Provisional Title

Authors[1,2,3] and Álvaro Sánchez[1,2,†]

[1]*Department of Ecology & Evolutionary Biology, Yale University, New Haven, CT, USA*
[2]*Microbial Sciences Institute, Yale University, New Haven, CT, USA*
[3]*Other affiliations...*
[†]*To whom correspondence should be addressed:* alvaro.sanchez@yale.edu

Abstract

The abstract goes here.

Introduction

Microbial communities often invade one another. This has been observed, for instance, in river courses where terrestrial microbial communities mix with aquatic microorganisms [1–3] or in soil communities being invaded as a result of tillage and outplanting [4] or by aerially dispersed bacteria and funghi [5]. The human digestive system can get invaded several times a day by the microbial consortia that reside on the ingested food, and the skin microbiota is also subject to invasions when making contact with environmental sources of microbes [6].

The phenomenon by which entire microbiomes invade one another has been termed *community coalescence* [7]. Ecologists have long contemplated the idea that interactions between multiple co-invading species can produce correlated invasional outcomes [7–13]. One example is the hypothesis known as *invasional meltdown*, which proposes that positive interactions between co-invading species can enhance their invasive success and facilitate future invasions [14–16]. However, and in spite of its clear potential importance, the role of coalescence in microbiome assembly is only beginning to be addressed and little is known about the mechanisms that govern it and its potential implications. Early mathematical models of community-community invasions [8, 17] as well as more recent work [18–21] suggest that high-order invasion effects are common during community coalescence. Communities that have a previous history of coexistence may exhibit an emergent "cohesiveness" which produces correlated invasional outcomes among species from the same community [11, 22]. The situation where ecological partners in the invading community recruit each other into the final coalesced community has been called *ecological co-selection* [22, 23].

The mechanisms of ecological co-selection during community coalescence are still poorly understood. Do a few key species recruit everyone else, or are collective interactions among all species (including the rarer members of the community) relevant for coalescence outcomes? While it is reasonable to expect species with larger population sizes to have a proportionally oversized effect, natural communities tend to be highly diverse [24] and the role played by the less abundant species has long been subject to debate [25]. Laboratory cultures have also been found to contain uneven distributions of multiple strains that feed off the metabolic secretions of the dominant species [26, 27]. The fate of these sub-dominant taxa may be dependent on the invasional success of their dominant species, or, alternatively, the dominant itself may owe its dominance (at least in part) to cross-feeding or other forms of facilitation from the rarer members of the population. These scenarios would give rise to "top-down" or "bottom-up" community cohesiveness, respectively. Either of these forms of co-selection could, in principle, be positive (recruitment) or negative (antagonism), as illustrated in Figure 1e. Which of these situations are typically found in nature? Addressing this question has been experimentally challenging in the past [22, 23].

In previous work, we have shown that a large amount of soil and plant microbiomes can be cultured *ex situ* in synthetic minimal environments with a single supplied limiting resource under serial growth-dilution cycles [27] (Figure 1a-b). Under these conditions, environmental microbiomes spontaneously re-assemble into complex multi-species communities sustained by dense cross-feeding facilitation networks [27]. In

addition, and just like in natural assemblies, species abundance distributions in these communities are generally long-tailed and uneven (Figure 1d and Figure S1), with the dominant (most abundant) species typically comprising most of the biomass (median = 46%, Figure S1). Because these communities are easy to manipulate and grow in high throughput, and are largely made up by culturable members, they represent good test cases to investigate ecological co-selection during community coalescence. Here we focus on the dominants and ask whether they can co-select or be co-selected by the sub-dominant species in their communities (henceforth referred to as their *cohorts*, Figure 1c). Our results confirm that positive top-down co-selection is common, but its effects are weak. In contrast, bottom-up co-selection can be very strong, and positive co-selection is far more common than negative co-selection. We then turn to a microbial consumer-resource model (microCRM) [27–29] that is able to capture the dynamics of microbial communities dominated by metabolic interactions, as is the case for the ones assembled in our experimental conditions [27]. We show that the empirically observed trends in ecological co-selection are reproduced with minimal model assumptions, and that tuning the complexity of the metabolic interactions in our *in silico* communities can modulate the recurrence of top-down or bottom-up co-selection. Our findings indicate that collective interactions play an important role at dictating community structure during coalescence.

Results & Discussion

We collected eight natural microbiomes from different soil and plant environmental samples (Figure 1a) and used them to inoculate our synthetic communities, which were stabilized in serial batch-culture bioreactors for 84 generations in synthetic minimal media containing either glutamine or citrate as the only supplied carbon source (Figure 1b, Methods: Stabilization of environmental communities in simple synthetic environments). We isolated the dominant species of every community (Methods: Isolation of dominant species) and identified them by Sanger-sequencing their 16S rRNA gene (Methods: Determination of community composition by 16S sequencing), which correctly matched the dominant Exact Sequence Variant (ESV) [30, 31] found through community-level 16S illumina sequencing (Figure S1). These dominants remained at high frequency after seven additional transfers with the exception of two of the citrate communities and one of the glutamine communities (where the dominants were presumably a transiently dominating species) that were excluded from further analysis (Figure S1). Similarly, pairs of communities where the dominants shared a same 16S sequence and had similar colony morphology were excluded (Figure S1).

