nature ecology & evolution

Article

https://doi.org/10.1038/s41559-024-02440-6

Metabolic complexity drives divergence in microbial communities

Received: 4 August 2023

Accepted: 14 May 2024

Published online: 02 July 2024

Check for updates

Michael R. Silverstein [1,2], Jennifer M. Bhatnagar [1,3] & Daniel Segrè [1,2,3,4]

Microbial communities are shaped by environmental metabolites, but the principles that govern whether different communities will converge or diverge in any given condition remain unknown, posing fundamental questions about the feasibility of microbiome engineering. Here we studied the longitudinal assembly dynamics of a set of natural microbial communities grown in laboratory conditions of increasing metabolic complexity. We found that different microbial communities tend to become similar to each other when grown in metabolically simple conditions, but they diverge in composition as the metabolic complexity of the environment increases, a phenomenon we refer to as the divergence-complexity effect. A comparative analysis of these communities revealed that this divergence is driven by community diversity and by the assortment of specialist taxa capable of degrading complex metabolites. An ecological model of community dynamics indicates that the hierarchical structure of metabolism itself, where complex molecules are enzymatically degraded into progressively simpler ones that then participate in cross-feeding between community members, is necessary and sufficient to recapitulate our experimental observations. In addition to helping understand the role of the environment in community assembly, the divergence-complexity effect can provide insight into which environments support multiple community states, enabling the search for desired ecosystem functions towards microbiome engineering applications.

Understanding how diverse microbial communities assemble is important for addressing open challenges in microbial ecology with applications that range from medicine[1,2] to climate change mitigation[3–5]. Studies in natural[6,7] and laboratory[8–15] settings have investigated the reproducibility of assembly dynamics across a range of environmental conditions leading to seemingly contradictory results. Under certain conditions, microbial community assembly appears to be highly deterministic, as different communities are driven by strong environmental selection towards a specific steady state independent of their initial composition[8]. Under other conditions, however, environmental selection is weaker, resulting in highly variable assembly of communities with more dependence on their initial composition[11]. Uncovering what properties govern this variability in microbial community assembly constitutes one of the fundamental questions of microbial ecology[16] and is crucial for successful microbiome engineering, which aims to steer communities towards a desired structure in a given environment[17].

Previous studies have explored how one community behaves across a range of conditions[9,10,18] or how communities behave under a single condition[11–15]. However, systematically comparing the fates of multiple communities grown in multiple conditions remains an underexplored avenue for understanding microbial

[1]Bioinformatics Program, Faculty of Computing and Data Science, Boston University, Boston, MA, USA. [2]Biological Design Center, Boston University, Boston, MA, USA. [3]Department of Biology, Boston University, Boston, MA, USA. [4]Department of Biomedical Engineering and Department of Physics, Boston University, Boston, MA, USA. e-mail: dsegre@bu.edu

