac{c_\text{m} R_a}{1 + \frac{c_\text{m} R_a}{\sigma_{\text{max}}}} \]  
  (Eq. S8)
- A Hill function (Type-III)
  \[ \sigma_{III}(R_a) = \frac{(c_\text{n} R_a)^n}{1 + \frac{c_\text{n} R_a}{\sigma_{\text{max}}}} \]  
  (Eq. S9)

where \( n \) is the Hill coefficient for functional response and \( \sigma_{\text{max}} \) 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 \( \sigma_{\text{max}} = 1 \) and \( n = 2 \). In Extended Data Fig. 3, 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 \( n_{\text{inoc}} = 10^6 \) cells from a different regional pool where the species abundance of all \( H \) species in each pool follows a power-law distribution:
  \[ P_{\text{power}}(x) = ax^{a-1} \]  
  (Eq. S10)
(ii) Each community is generated by sampling \( n_{\text{inoc}} = 10^6 \) cells from a different regional pool where the species abundance of all \( H \) species in each pool follows a log-normal distribution
  \[ P_{\text{log-normal}}(x) = \frac{1}{x\alpha\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} \]  
  (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 \( n_{\text{inoc}} = 10^6 \)
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 Supplementary Fig. 11 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 9. In addition, we show in Extended Data Fig. 4, 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 $\phi_i \sim \text{Normal}(\text{mean} = 0, sd = 1)$. 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 $\phi_i$:

1. Most contributions are positive but strong contributors remain rare $\phi_i \sim \text{Normal}(\text{mean} = 1, sd = 1)$
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: $\phi_i \sim \text{Normal}(\text{mean} = 0, sd = 1) \times \text{Bernoulli}(0.2)$

In Extended Data Fig. 5 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.

**REFERENCES SUPPLEMENTARY TEXT:**

1. Rubin, B. E. *et al.* Targeted Genome Editing of Bacteria Within Microbial Communities. *bioRxiv* (2020) doi:10.1101/2020.07.17.209189.

2. Sheth, R. U., Cabral, V., Chen, S. P. & Wang, H. H. Manipulating Bacterial Communities by in situ Microbiome Engineering. *Trends Genet.* **32**, 189–200 (2016).
3. Ronda, C., Chen, S. P., Cabral, V., Yaung, S. J. & Wang, H. H. Metagenomic engineering of the mammalian gut microbiome in situ. *Nat. Methods* **16**, 167–170 (2019).

4. Franklin, R. B. & Mills, A. L. Structural and functional responses of a sewage microbial community to dilution-induced reductions in diversity. *Microb. Ecol.* **52**, 280–288 (2006).

5. Lee, D.-J., Show, K.-Y. & Wang, A. Unconventional approaches to isolation and enrichment of functional microbial consortium—a review. *Bioresour. Technol.* **136**, 697–706 (2013).

6. Shepherd, E. S., DeLoache, W. C., Pruss, K. M., Whitaker, W. R. & Sonnenburg, J. L. An exclusive metabolic niche enables strain engraftment in the gut microbiota. *Nature* **557**, 434–438 (2018).

7. Doulcier, G., Lambert, A., De Monte, S. & Rainey, P. B. Eco-evolutionary dynamics of nested Darwinian populations and the emergence of community-level heredity. *Elife* 53433 (2020).

8. Xie, L., Yuan, A. E. & Shou, W. Simulations reveal challenges to artificial community selection and possible strategies for success. *PLoS Biol.* **17**, e3000295 (2019).

9. Goldford, J. E. *et al.* Emergent simplicity in microbial community assembly. *Science* **361**, 469–474 (2018).

10. Wright, R. J., Gibson, M. I. & Christie-Oleza, J. A. Understanding microbial community dynamics to improve optimal microbiome selection. *Microbiome* **7**, 85 (2019).

11. Kang, D. *et al.* Enrichment and characterization of an environmental microbial consortium displaying efficient keratinolytic activity. *Bioresour. Technol.* **270**, 303–310 (2018).

12. Lazuka, A., Auer, L., O’Donohue, M. & Hernandez-Raquet, G. Anaerobic lignocellulolytic microbial consortium derived from termite gut: enrichment, lignocellulose degradation and community dynamics. *Biotechnol. Biofuels* **11**, 284 (2018).