Top-dow co-selection is weak

If communities being coalesced were highly cohesive from the top-down, the dominant species would co-select the rarer members of its community during coalescence (Figure 1e, left panels). In this scenario, we would expect the outcome of community coalescence to be predicted by which of the two dominants is most competitive in pairwise competition. Coalescence outcomes can be quantified by the similarity between the coalesced and the invasive communities (Methods: Metrics of community distance). To test this hypothesis, we performed all pairwise competitions between dominant species in glutamine and citrate environments by mixing them 1:1 on their native media and propagating the cultures for seven serial transfers, roughly 42 generations (Methods: Dominant-dominant and community-community competitions). We then performed all possible pairwise community coalescence experiments by mixing equal volumes of the communities and propagating the resulting cultures for seven extra transfers (Figure 1f). The frequencies of all species in both community-community and dominant-dominant competitions were determined by 16S illumina sequencing (Methods: Determination of community composition by 16S sequencing). We found that the pairwise competitive ability of an invasive dominant is only weakly predictive of the performance of the invasive community in coalescence, as quantified by the relative Bray-Curtis similarity between the coalesced and invasive communities (Figure 2a, $R^2 = 0.15$, $p <$ for glutamine and $R^2 = 0.15$, $p <$ for citrate). Alternative quantifications of community distance also yield weak effects, and even more so when the metric used accounts only for the presence/absence of specific species and not for their relative abundance in the communities (Figure S2).

Methods

Stabilization of environmental communities in simple synthetic environments

Communities were stabilized *ex situ* as described in [27]. In short, environmental samples (soil, leaves...) within one meter radius in eight different geographical locations were collected with sterile tweezers or spatulas into 50mL sterile tubes (Fig. **[missing ref(s)]**). One gram of each sample was allowed to sit at room temperature in 10mL of phosphate buffered saline (1×PBS) containing 200µg/mL cycloheximide to suppress eukaryotic growth. After 48h, samples were mixed 1:1 with 80% glycerol and kept frozen at −80°C. Starting microbial communities were prepared by scrapping the frozen stocks into 200µL of 1×PBS and adding a volume of 4µL to 500µL of synthetic minimal media (1×M9) supplemented with 200µg/mL cycloheximide and 0.07 C-mol/L glutamine or sodium citrate as the carbon source in 96 deep-well plates (1.2mL; VWR). Cultures were then incubated still at 30°C to allow for re-growth. After 48h, samples were fully homogenized and biomass increase was followed by measuring the optical density (620nm) of 100µL of the cultures in a Multiskan FC plate reader (Thermo Scientific). Communities were stabilized [27] by passaging 4µL of the cultures into 500µL of fresh media (1×M9 with the carbon source) every 48h for a total of 12 transfers at a dilution factor of 1:100, roughly equivalent to 80 generations per culture (Fig. **[missing ref(s)]**). Cycloheximide was not added to the media after the first two transfers.

Isolation of dominant species

For each community, the most abundant colony morphotype at the end of the ninth transfer was selected, resuspended in 100µL 1×PBS and serially diluted (1:10). Next, 20µL of the cells diluted to 10^{-6} were plated in the corresponding synthetic minimal media and allowed to regrow at 30°C for 48h. Dominants were then inoculated into 500µL of fresh media and incubated still at 30°C for 48h. After this period, the communities stabilized for eleven transfers and the isolated dominants were ready for the competition experiments (Fig **[missing ref(s)]**) at the onset of the twelfth transfer.

Dominant-dominant and community-community competitions

All possible pairwise dominant-dominant and community-community competition experiments were performed by mixing equal volumes (4µL) of each of the eight communities or eight dominants at the onset of the twelfth transfer. Competitions were set up in their native media, i.e. in 500µL of 1×M9 supplemented with 0.07 C-mol/L of either glutamine or citrate in 96 deep-well plates. Plates were incubated at 30°C for 48h. Pairwise competitions were further propagated for seven serial transfers (roughly 42 generations; Fig. **[missing ref(s)]**) by transferring 8µL of each culture to fresh media (500µL).