https://doi.org/10.1038/s41559-024-02440-6

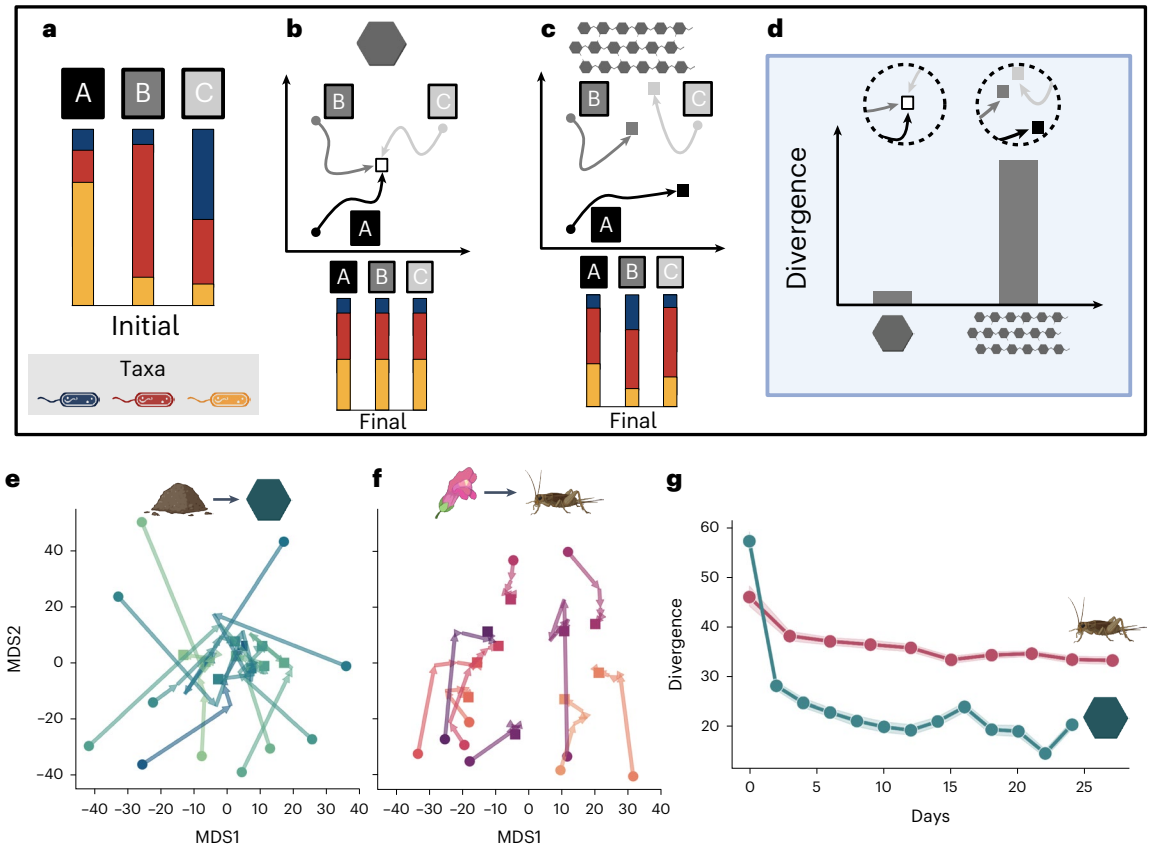


Fig. 1 | Microbial communities may diverge in environments with increasing metabolic complexity. a–**d**, Hypothesis of microbial community divergence in theoretical simple (**b**) and complex (**c**) metabolic conditions. Microbial communities A, B and C are initially composed of different compositions of the same three microbial species (**a**; blue, red and yellow). Over time, communities grown on a simple substrate (**b**) converge, while these same communities grown on a complex substrate (**c**) diverge. The lines in **b** and **c** show the trajectory of each community from the initial composition (circles) to final compositions (squares). **d**, Quantification of divergence (distance between communities in the same condition) at the final timepoint (trajectories arriving at squares shown in dashed circles above each bar for each condition) for hypothetical scenarios in **a**–**c**. **e**–**g**, Divergence observed in two independent experimental studies, one where microbial communities were sourced from soils or leaves and grown on glucose (**e**; a relatively simple metabolic environment from Goldford et al., $N = 11$ communities) and another where communities were sourced from pitcher plants and grown on acidified cricket media (**f**; a more complex metabolic environment from Bittleston et al., $N = 10$ communities). Each coloured line in **e** and **f** represents the trajectory of a community's composition over time in separately computed MDS projections. The circles indicate the initial community composition, and the squares indicate the final community composition. **g**, The divergence for each metabolic environment, calculated as the pairwise distances between all communities within a given condition at each timepoint. Each point is the mean pairwise distance within condition at each timepoint, and shading represents the 95% confidence interval over all pairwise distances within each environment at each timepoint.

ecosystem dynamics. Culturing multiple communities in a given condition is necessary for understanding how closely those communities converge or diverge taxonomically from each other over time. Conversely, exploring these dynamics across many conditions is necessary for understanding how divergence depends on properties of the environment.

In this Article, we combined experimental measurements and computational modelling to investigate the interplay of initial composition and environmental selection in determining the extent to which the same set of distinct initial communities tend to resemble each other within environments that span a broad range of metabolic complexity. We followed the dynamic assembly of diverse microbial communities inoculated from different soil samples grown on carbon sources of increasing metabolic complexity. We then measured the divergence (beta-diversity) between communities within each condition. Since metabolic complexity could be measured by the potential for a single substrate to give rise to new byproducts and the number of different substrates present in the environment[18], we investigated single- and mixed-metabolite conditions, separately. While other studies have explored the relationship between metabolic complexity and (alpha) diversity within a community[9,18], we now investigate how metabolic complexity relates to the divergence between the same set of communities (beta-diversity).

By tracking how closely the taxonomic structures of these communities resembled each other over time, we found that the effect of environmental selection on communities depended on the metabolic complexity of the environment itself. Specifically, different microbial communities diverged in their taxonomic composition across a gradient of increasingly complex metabolic conditions, suggesting that the forces dominating microbial community assembly shift from strong to weak environmental selection in increasingly complex conditions. We tested whether a commonly used ecological model for microbial ecosystem dynamics[8,19–21] could reproduce this effect and found that this occurs only upon incorporating a representation of the hierarchical structure of metabolite transformations (for example, polysaccharides to oligosaccharides to monosaccharides)[14,22] and does not depend on how metabolite preferences are distributed across taxa. Our results point to an ecosystem organization principle that can help reconcile seemingly incompatible observations of divergence in different conditions and provide guidelines for which environments may be more conducive to microbiome engineering.

https://doi.org/10.1038/s41559-024-02440-6

Results

The divergence-complexity effect hypothesis

To assess the strength of environmental selection on community assembly, one would ideally compare how the trajectories of multiple distinct microbial communities diverge in taxonomic composition across a set of conditions. A key question we ask is whether distinct communities assembled in the same condition tend to become taxonomically similar and how the degree of similarity depends on the metabolic complexity of the environment. For example, different microbial communities that initially vary in taxonomic composition (Fig. 1a) may converge in composition over time when grown in one environment (strong environmental selection; Fig. 1b), while those same communities may diverge in another environment, arriving at alternative stable states (weak environmental selection; Fig. 1c). To quantify the degree to which different communities differ taxonomically from each other when grown in a given condition, we calculate the divergence in community compositions (i.e. beta-diversity within each condition; Fig. 1d; Methods and 'Measuring divergence with the Aitchison distance metric' section in Supplementary Information).

We initially identified existing data that could indicate whether and how community divergence would depend on environmental conditions. Specifically, we re-analysed two independent studies that explored how a collection of diverse microbial communities assembled over time, but did so under very different conditions. When one study, Goldford et al.[8], cultured communities in (simple) glucose media, communities converged (Fig. 1e). By contrast, when Bittleston et al.[11] cultured communities in (complex) acidified cricket media, they diverged (Fig. 1f). In both cases, the initial communities differed substantially from each other and then immediately became more similar; however, communities enriched on glucose ultimately converged significantly more, despite starting with greater variation in initial community composition (Fig. 1g). Based on the striking discrepancy in the degree of divergence across these two studies, we formulated the hypothesis that divergence increases with the metabolic complexity of the provided resources (Fig. 1d), a relationship that we will refer to as the divergence-complexity effect.

Community divergence increases with metabolic complexity

To directly test the divergence-complexity effect, we designed an experiment to quantify the divergence of microbial communities grown in conditions of increasing metabolic complexity (Methods and Fig. 2a). To assess divergence, we sourced six microbial communities from forest soils, which are generally diverse and distinct from each other[23], even over small (centimetre) spatial scales[7]. Each microbial community was grown in nine different minimal media, each supplemented with equimolar carbon concentrations of at least one carbon source commonly found in soils[24]: (1) citrate, (2) glucose, (3) cellobiose, (4) cellulose, (5) lignin, (6) citrate + glucose, (7) citrate + glucose + cellobiose, (8) citrate + glucose + cellobiose + cellulose or (9) citrate + glucose + cellobiose + cellulose + lignin (Methods).