13. Estrela, S. *et al.* Metabolic rules of microbial community assembly. *bioRxiv* 2020.03.09.984278
14. Swenson, W., Wilson, D. S. & Elias, R. Artificial ecosystem selection. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9110–9114 (2000).

15. Mueller, U. G. *et al.* Artificial Microbiome-Selection to Engineer Microbiomes That Confer Salt-Tolerance to Plants. *bioRxiv* 081521 (2016) doi:10.1101/081521.

16. Jochum, M. D., McWilliams, K. L., Pierson, E. A. & Jo, Y.-K. Host-mediated microbiome engineering (HMME) of drought tolerance in the wheat rhizosphere. *PLoS One* 14, e0225933 (2019).

17. Arora, J., Mars Brisbin, M. A. & Mikheyev, A. S. Effects of microbial evolution dominate those of experimental host-mediated indirect selection. *PeerJ* 8, e9350 (2020).

18. Niehaus, L. *et al.* Microbial coexistence through chemical-mediated interactions. *Nat. Commun.* 10, 2052 (2019).

19. Xie, L. & Shou, W. Steering ecological-evolutionary dynamics during artificial selection of microbial communities. *bioRxiv* 264697 (2020) doi:10.1101/264697.

20. Panke-Buisse, K., Poole, A. C., Goodrich, J. K., Ley, R. E. & Kao-Kniffin, J. Selection on soil microbiomes reveals reproducible impacts on plant function. *ISME J.* 9, 980–989 (2015).

21. Fisher, C. K. & Mehta, P. The transition between the niche and neutral regimes in ecology. *Proc. Natl. Acad. Sci. U. S. A.* 111, 13111–13116 (2014).

22. Amor, D. R., Ratzke, C. & Gore, J. Transient invaders can induce shifts between alternative stable states of microbial communities. *Sci Adv* 6, eaay8676 (2020).

23. Locey, K. J. & Lennon, J. T. A Residence Time Theory for Biodiversity. *Am. Nat.* 194, 59–72 (2019).

24. Bunin, G. Ecological communities with Lotka-Volterra dynamics. *Phys Rev E* 95, 042414 (2017).
25. Biroli, G., Bunin, G. & Cammarota, C. Marginally stable equilibria in critical ecosystems. New J. Phys. **20**, 083051 (2018).

26. Pearce, M. T., Agarwala, A. & Fisher, D. S. Stabilization of extensive fine-scale diversity by ecologically driven spatiotemporal chaos. Proc. Natl. Acad. Sci. U. S. A. **117**, 14572–14583 (2020).

27. Wei, Z. et al. Trophic network architecture of root-associated bacterial communities determines pathogen invasion and plant health. Nat. Commun. **6**, 8413 (2015).

28. Wade, M. J. Group selections among laboratory populations of Tribolium. Proc. Natl. Acad. Sci. U. S. A. **73**, 4604–4607 (1976).

29. Wade, M. J. An experimental study of group selection. Evolution **31**, 134–153 (1977).

30. Wade, M. J. A Critical Review of the Models of Group Selection. Q. Rev. Biol. **53**, 101–114 (1978).

31. Goodnight, C. J. Experimental Studies of Community Evolution I: The Response to Selection at the Community Level. Evolution **44**, 1614–1624 (1990).

32. Goodnight, C. J. Evolution in metacommunities. Philos. Trans. R. Soc. Lond. B Biol. Sci. **366**, 1401–1409 (2011).

33. Torsvik, V. & Øvreås, L. Microbial diversity and function in soil: from genes to ecosystems. Curr. Opin. Microbiol. **5**, 240–245 (2002).

34. Blouin, M., Karimi, B., Mathieu, J. & Lerch, T. Z. Levels and limits in artificial selection of communities. Ecol. Lett. **18**, 1040–1048 (2015).

35. Goodnight, C. J. Heritability at the ecosystem level. Proceedings of the National Academy of Sciences of the United States of America vol. 97 9365–9366 (2000).

36. Chang, C.-Y., Osborne, M. L., Bajic, D. & Sanchez, A. Artificially selecting bacterial communities using propagule strategies. Evolution (2020) doi:10.1111/evo.14092.
37. Swenson, W., Arendt, J. & Wilson, D. S. Artificial selection of microbial ecosystems for 3-chloroaniline biodegradation. *Environ. Microbiol.* **2**, 564–571 (2000).

38. Raynaud, T., Devers, M., Spor, A. & Blouin, M. Effect of the Reproduction Method in an Artificial Selection Experiment at the Community Level. *Frontiers in Ecology and Evolution* **7**, 416 (2019).