Determination of community composition by 16S sequencing

The sequencing protocol was identical to that described in [27]. Community samples were collected by spinning down at 3500rpm for 25min in a bench-top centrifuge at room temperature; cell pellets were stored at −80°C before processing. To maximize Gram-positive bacteria cell wall lysis, the cell pellets were re-suspended and incubated at 37°C for 30min in enzymatic lysis buffer (20mM Tris-HCl, 2mM sodium EDTA, 1.2% Triton X-100) and 20mg/mL of lysozyme from chicken egg white (Sigma-Aldrich). After cell lysis, the DNA extraction and purification was performed using the DNeasy 96 protocol for animal tissues (Qiagen). The clean DNA in 100µL elution buffer of 10mM Tris-HCl, 0.5mM EDTA at pH 9.0 was quantified using Quan-iT PicoGreen dsDNA Assay Kit (Molecular Probes, Inc.) and normalized to 5ng/µL in nuclease-free water (Qiagen) for subsequent 16S rRNA illumina sequencing. 16S rRNA amplicon library preparation was performed following a dual-index paired-end approach [32]. Briefly, PCR amplicon libraries of V4 regions of the 16S rRNA were prepared sing dual-index primers (F515/R805), then pooled and sequenced using the Illumina MiSeq chemistry and platform. Each sample went through a 30-cycle PCR in duplicate of 20µL reaction volumes using 5ng of DNA each, dual index primers, and AccuPrime Pfx SuperMix (Invitrogen). The thermocycling procedure includes a 2min initial denaturation step at 95°C, and 30 cycles of the following PCR scheme: (a) 20-second denaturation at 95°C, (b) 15-second annealing at 55°C, and (c) 5-minute extension at 72°C. The duplicate PCR products of each sample were pooled, purified, and normalized using SequalPrep PCR cleanup and normalization kit (Invitrogen). Barcoded amplicon libraries were then pooled and sequenced using Illumina Miseq v2 reagent kit, which

generated 2×250bp paired-end reads at the Yale Center for Genome Analysis (YCGA). The sequencing reads were demultiplexed on QIIME 1.9.0 [33]. The barcodes, indexes, and primers were removed from raw reads, producing FASTQ files with both the forward and reverse reads for each sample, ready for DADA2 analysis [31]. DADA2 version 1.1.6 was used to infer unique biological exact sequence variants (ESVs) for each sample and naïve Bayes was used to assign taxonomy using the SILVA version 123 database [34, 35].

Metrics of community distance

Beta-diversity indexes between the invasive and coalesced communities or the resident and coalesced communities were computed using various similarity metrics. For two arbitrary communities with ESV abundances represented by the vectors $\mathbf{x} = (x_1, x_2, \cdots, x_N)$ and $\mathbf{y} = (y_1, y_2, \cdots, y_N)$ (where x_i and y_i represent the relative abundance of the ith ESV in each community respectively and N is the total number of ESVs), the Bray-Curtis similarity $BC(\mathbf{x}, \mathbf{y})$ is calculated as [36]

$$BC(\mathbf{x}, \mathbf{y}) = \sum_i \min(x_i, y_i) \tag{1}$$

The Jensen-Shannon similarity $JS(\mathbf{x}, \mathbf{y})$ is defined as one minus the Jensen-Shannon distance (which is, in turn, the square root of the Jensen-Shannon divergence [37])

$$JS(\mathbf{x}, \mathbf{y}) = 1 - \sqrt{\frac{1}{2}KL(\mathbf{x}, \mathbf{m}) + \frac{1}{2}KL(\mathbf{y}, \mathbf{m})} \tag{2}$$

where $\mathbf{m} = (\mathbf{x} + \mathbf{y})/2$ and KL denotes the Kullback-Leibler divergence [38]

$$KL(\mathbf{x}, \mathbf{y}) = \sum_i x_i \log_2\left(\frac{x_i}{y_i}\right) \tag{3}$$

The Jaccard similarity is given by $J(\mathbf{x}, \mathbf{y})$ [39]

$$J(\mathbf{x}, \mathbf{y}) = \frac{|\,\mathbf{x} \cap \mathbf{y}\,|}{|\,\mathbf{x} \cup \mathbf{y}\,|} \tag{4}$$

Additionally, we quantify coalescence outcomes by examining the fraction of the endemic cohort of the original communities that persists in the coalesced one. We call $E(\mathbf{x}, \mathbf{y})$ to the fraction of endemic species of $\mathbf{x}$ that are also found in $\mathbf{y}$.

For all the metrics above, we quantify the relative similarity between the invasive and the coalesced communities using relative metrics (Q):

$$Q(\mathbf{x}_\mathrm{I}, \mathbf{x}_\mathrm{R}, \mathbf{x}_\mathrm{C}) = \frac{F(\mathbf{x}_\mathrm{I}, \mathbf{x}_\mathrm{C})}{F(\mathbf{x}_\mathrm{I}, \mathbf{x}_\mathrm{C}) + F(\mathbf{x}_\mathrm{R}, \mathbf{x}_\mathrm{C})} \tag{5}$$

where the subindices I, R and C correspond to the invasive, resident and coalesced communities respectively, and F represents one of BC (Bray-Curtis similarity), JS (Jensen-Shannon similarity), J (Jaccard similarity) or E (endemic survival) defined above.

Simulations

We used the Community Simulator package [29] and included new features for our simulations. In the package, species are characterized by their resource uptake rates ($c_{i\alpha}$ for species i and resource α), and they all share a common metabolic matrix $\mathbf{D}$. The element $D_{\alpha\beta}$ of this matrix represents the fraction of energy in the form of resource α secreted when resource β is consumed. Here we implemented a new operation mode in which species can secrete different metabolites (and/or in different abundances) when consuming a same resource. Experimental observations support the idea of distinct species producing different sets of byproducts when growing in the same primary resource **[missing ref(s)]**. We call $D_{i\alpha\beta}$ to the fraction of energy in the form of resource α secreted *by species* i when consuming resource β —note that now $D_{i\alpha\beta}$ need not be equal to $D_{j\alpha\beta}$ if $i \neq j$, unlike in the original Community Simulator. In the package's underlying Microbial Consumer Resource Model [27, 28], this just means that the energy flux $J^{\mathrm{out}}_{i\beta}$ now takes the form

$$J^{\mathrm{out}}_{i\beta} = \sum_\alpha D_{i\beta\alpha} l_\alpha J^{\mathrm{in}}_{i\alpha} \tag{6}$$

The documentation for the Community Simulator contains detailed descriptions of the model, parameters and package use. For the updated package with the new functionality, see Data & code availability.