In testing the divergence-complexity effect, we consider metabolic complexity to increase from citrate to lignin (in line with the number of metabolic byproducts expected from each metabolite[18]), enabling us to define a clear gradient of complexity separately for single metabolites, from condition 1 to 5, and mixed metabolites from condition 6 to 9. Lignin can be depolymerized extracellularly by a variety of enzymes into a diversity of polyphenolic compounds, which can be broken down into aromatic compounds, and eventually cleaved into respirable or fermentable byproducts[25–29]. Cellulose, a polysaccharide composed of repeating glucose subunits, can be degraded by cellulases[30] into cellobiose (a disaccharide of a pair of glucose molecules) and glucose itself[31]. Citrate[32], the simplest carbon source in our experiment, is an organic acid related to each of the metabolites in our study as an intermediate of respiration. Given that our culture conditions were probably not fully aerated, the above metabolites are expected to ultimately be used by the communities through a mixture of aerobic and anaerobic metabolism, not unlike what is experienced by soil microbial communities[33–36].

Each microcosm, containing one source community growing in one condition, was serially passaged ten times, in intervals of 3 days. 16S ribosomal RNA sequencing was performed and amplicon sequence variant (ASV) counts were generated for the initial soil inocula and for microcosm communities at days 3, 6, 9, 12 and 33 (excluding 48/384 samples; Extended Data Table 1 and Methods). The divergence for each condition was then calculated by computing the distance between all pairs of communities within each condition for each day (Methods). For increased accuracy at our final timepoint, the divergence for day 33 communities was computed by averaging the distances between three replicates for each community (excluding one community, where only two replicates were available; Extended Data Table 1).

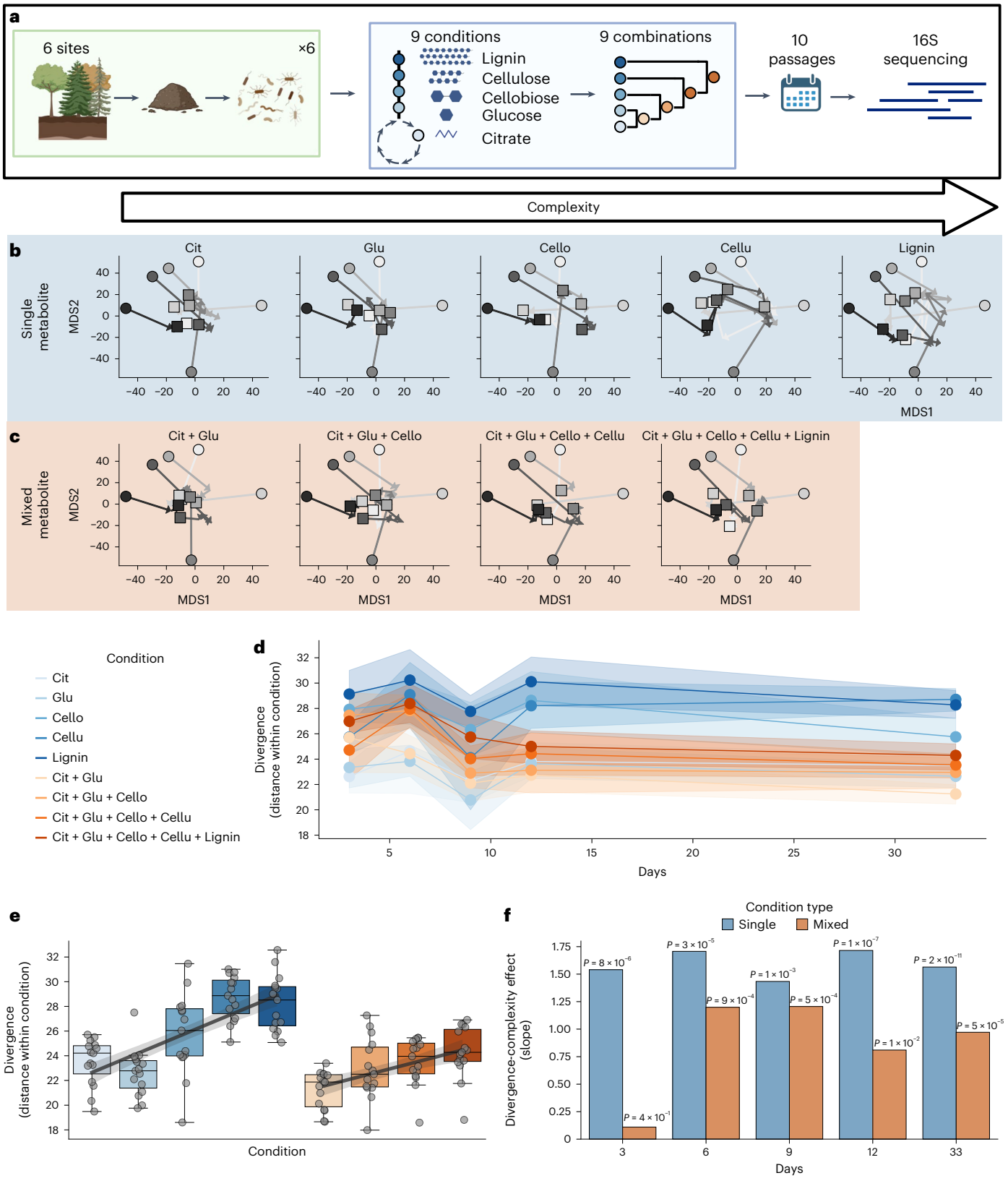
Supporting our hypothesis of the divergence-complexity effect, we observed that divergence increased for both of our axes of metabolic complexity (single and mixed conditions; Fig. 2b–f). In accordance with previous studies (Fig. 1e,f), our source communities started differently from each other and initially converged (Fig. 2b,c) and stabilized once introduced to laboratory conditions (Fig. 2d and Extended Data Fig. 1). Within single-metabolite conditions, the communities converged strongly on simple metabolites, while they diverged to increasingly distinct states on the more complex metabolites (Fig. 2b,d–f). At the final timepoint, divergence occurred in a discrete fashion with similarly low levels between citrate and glucose, similarly high levels between cellulose and lignin, and an intermediate level of divergence on cellobiose (Fig. 2e), which may be related to where each metabolite is enzymatically processed (citrate and glucose: intracellularly, cellulose and lignin: extracellularly, and cellobiose: both ways[37]). Similar to single-metabolite conditions, community divergence increased from the least (citrate + glucose) to the most diverse (all metabolites) mixed-metabolite conditions (Fig. 2c–f). For single- and mixed-metabolite conditions, the divergence-complexity effect