39. Shade, A. *et al.* Fundamentals of microbial community resistance and resilience. *Front. Microbiol.* **3**, 417 (2012).

40. Marsland, R., Cui, W., Goldford, J. & Mehta, P. The Community Simulator: A Python package for microbial ecology. *PLoS One* **15**, e0230430 (2020).

41. Marsland, R., 3rd, Cui, W. & Mehta, P. A minimal model for microbial biodiversity can reproduce experimentally observed ecological patterns. *Sci. Rep.* **10**, 3308 (2020).

42. Enke, T. N. *et al.* Modular Assembly of Polysaccharide-Degrading Marine Microbial Communities. *Curr. Biol.* **29**, 1528–1535.e6 (2019).

43. Cordero, O. X., Ventouras, L.-A., DeLong, E. F. & Polz, M. F. Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 20059–20064 (2012).

44. Griffin, A. S., West, S. A. & Buckling, A. Cooperation and competition in pathogenic bacteria. *Nature* **430**, 1024–1027 (2004).

45. West, S. A. & Buckling, A. Cooperation, virulence and siderophore production in bacterial parasites. *Proc. Biol. Sci.* **270**, 37–44 (2003).

46. Sierocinski, P. *et al.* A Single Community Dominates Structure and Function of a Mixture of Multiple Methanogenic Communities. *Curr. Biol.* **27**, 3390–3395.e4 (2017).

47. Embree, M., Liu, J. K., Al-Bassam, M. M. & Zengler, K. Networks of energetic and metabolic interactions define dynamics in microbial communities. *Proceedings of the National Academy of Sciences* **111**, 8281–8286 (2014).
48. Piccardi, P., Vessman, B. & Mitri, S. Toxicity drives facilitation between 4 bacterial species. *Proceedings of the National Academy of Sciences* **116**, 15979–15984 (2019).

49. van der Gast, C. J., Knowles, C. J., Starkey, M. & Thompson, I. P. Selection of microbial consortia for treating metal-working fluids. *J. Ind. Microbiol. Biotechnol.* **29**, 20–27 (2002).

50. Yin, C. *et al.* Rhizosphere Community Selection Reveals Bacteria Associated With Reduced Root Disease. *In Review* (2020) doi: 10.21203/rs.3.rs-64051/v1.

51. Tilman, D. *Resource Competition and Community Structure*. (Princeton University Press, 1982).

52. Acosta, F., Zamor, R. M., Najar, F. Z., Roe, B. A. & Hambright, K. D. Dynamics of an experimental microbial invasion. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11594–11599 (2015).

53. 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.
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 $\rho = 0.25$ (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) $F_{\text{max}}$ between the AS and RS lines. Only experimental protocols that have described RS are shown (Supplementary Table 1). 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 (NS). 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 $\rho$ between Rank F at Start and Rank F at equilibrium for this dark blue points (C) Distribution of Spearman’s $\rho$ for all 100 NS lines. Dotted red line corresponds to Spearman’s $\rho$ 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 $Q=F_{\text{max}}[\text{AS}]-F_{\text{max}}[\text{NS}]$. 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). ****: $p<0.0001$. 