For our simulations, we first generate a library of 660 species (divided into three specialist families of 200 members each and a generalist family of 60 members) and 30 resources (divided into three classes of 10 members each). We split this library into two non-overlapping pools of 330 species each. We randomly sample 50 species from each pool in equal ratios to seed 100 resident and 100 invasive communities respectively. We then grow and dilute the communities serially, replenishing the primary resource after each dilution. We repeat the process 20 times to ensure generational equilibrium is achieved [27]. We then perform the *in silico* experiments by using the generationally stable communities to seed 100 coalesced communities that we again stabilize as described previously. Similarly, we identify the dominant (most abundant) species of every resident and invasive community to carry out pairwise competition and single invasion simulations. Most parameters are set to the defaults of the original Community Simulator package. Table **[missing ref(s)]**shows those that are given non-default values to ensure enough variation in the primary communities.

Data & code availability

Experimental data and code for the analysis, as well as code for the simulations and the updated Community Simulator package with instructions for the new features can be found in `github.com/jdiazc9/coalescence`.

Acknowledgements

The authors wish to thank Joshua Goldford, Pankaj Mehta, Wenping Cui, Robert Marsland and all members of the Sanchez laboratory for many helpful discussions. We also wish to express our gratitude to the Goodman laboratory at Yale for technical help during the early stages of this project. The funding for this work partly results from a Scialog Program sponsored jointly by the Research Corporation for Science Advancement and the Gordon and Betty Moore Foundation through grants to Yale University by the Research Corporation and the Simons Foundation.

References

1. Mansour I, Heppell CM, Ryo M and Rillig MC (2018). Application of the microbial community coalescence concept to riverine networks. *Biological Reviews* **93(4)**:1832–1845

2. Luo X, Xiang X, Yang Y, Huang G, Fu K, Che R and Chen L (2020). Seasonal effects of river flow on microbial community coalescence and diversity in a riverine network. *FEMS Microbiology Ecology* **96(8)**:fiaa132

3. Vass M, Székely AJ, Lindström ES, Osman OA and Langenheder S (2021). Warming mediates the resistance of aquatic bacteria to invasion during community coalescence. *Molecular Ecology* **30(5)**:1345–1356

4. Rillig MC, Lehmann A, Aguilar-Trigueros CA, Antonovics J, Caruso T, Hempel S, Lehmann J, Valyi K, Verbruggen E et al. (2016). Soil microbes and community coalescence. *Pedobiologia* **59(1-2)**:37–40

5. Evans SE, Bell-Dereske LP, Dougherty KM and Kittredge HA (2019). Dispersal alters soil microbial community response to drought. *Environmental Microbiology* **22(3)**:905–916

6. Vandegrift R, Fahimipour AK, Muscarella M, Bateman AC, Wymelenberg KVD and Bohannan BJ (2019). Moving microbes: the dynamics of transient microbial residence on human skin. *biorXiv*

7. Rillig MC, Antonovics J, Caruso T, Lehmann A, Powell JR, Veresoglou SD and Verbruggen E (2015). Interchange of entire communities: microbial community coalescence. *Trends in Ecology & Evolution* **30(8)**:470–476

8. Gilpin M (1994). Community-level competition: asymmetrical dominance. *Proceedings of the National Academy of Sciences* **91(8)**:3252–3254

9. Grosholz ED (2005). Recent biological invasion may hasten invasional meltdown by accelerating historical introductions. *Proceedings of the National Academy of Sciences* **102(4)**:1088–1091

10. Green PT, O'Dowd DJ, Abbott KL, Jeffery M, Retallick K and Nally RM (2011). Invasional meltdown: Invader–invader mutualism facilitates a secondary invasion. *Ecology* **92(9)**:1758–1768

11. Livingston G, Jiang Y, Fox JW and Leibold MA (2013). The dynamics of community assembly under sudden mixing in experimental microcosms. *Ecology* **94(12)**:2898–2906

12. O'Loughlin LS and Green PT (2017). Secondary invasion: When invasion success is contingent on other invaders altering the properties of recipient ecosystems. *Ecology and Evolution* **7(19)**:7628–7637

13. Castledine M, Sierocinski P, Padfield D and Buckling A (2020). Community coalescence: an eco-evolutionary perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences* **375(1798)**:20190252

14. Simberloff D and Holle BV (1999). Positive Interactions of Nonindigenous Species: Invasional Meltdown? *Biological Invasions* **1(1)**:21–32

15. Simberloff D (2006). Invasional meltdown 6 years later: important phenomenon, unfortunate metaphor, or both? *Ecology Letters* **9(8)**:912–919