Fig. 2 | Divergence of microbial communities increases in environments of increasing metabolic complexity. a, Study design: microbial communities were extracted from six forest soils and were then grown in nine conditions (citrate, glucose, cellobiose, cellulose, lignin, citrate + glucose, citrate + glucose + cellobiose, citrate + glucose + cellobiose + cellulose and citrate + glucose + cellobiose + cellulose + lignin). Communities were passaged ten times once every 3 days and sequenced on days 0, 3, 6, 9, 12 and 33 (*N* varies for each community; Extended Data Table 1). **b**,**c**, MDS projections of community trajectories over time in each single-metabolite condition (**b**) and mixed-metabolite condition (**c**). MDS was calculated on all samples together for ease of visually comparing trajectories between conditions. The circles indicate the initial community composition, and the squares indicate the final community composition. **d**, Divergence of communities within each condition over time from day 3 onwards. Initial communities are a distance of 58.1 ± 3.5 (not shown for clarity). Single-metabolite conditions are in blue, mixed conditions are in orange and colours darken with complexity. The points on each line represent the mean divergence, and the shaded region represents the 95% confidence interval for pairwise distances between all six communities within each condition. **e**, Distribution of divergence between all six communities at the final timepoint where divergence increases with metabolic complexity for single- and mixed-metabolite conditions (same colours as **d**). Each point is a pairwise distance between two communities (*N* = 15 per condition). The boxes are bound by the interquartile range, divided by the median, and whiskers extend to a maximum of 1.5 times the interquartile range. **f**, Divergence-complexity effect by condition type (slopes shown in **e**) for all timepoints (one-sided Wald test on effect (slope) >0, *P* values shown in figure for each day and condition type). All divergence results for communities at day 33 are the mean distances between up to three replicates for each community (see list of all samples and conditions in Extended Data Table 1). Cit, citrate; Glu, glucose; Cello, cellobiose; Cellu, cellulose.

https://doi.org/10.1038/s41559-024-02440-6

was detectable at each sampled timepoint (Fig. 2f and Extended Data Fig. 2), suggesting that the effect occurred rapidly (as early as day 3) and was sustained irrespective of possible evolutionary changes. While evolutionary adaptation to utilizing new nutrients tends to occur at much longer timescales[38–40], we cannot rule out the role of strain evolution in our longitudinal analysis[41].

Interestingly, the effect was stronger in single-metabolite conditions than in mixed-metabolite conditions (Fig. 2e,f). One possible explanation for this trend includes priority effects[42], where assembly dynamics are dependent on the order in which different metabolites become available through trophic interactions[10]. Another possibility is that divergence was sensitive to the concentration of complex



https://doi.org/10.1038/s41559-024-02440-6

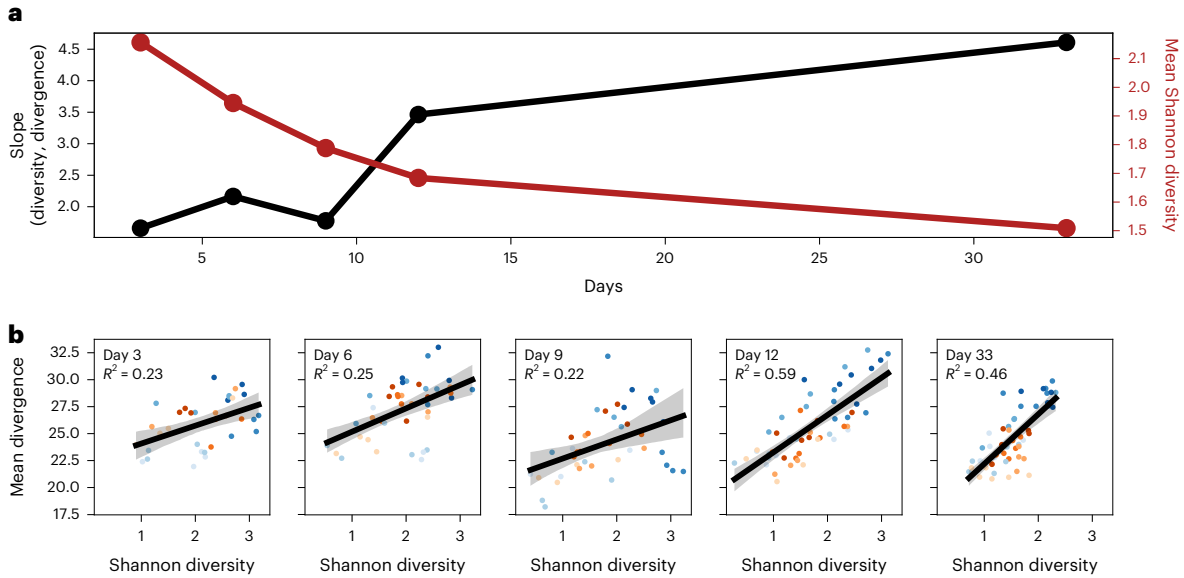


Fig. 3 | Divergence dynamically correlates with diversity. a, The slope of the relationship between community alpha-diversity and divergence (black) and the mean community alpha-diversity (red) over time. **b**, The data underlying the relationship in **a** over time. Each point is the diversity of a community in a condition (*x* axis; same colours per condition as Fig. 2) and the mean divergence of that community from all others within a condition (*y* axis). All divergence results for communities at day 33 are the mean distances between up to three replicates for each community (Extended Data Table 1). The coefficient of determination (R^2) for each day is reported from linear regression of mean divergence on diversity. The shaded areas around each regression line represent the 95% confidence interval. While diversity is expected to increase with metabolic complexity, it is not clear if different communities will increase in diversity in the same ways (with coordinated changes to the same taxa). See 'Divergence is sensitive to richness and evenness' section in Supplementary Information for more details about the relationship between divergence and diversity.

metabolites[43]. Since the total carbon concentration was kept constant across all conditions, the concentration of complex metabolites was highest in single-metabolite conditions and decreased incrementally for each subsequently complex mixed-metabolite condition (Methods).