$\rho = 0.289$
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 (Supplementary Fig. 7). (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 \( (iN_i) \) for the top one of 96 communities in Supplementary Figure 5A 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 \( F = \sum_i N_i \) from 841 to 526 shown in Supplementary Fig 5A 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 ($d_{bot}=10^3$) (top left), resource shifts ($\delta = 1$) (bottom left), or species knock-outs (bottom right).
Supplementary Figure 10. Directly evolved communities are enriched for taxa with higher per-capita function, but are not necessarily enriched with taxa that are more competitive in monoculture. For 20 replicate simulations we compare the taxa present in the highest functioning directly evolved communities with the taxa in the highest functioning (stabilized) ancestral community. The 20 communities plotted correspond to 20 of the 100 replicate communities generated by iteratively combining bottlenecks and migration as shown in Fig. 3. (A) Distribution of $\phi_i$ for ancestral and directly evolved communities. (B) Distribution of $N_i$ in monoculture of taxa found in ancestral and directly evolved communities. For both plots annotations represent significance values for welch’s t-test. *:p<0.05, **:p<0.01, ***:p<0.001 ****:p<0.0001
Supplementary Figure 11. Rarefaction curves of metacommunity sampling approaches. An empirical rarefaction curve 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 Extended Data Fig. 4 we show results when using either a lognormal distribution or assuming uniform initial abundances.
Supplementary Figure 12. 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 strategies (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 $2 \times 10^6$ 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 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 | Random selection matrix |
|----------------|---------------------------|----------------------------------------|----------------------------------------|--------------------------|----------------------|--------------------------|-------------------------------|----------------------------------|---------------------|---------------------|------------------------|
| migrant pool   | Swenson2000a              | plant-associated soil                  | soil microbiome, plant host biomass   | yes                      | 16                   | 2                       | 15                           | 3                               | 0.20                | 14 and 1417         |                       |
|                | Blount2015                | wastewater treatment plant             | lowest CO2 emission                   | yes                      | 20                   | 6                       | 30                           | 3                               | 0.10                | 16                  |                       |
|                | Parke-Bissell2015         | plant-associated soil, microbiome      | early or late flowering time in plant host | not available            | 10                   | 1                       | 14                           | 4                               | 0.28                | unclear             |                       |
|                | Jochum2019                | soil microbiome associated to drought-tolerant 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               | rhizosphere                            | 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                 | fruity gull                             | shortest host excision time           | yes                      | 4                    | 10                      | 3                            | 1                               | 0.33                | unclear             |                       |
|                | Raynaud2019a              | topsoil                                 | biomass estimated by OD              | yes                      | 14                   | 3                       | 10                           | 1                               | 0.10                | 20                  |                       |
|                | Chang2008a                | synthetic communities with four known strains | amylolytic activities                | yes                      | 17                   | 1                       | 24                           | 4                               | 0.15                | 10                  |                       |
|                | Chang2008b                | soil and leaves                         | cross-feeding potential              | yes                      | 7                    | 1                       | 92                           | 23                              | 0.25                | 125                 |                       |

For illustration convenience, the selection matrices shown here are designed for 24 communities rather than 98 that are used otherwise in the main text.

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

The 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     |
| $R_\alpha$ | Concentration of resource $\alpha$ (mass/volume) | a     |
| $C_{\alpha i}$ | Uptake rate per unit concentration of resource $\alpha$ by species i (volume/time) | b     |
| $D_{\alpha \beta}$ | Fraction of byproducts from resource $\beta$ converted to $\alpha$ (unitless) | bc    |
| $g_i$     | Conversion factor from energy uptake to growth rate (1/energy) | 1     |
| $w_\alpha$ | Energy content of resource $\alpha$ (energy/mass) | 1     |
| $l_\alpha$ | Leakage fraction for resource $\alpha$ (unitless) | 0     |
| $m_i$     | Minimal energy uptake for maintenance of species i (energy/time) | 0     |
| $\sigma_\alpha$ | Functional response of utilization on resource $\alpha$ | 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 $l_\alpha$ 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, $n_{\text{inoc}}$, and $\alpha$.

| 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 |
| $R_{\text{tot}}$ | 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   |
| $\sigma_c$ | Standard deviation of sum over a row of the preference matrix $c_{ia}$ | 3   |
| $c_0$     | Low consumption level for Binary $c_{ia}$ | 0    |
| $c_i$     | 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 |
| $\psi$    | Number of cells when $N_i = 1$ | 1e+06 |
| $n_{\text{inoc}}$ | Number of cells in the initial inoculum | 1e+06 |
| $\alpha$  | Relative functional contribution of species interaction to the additive case | 1    |

$^a$These values do not matter if $S_f$ is 1  
$^b$This value does not matter if $I_e$ 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; Supplementary Table 1).

| Parameter | Description and units                               | Value |
|-----------|-----------------------------------------------------|-------|
| d         | Dilution factor in the batch culture                | 0.001 |
| t         | Incubation time                                     | 1     |
| n<sub>well</sub> | Number of wells; number of metacommunities     | 96    |
| T<sub>tot</sub> | Number of total transfers (generations)          | 40    |
| T<sub>selc</sub> | Number of selection transfers (generations)      | 20    |
Supplementary Table 5. Parameters for directed selection in Figure 2D-2I.

| Parameter | Description and units                                                                 | Value  |
|-----------|--------------------------------------------------------------------------------------|--------|
| $\theta$  | 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    |
| $\delta$  | Tunes the magnitude of resource perturbation. The fraction from depleting a resource and move the same amount to another | 1      |