16. Gurevitch J (2006). Commentary on Simberloff (2006): Meltdowns, snowballs and positive feedbacks. *Ecology Letters* **9(8)**:919–921

17. Toquenaga Y (1997). Historicity of a Simple Competition Model. *Journal of Theoretical Biology* **187(2)**:175–181

18. Tikhonov M (2016). Community-level cohesion without cooperation. *eLife* **5**:e15747

19. Tikhonov M and Monasson R (2017). Collective Phase in Resource Competition in a Highly Diverse Ecosystem. *Physical Review Letters* **118(4)**:048103

20. Vila JCC, Jones ML, Patel M, Bell T and Rosindell J (2019). Uncovering the rules of microbial community invasions. *Nature Ecology & Evolution* **3(8)**:1162–1171

21. Lechón P, Clegg T, Cook J, Smith TP and Pawar S (2021). The role of competition versus cooperation in microbial community coalescence. *biorXiv*

22. Sierocinski P, Milferstedt K, Bayer F, Großkopf T, Alston M, Bastkowski S, Swarbreck D, Hobbs PJ, Soyer OS et al. (2017). A Single Community Dominates Structure and Function of a Mixture of Multiple Methanogenic Communities. *Current Biology* **27(21)**:3390–3395.e4

23. Rillig MC and Mansour I (2017). Microbial Ecology: Community Coalescence Stirs Things Up. *Current Biology* **27(23)**:R1280–R1282

24. Louca S, Jacques SMS, Pires APF, Leal JS, Srivastava DS, Parfrey LW, Farjalla VF and Doebeli M (2016). High taxonomic variability despite stable functional structure across microbial communities. *Nature Ecology & Evolution* **1(1)**:0015

25. Winfree R, Fox JW, Williams NM, Reilly JR and Cariveau DP (2015). Abundance of common species, not species richness, drives delivery of a real-world ecosystem service. *Ecology Letters* **18(7)**:626–635

26. Rosenzweig RF, Sharp RR, Treves DS and Adams J (1994). Microbial evolution in a simple unstructured environment: genetic differentiation in Escherichia coli. *Genetics* **137(4)**:903–917

27. Goldford JE, Lu N, Bajić D, Estrela S, Tikhonov M, Sanchez-Gorostiaga A, Segrè D, Mehta P and Sanchez A (2018). Emergent simplicity in microbial community assembly. *Science* **361(6401)**:469–474

28. Marsland III R, Cui W, Goldford J, Sanchez A, Korolev K and Mehta P (2019). Available energy fluxes drive a transition in the diversity, stability, and functional structure of microbial communities. *PLoS Computational Biology* **15(2)**:e1006793

29. Marsland R, Cui W, Goldford J and Mehta P (2020). The Community Simulator: A Python package for microbial ecology. *PLoS ONE* **15(3)**:e0230430

30. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA and Holmes SP (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* **13(7)**:581–583

31. Callahan BJ, McMurdie PJ and Holmes SP (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal* **11**:2639–2643

32. Kozich JJ, Westcott SL, Baxter NT, Highlander SK and Schloss PD (2013). Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Applied and Environmental Microbiology* **79(17)**:5112–5120

33. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**:335–336

34. Wang Q, Garrity GM, Tiedje JM and Cole JR (2007). Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology* **73(16)**:5261–5267

35. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J and Glöckner FO (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41(D1)**:D590–D596

36. Curtis JT and Bray JR (1957). An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* **27(4)**:325–349

37. Lin J (1991). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory* **37(1)**:145–151

38. Kullback S and Leibler RA (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics* **22(1)**:79–86

39. Jaccard P (1912). The distribution of the flora in the alpine zone. *New Phytologist* **11(2)**:37–50

40. Yang T, Wei Z, Friman VP, Xu Y, Shen Q, Kowalchuk GA and Jousset A (2017). Resource availability modulates biodiversity-invasion relationships by altering competitive interactions. *Environmental Microbiology* **19(8)**:2984–2991

41. Mallon CA, Poly F, Roux XL, Marring I, van Elsas JD and Salles JF (2015). Resource pulses can alleviate the biodiversity–invasion relationship in soil microbial communities. *Ecology* **96(4)**:915–926