Since it is common for many taxa to engage in functionally redundant actitivies[44,45] and since our data are generated from amplicon sequencing, we cannot conclude whether these taxonomically divergent communities are utilizing similarly divergent metabolic pathways[46]. However, when we recomputed divergence at the family taxonomic level (Extended Data Figs. 3 and 4) where bacteria often differ in metabolic functions[15], we still observed the divergence-complexity effect, which suggests that our communities that assembled to distinct taxonomic compositions may have also engaged in distinct metabolic activities.

Divergence, which was high across different source communities in complex conditions, was significantly lower across replicates from the same source community. Replicate microcosms assembled from the same source converge to similar final states after 33 days (Extended Data Fig. 5). Conversely, communities sourced from different locations, even if they displayed similar initial composition, did not necessarily converge (Fig. 2b,c). This suggests that, while assembly dynamics are indeed complex (initially similar communities can diverge), they are reproducible (replicate microcosms converge). While previous studies have reported that stochasticity can be responsible for multistability, this similarity across replicate experiments suggests that stochasticity is not the main driver of divergence in our communities[15]. Our observations are consistent with simulation work that suggests that multistability can take place without stochastic fluctuations in environments containing multiple essential nutrients[47]. Ultimately, the community assembly trajectories we observed in our experiment show that complex environments can support a greater number of discrete states than simple environments but that these discrete environments may be achievable only from distinct initial conditions.

Diversity correlates with divergence, implicating specialists

To gain a deeper understanding of the divergence-complexity effect, we investigated how alpha-diversity within each individual community correlates with divergence across communities. In particular, two separate principles could jointly give rise to the divergence-complexity effect. The first principle, 'metabolic complexity begets diversity', where community diversity increases with increasing metabolic complexity, has been experimentally documented in both natural and synthetic communities[9,18]. A yet unexplored second principle that we investigate in our study, 'diversity begets divergence', could result from the expectation that more diverse communities generally diverge more from one another (see 'Divergence is sensitive to richness and evenness' section in Supplementary Information; Fig. 4d). In general, the amount of divergence we observe is greater than expected from richness alone, suggesting that differing distribution of taxa (skewness) between communities substantially contributes to the divergence-complexity effect (Supplementary Fig. 1). If metabolic complexity were to yield diversity and diversity were to yield divergence, we would expect higher divergence in increasingly complex conditions, contributing to the divergence-complexity effect.

Consistent with these expectations, we observed a strong linear relationship between diversity and divergence, which strengthened over time, indicating that specific changes in community assembly drive the rise of divergence. The slope of the diversity–divergence relationship increased over time (Fig. 3a, black line, Fig. 3b and Extended Data Fig. 6), despite the fact that diversity itself, on average, decreased (Fig. 3a, red line). In other words, over time, the same degree of divergence is maintained by communities with reduced diversity. For divergence to remain relatively stable (Fig. 2d) while diversity decreases, taxa endemic (that is, specific) to each community must persist while a set of species shared across communities universally go extinct within each condition. One possible explanation for this trend is that these persistent taxa are metabolic specialists, which produce enzymes that target specific biochemical bonds[48]. In conjunction with our previous

https://doi.org/10.1038/s41559-024-02440-6

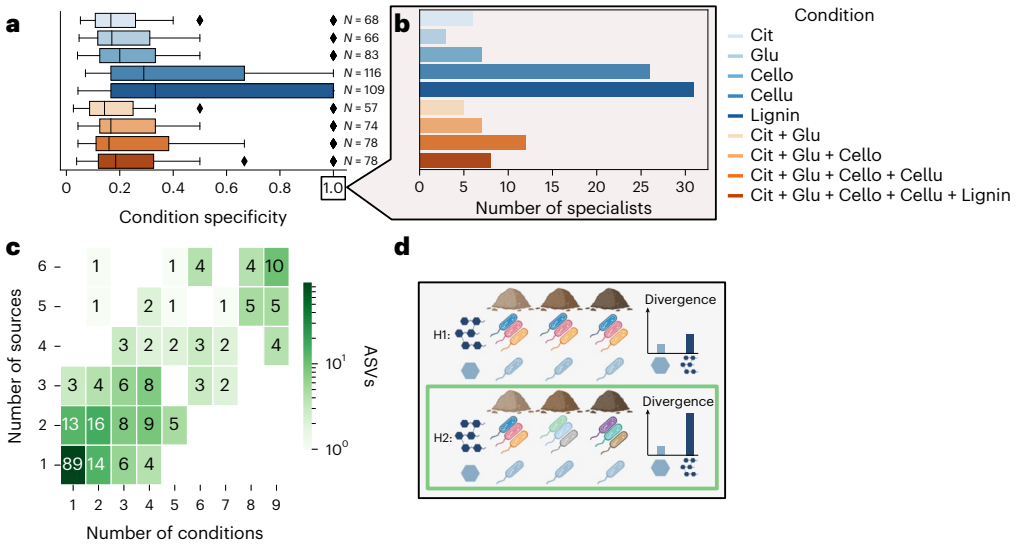


Fig. 4 | Endemic taxa are enriched and unevenly distributed in complex conditions. a, The distribution of taxon condition specificity per condition for day 33. The number of taxa (*N*) for each condition is indicated in the figure. Condition specificity is calculated as the fraction of occurrences of a taxon that is attributed to a particular condition, such that a specificity of 1 means that taxon occurs in only one condition (a specialist). The boxes are bound by the interquartile range, divided by the median, and whiskers extend to a maximum of 1.5 times the interquartile range. Diamond symbols indicate outliers passed 1.5× beyond the interquartile range. **b**, The number of specialists per condition (taxa with condition specificity = 1). **c**, Taxon occurrence by number of conditions and number of source communities. ASVs found in fewer conditions are less evenly distributed across source communities (found in fewer source communities), and taxa found in more conditions are more evenly distributed across source communities. **d**, Two hypotheses for single-metabolite conditions following from **a** and **b**, where H2 is supported and H1 is not. H1: more complex conditions are enriched for specialists, and when those taxa are evenly distributed across source communities, it results in relatively similar divergence for complex and simple metabolic conditions. H2: when more complex conditions are enriched for specialists and these taxa are less evenly distributed across communities, more complex conditions result in greater divergence. Cit, citrate; Glu, glucose; Cello, cellobiose; Cellu, cellulose.

observations that communities diverge more in complex conditions, we hypothesized that specialists that target complex metabolites were less evenly distributed across communities than taxa that specialize on simpler metabolites, driving the divergence-complexity effect.

Specialists are more endemic in complex conditions

To investigate the role of specialists in the divergence-complexity effect, we explored the distribution of taxa across experimental conditions and source communities. If specialists drive the diversity-complexity effect, we would expect to see that specialists are increasingly endemic, or unevenly distributed across source communities, in more complex conditions.

To quantify the degree of specialization, we computed a condition-specificity metric for each taxon (ASV) in each condition and then assessed whether specialization and endemism depended on metabolic complexity. We defined condition specificity for each taxon and condition as the fraction of communities in which that taxon was found at the final sampling timepoint for that given condition. In particular, if a taxon occurs in only one condition, its condition specificity is 1 and will be referred to as a 'specialist'. In accordance with our expectations, we observed that more complex conditions (particularly single-metabolite ones) had greater condition specificity (Fig. 4a) and more specialists (Fig. 4b). Given the broad phylogenetic distribution of traits associated with degrading and utilizing the byproducts of cellulose[49] and lignin[25–27], these condition-specific taxa may be performing redundant metabolic functions in their respective communities.

These results alone are encouraging but are not sufficient for linking specialists to the divergence-complexity effect, which would additionally require specialists to differ between communities in the same condition. While we observed an enrichment of condition-specific taxa in complex conditions, it hypothetically could be the case that these same taxa were found across all source communities, in which case communities in complex conditions would not necessarily diverge substantially more than those in simple conditions (Fig. 4d, H1 and Supplementary Fig. 1). However, when we count the occurrence of each taxon in each condition and source community, we find that condition-specific taxa are also source community specific (endemic; Fig. 4c). As a result, taxa that specialize on complex metabolites are less evenly distributed across communities than taxa that specialize on simpler metabolites and are therefore heavily implicated in mediating the divergence-complexity effect (Fig. 4d, H2 and Supplementary Fig. 1).

Trophic resource transformations reproduce divergence

To better understand what properties of microorganisms and their environment are necessary for the diversity-complexity effect, we performed a series of simulations with microbial consumer-resource models (CRMs; Methods)[19]. In particular, we wanted to corroborate our hypothesized mechanism of the diversity-complexity effect: that divergence correlates with metabolic complexity and emerges from endemism of specialist taxa.

CRMs are dynamical ecological models where consumers are defined by the set of resources they prefer (consumer preferences) and resources are transformed by consumers into other resources following consumption (resource transformations). Overlapping consumer preferences give rise to competition, while the exchange of transformed products can generate cross-feeding interactions[19]. Cross-feeding can take many different shapes in natural communities, including adaptive cooperation among auxotrophs[50], secretion of central carbon metabolism intermediates[51–56] or fermentation products[57], extracellular degradation of complex macromolecules whose byproducts are shared with community members[58], or stress-induced cross-feeding[59]. While cross-feeding in CRMs is often assumed to reflect exudation of intracellular intermediates[19], CRMs can still be interpreted as approximating a broader set of mechanisms, including those mentioned above (see 'Generality of cross-feeding mechanism in our consumer-resource model' section in Supplementary Information).