Figures

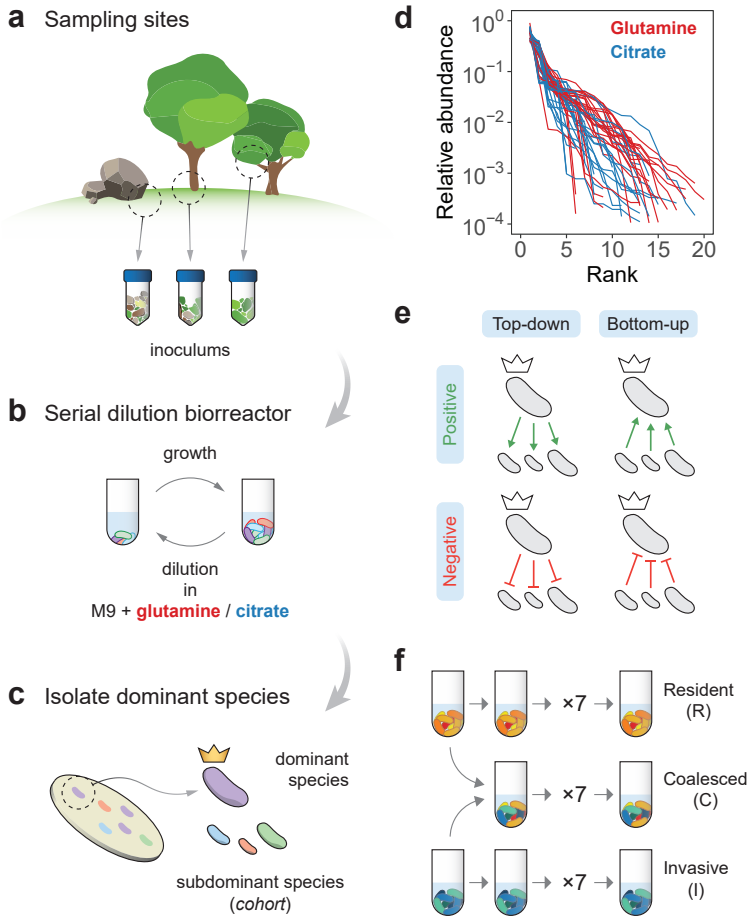

Figure 1. Overview of the experimental protocol. a. Environmental samples collected from eight different locations were used to inoculate our communities. **b.** Communities were stabilized in serial batch culture bioreactors [27] in minimal synthetic media with glutamine or citrate as the only supplied carbon source. **c.** Communities were plated in minimal media agar plates and the most abundant species (the "dominants") from each community were isolated. We refer to the set of sub-dominant species as the "cohorts". **d.** Rank-frequency distributions of all eight communities stabilized in either glutamine (red) or citrate (blue), sequenced at a depth of 10^{-4} reads. Three biological replicates per community are shown. Community compositions are skewed and long-tailed. **e.** Our hypothesis is that ecological co-selection can take place from the top-down, i.e. the dominant co-selecting the cohort, or from the bottom-up, i.e. the cohort co-selecting the dominant. Both forms of co-selection can be positive (recruitment) or negative (antagonism). **f.** Illustration of the protocol of our coalescence experiments. All pairs of communities were inoculated into fresh minimal media supplemented with the same carbon source where communities had been previously stabilized. The coalesced (C) and original resident (R) and invasive (I) communities were then serially diluted and allowed to grow for seven additional transfers.

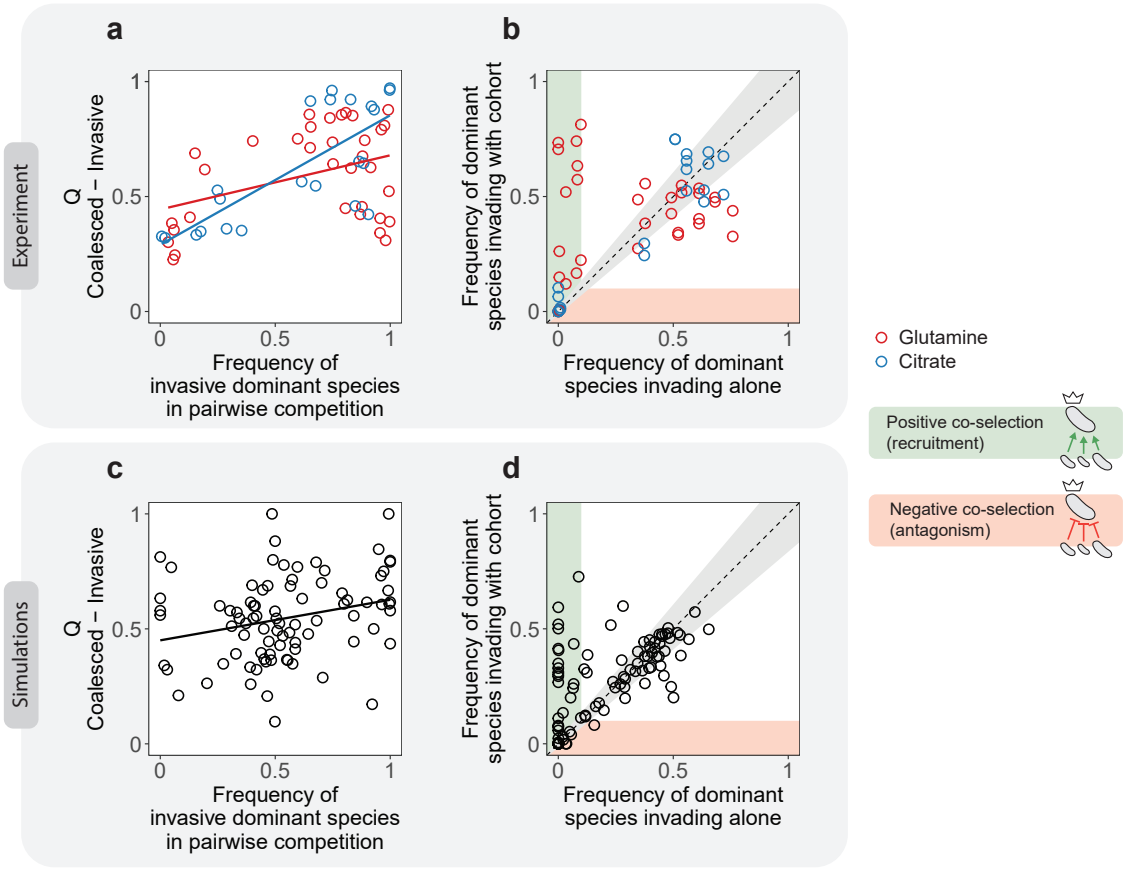

Figure 2. Co-selection in microbial community coalescence. a. Coalescence outcomes are quantified by the relative Bray-Curtis similarity (Q) between the coalesced and invasive communities (Methods: Metrics of community distance). These outcomes are predicted by the pairwise competition between the invasive and resident dominant species (R^2 = 0.15 for glutamine and R^2 = 0.57 for citrate). This is consistent with a scenario of top-down positive co-selection where dominants recruit their cohorts for the final coalesced community. Two biological replicates per experiment are plotted individually. **b.** We represent the frequency reached by the invasive dominant species when they invade the resident communities on their own versus when they are in the company of their cohort. Three scenarios are possible: green and red shaded areas represent limit cases of positive (recruitment) or negative (antagonism) bottom-up co-selection, gray area corresponds to situations where invasive dominant species can invade with equal success regardless of the presence of their cohorts. Data shows that positive co-selection is common, whereas antagonistic co-selection is rare in our experiments. Two biological replicates per experiment are plotted individually. **c-d.** Simulations of community coalescence with a consumer-resource model are able to capture these trends.