https://doi.org/10.1038/s41559-024-02440-6

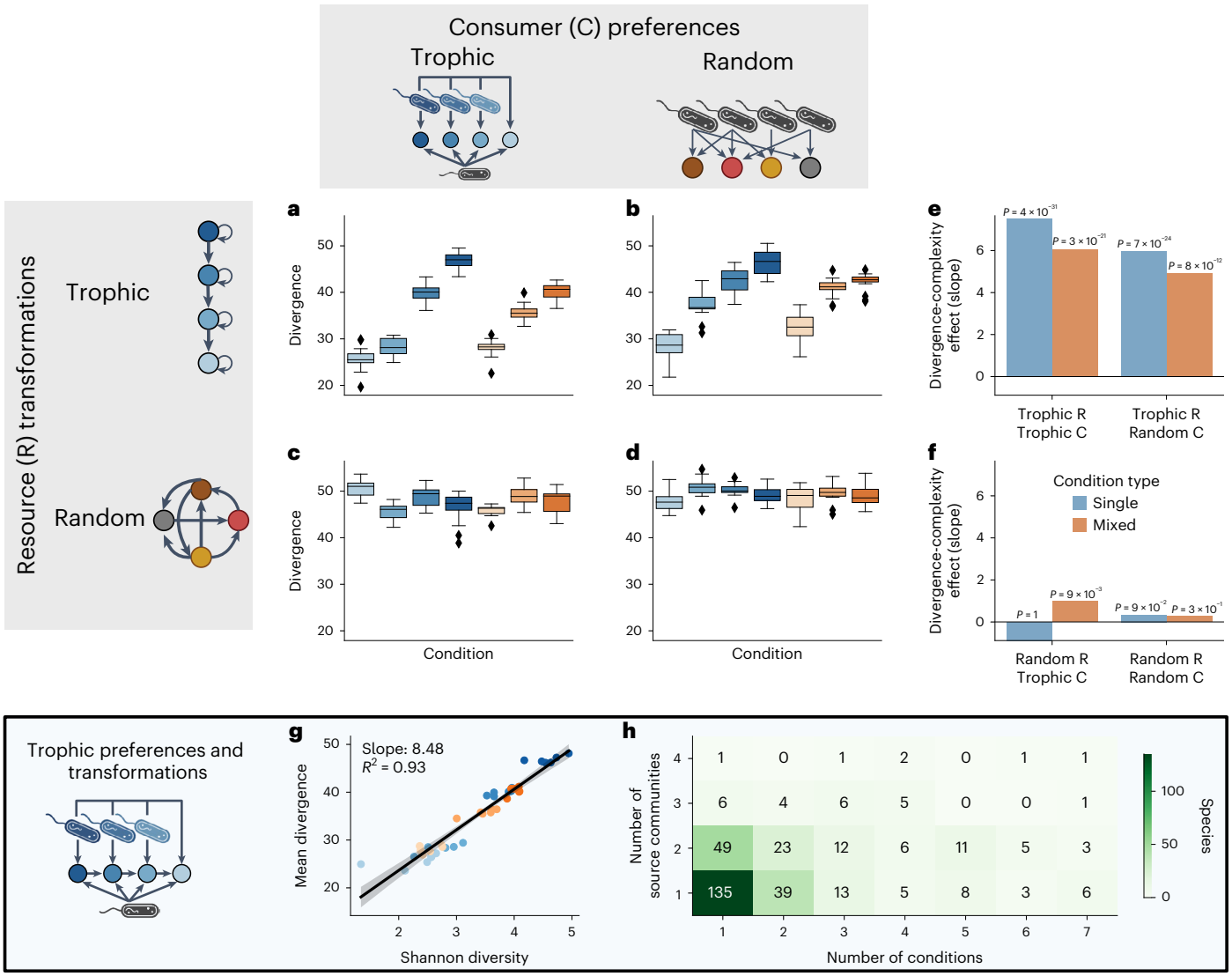


Fig. 5 | Trophic resource transformations reproduce divergence-complexity effect with CRM simulations. The distribution of divergence for six communities simulated with CRMs with and without trophic structure in resource transformations and consumer preferences. Mimicking our experiment, community growth was simulated in single-metabolite (blue) and mixed-metabolite (orange) conditions of increasing complexity (from lighter to darker). **a**–**d**, Divergence for communities simulated with trophic resource transformations and consumer preferences (fully structured; **a**), trophic resource transformations and random consumer preferences (resource structured; **b**), random resource transformations and trophic consumer preferences (consumer structured; **c**) and random resource transformations and consumer preferences (fully random; **d**). Each box describes the distribution of pairwise distances between all six simulated communities ($N = 15$ per box). The boxes are bound by the interquartile range, divided by the median, and whiskers extend to a maximum of 1.5 times the interquartile range. Diamond symbols indicate outliers passed 1.5× beyond the interquartile range. **e**,**f**, The effect of metabolic complexity on divergence for single- and mixed-metabolite conditions with trophic resource transformations (**e**) and random transformations (**f**) for resource transformations (R) and consumer preferences (C) (one-sided Wald test, P values shown in figure for each parameterization and condition type). **g**,**h**, Using the fully structured configuration, the relationship between diversity and mean divergence where the shaded region represents the 95% confidence interval (**g**) and the relationship between occupancy in conditions and number of source communities (**h**).

In implementing a CRM, we sought to investigate how the divergence-complexity effect depends on consumer preferences and resource transformations. These consumers and resources are typically arranged in a trophic structure, where taxa specialize in the hierarchical consumption of environmentally available metabolites and cross-feed the resulting (simpler) byproducts to taxa at subsequently lower trophic levels[10,22]. Trophic interactions are implicated in explaining the immense diversity of microbial communities since only certain organisms can break down the most complex resources in a given environment[22] and, on average, half of a microbial community's members are auxotrophic and require a cooperative interaction with another organism for essential nutrients[50].

Taxonomic structure in CRMs can be represented by specifying what resources each consumer prefers (Extended Data Fig. 7c,d). In our models, we defined trophic consumer preferences with 'families' of specialists, or modules of consumers that have similar resource preferences (specialization)[10], and generalists, which have a broad distribution of preferences (Extended Data Fig. 7c). In line with our experimental observations (Fig. 4), we set the number of different specialists to be proportional to the complexity of the resource type they prefer. We also generated random controls where consumer preferences were assigned randomly (Extended Data Fig. 7d). The use of such random structure has been used successfully to reproduce certain ecological patterns, suggesting that community assembly can be well approximated by random, dense, complex interactions between microbes and their environment[20,60].

Metabolic structure can be represented by assuming that, upon metabolization, resources of a given type transform into resources

of another specific type (Extended Data Fig. 7a,b). In our models, we defined trophic resource transformations such that complex resources transformed into simpler ones in a hierarchical fashion[20] (Extended Data Fig. 7a). Similar to random consumer preferences, we defined random resource transformations where resources of one type could transform into each other randomly (Extended Data Fig. 7b).

To investigate the role of taxonomic and metabolic structure in community divergence, we simulated communities growing in increasingly complex conditions. We performed these simulations (Fig. 5) using four different CRM configurations (Extended Data Fig. 7): trophic consumer preferences and trophic resource transformations (fully structured; Fig. 5a), random preferences and trophic transformations (resource structured; Fig. 5b), trophic preferences and random transformations (consumer structured; Fig. 5c), and random preferences and transformations (fully random; Fig. 5d). Similar to our experimental design, all four model configurations were initialized with six source communities and seven conditions (four single- and three mixed-resource conditions; Extended Data Fig. 7e,f) and growth dynamics were simulated until reaching a steady state. To understand whether the divergence-complexity effect could emerge solely from the ecological forces encoded in our model (consumer preferences and resource transformations), we assumed physiological parameters, such as rates of consumer growth, consumer maintenance, resource utilization, resource energy density and leakage, to be uniform across all consumers and resources[19,20].

Surprisingly, our simulations showed that trophic structure of the resource transformations alone was necessary and sufficient to reproduce the divergence-complexity effect in both single- and mixed-resource conditions (Fig. 5a–f). Even when consumer preferences were random, the divergence-complexity effect was still observed as long as resource transformations were structured (Fig. 5a,b,e). However, whenever resource transformations were random, all communities diverged equally, irrespective of metabolic complexity, and, thus, there was no divergence-complexity effect (Fig. 5c,d,f).

In addition to reproducing the divergence-complexity effect, our model recovered, as emergent properties, further non-trivial trends detected in our experiment. For example, in model configurations with trophic resource transformations, the divergence-complexity effect was greater for single-resource conditions than mixed ones (Fig. 5e), as observed experimentally (Fig. 2f). Additionally, the maximum divergence in single resource conditions exceeded that of mixed conditions (Figs. 5a,b and Fig. 2e). These model configurations also reproduced our downstream analyses, such as the correlation between divergence and diversity (Figs. 5g and 3b) and the tendency for specialists to be endemic (Figs. 5h and 4d). The reproduction of all of these patterns with physiologically neutral consumers (uniform physiological parameters) and resources implicates the trophic metabolic structure in resource transformations as the driving mechanism of the divergence-complexity effect.