Supplementary Figures

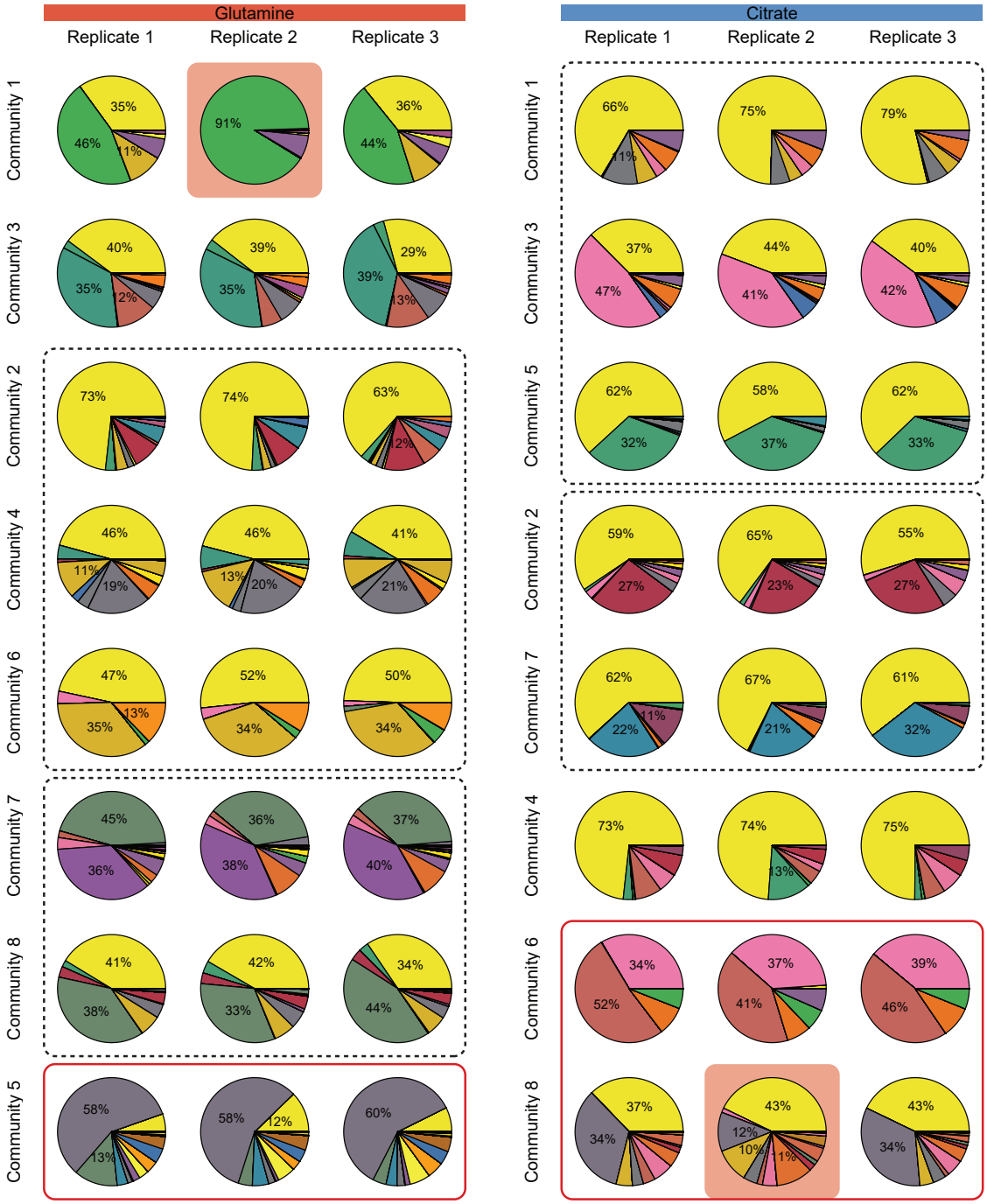

Figure S1. Community compositions after seven additional transfers without coalescence. Each color of the pie plots corresponds to a different exact sequence variant (Methods: Determination of community composition by 16S sequencing). Replicate 2 of community 1 from glutamine, as well as replicate 2 of community 8 from citrate (highlighted) were removed based on their dissimilarity to the other two replicates (details in code for data analysis, see Data & code availability). Communities clustered in dashed boxes shared the same dominant species as revealed by sequencing data. For communities enclosed in red boxes, sequencing data showed that the species isolated by plating was not detectable in the community after seven additional transfers (i.e. the dominant was incorrectly identified) and were therefore excluded from downstream analyses.

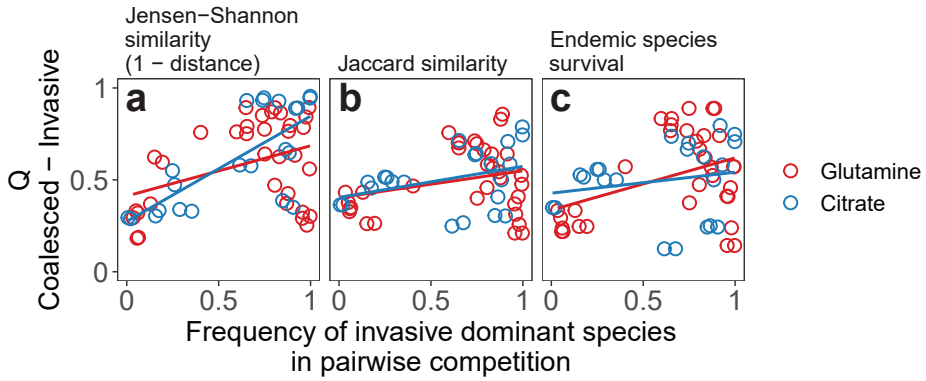

Figure S2. Alternative metrics of community distance. Quantifying coalescence outcomes using different metrics of community similarity (Methods: Metrics of community distance) gives similar results to those shown in Figure 2a. Metrics that account for the relative species abundances (**a.** Jensen-Shannon similarity or Figure 2a Bray-Curtis similarity) yield higher correlations than less quantitative metrics that only account for species presence/absence (**b.** Jaccard similarity or **c.** Fraction of endemic invasive species persisting in the coalesced community).

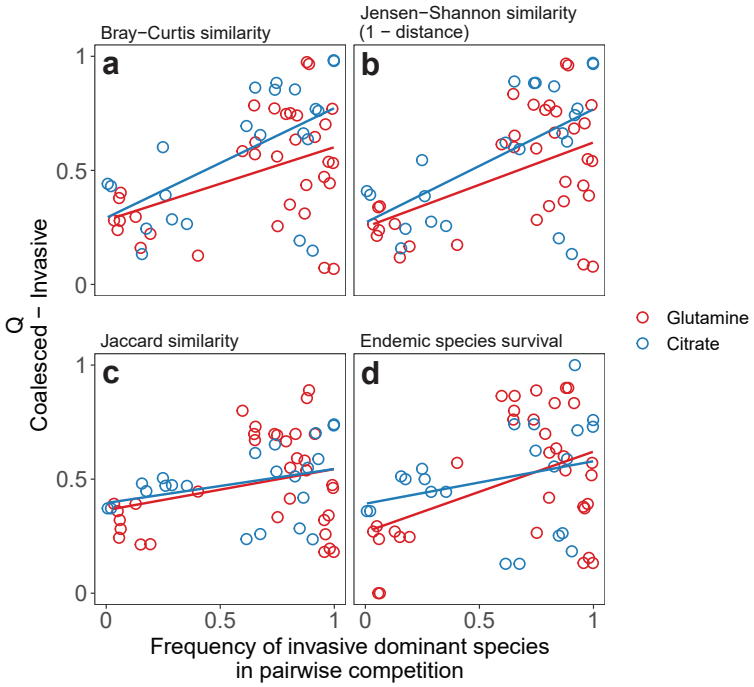

Figure S3. Dominant species have limited effects on coalescence outcomes. We repeated the analyses shown in Figure 2a and Figure S2, but this time we removed the dominants from the compositional data prior to quantifying community distances. The trends observed before are maintained.

Test section

This is an example cite [2, 3]. This is how you refer to Figure 1. This is how you refer to Figure 1a. This is how you refer to Table 1. This is how you refer to Figure S1. This is how you refer to the section Simulations of the Methods. This is a cleaner way to refer to the Methods: Simulations.

Check this reference [13]

Resource availability modulates biodiversity–invasion relationships by altering competitive interactions [40]

Resource pulses can alleviate the biodiversity–invasion relationship in soil microbial communities [41]