Discussion

Compelled by recent experiments that found that microbial community diversity increases with metabolic complexity[9,18], we sought to reconcile contradictory interpretations of whether microbial communities tend to converge[8] or diverge[11] in the same conditions. By jointly revisiting these two propositions, we hypothesized and provided support for the divergence-complexity effect as a quantitative ecological principle that resolves this contradiction. While previous work explored community assembly by modulating the complexity of metabolic conditions[9,10,18] or the variability of source communities[8,11,12], the divergence-complexity effect could be observed only by systematically varying both, that is, analysing multiple source communities under increasingly complex conditions. We found that divergence correlates strongly with diversity, which is driven by an enrichment of specialists in complex conditions. We concluded our analysis by reproducing these results using CRM simulations, which provide insights into the potential ecological mechanisms of the divergence-complexity effect.

While our experimental results are robust and reproducible, they necessarily rely on specific design constraints and assumptions. Experimental choices that could be revisited in future studies include the passaging time, chosen here to be 3 days, as used in other microbial community assembly studies with complex metabolites[11]; the selection of metabolites, which constitute a representative, but oversimplified version of the metabolic complexity of soil environments; our plate shaking rate (Methods), which may have affected oxygen availability[61] and the balance of aerobic and anaerobic metabolism; and the focus on taxonomic divergence (through 16S amplicon sequencing) rather than functional divergence, which would require a comprehensive profiling of microbial functions with metagenomics, metatranscriptomics or metabolomics. Furthermore, future research could explore which environmental parameters, in addition to metabolic complexity, may drive divergence[62,63].

The most surprising result from our simulations was how structured resource transformations (where complex metabolites are progressively degraded into simpler ones), but not consumer preferences, were required for reproducing the divergence-complexity effect (Fig. 5). A possible interpretation of this result is that microbial community assembly and dynamics are strongly dependent on the actual structured architecture of metabolism, which differs substantially from a network of random transformations[20,60]. To test the generality of the divergence complexity, we tried to minimize the use of data-derived parameters in the model. One exception is the parameterization of structure in consumer preferences based on our observed distribution of taxa, which is in any case ultimately shown to be unnecessary to yield the divergence-complexity effect. Conversely, we cannot rule out that additional or alternative data inputs into the model could affect its performance. For example, our CRMs lacked the encoding of certain processes that are known to affect microbial community assembly such as nuanced cross-feeding mechanisms (Supplementary Information), diffusion[64], transcriptional regulation[65], trade-offs between growth rate and enzyme production[66,67], evolution[41] and antibiotics[68]. Explicit measurement of these processes in future studies or inclusion with other ecological models[69] could help to reveal further mechanistic insights into the diversity-complexity effect. However, the fact that our current model captures so many of our experimental observations lends confidence to the dominant role that the architecture of metabolism plays in community structure, corroborating previous reports[18].

Importantly, the divergence-complexity effect has direct implications for the engineering of microbial communities towards any target, suggesting that metabolically complex environments may be more susceptible to microbiome engineering than simple ones. Potential targets for microbiome engineering include correcting the dysbiosis in the human gut[2] and increasing the carbon stabilization capacity of soils[5], among many other microbially regulated traits. The consequences of the divergence-complexity effect are encouraging for efforts along these lines, since complex environments may be more likely to support an alternative community that is equally stable as the original one but with potentially increased expression of a trait of interest. Approaches such as directed evolution, where a set of microbial communities undergoes iterative rounds of perturbation and artificial selection to assemble high-performing communities[70], offer an ideal strategy for exploring the different alternative states that a complex environment can support. Future research is required to understand how, in light of functional redundancy, the divergence in taxonomic composition that we observe relates to divergence in functional composition, since modifying functional activity is commonly the goal of microbiome engineering efforts. Ultimately, we envisage that the awareness of the divergence-complexity effect may help microbial ecologists reframe the role of environmental selection in microbial

community assembly and enable further research into microbiome engineering in complex environments.

Methods

Media preparation

Eleven different media were generated at equimolar (50 mM) concentrations of carbon (C) in increasing levels of complexity. Stocks of citrate, glucose, cellobiose, cellulose and lignin were generated at 1 mol C l^{-1} (1 M C) and then sterilized. Citrate, glucose and cellobiose stocks were sterilized through 0.2 µm RapidFlow filters, while cellulose and lignin stocks, whose particle sizes were too large for filters, were autoclaved. C source stocks were then mixed with M9 minimal media (5× M9 salts (BD Difco), 1 M $MgSO_4$, 1 M $CaCl_2$, 1,000× trace minerals)[71] to form the following nine conditions, each made at a final concentration of 50 mM C and with equal ratios of each C source (Sigma-Aldrich) for each condition: citrate, glucose, cellobiose, cellulose, lignin, citrate + glucose, citrate + glucose + cellobiose, citrate + glucose + cellobiose + cellulose and citrate + glucose + cellobiose + cellulose + lignin. All media were stored in glass bottles, wrapped in foil and stored at 4 °C.

Sample collection and microbial community extraction

On 27 October 2022, about half a pound of organic horizon soil (5–10 cm deep) was collected from six sites at Harvard Forest in Petersham, Massachusetts, United States, following sampling approval from Harvard Forest staff. Two were pine dominated, two were hardwood dominated and two were mixed. Samples were collected 15 m from the forest edge and kept on ice until transported back to the laboratory the same day. Fresh soils were sieved through a 2 mm mesh and then stored at 4 °C. On 21 November 2022, 20 g of each sieved soil was individually combined with 100 ml of sodium pyrophosphate to separate cells from soils[72] and was blended for three cycles of 10 s at ~22,000 RPM and then off for 10 s, and then 25 ml of the resulting slurry was transferred to a centrifuge tube. The blender was washed between each sample by blending in 500 ml of diluted bleach. Following the blending of all soils, each slurry was centrifuged for 10 min at 20,000*g*, resuspended in 30 ml of phosphate-buffered saline and rocked on an orbital shaker for 1 h (ref. 18) at 4 °C. After rocking, samples were allowed to settle for 5 min and then passed through a 100 µm cell straining filter. Optical density (OD) measurements were performed at 600 nm at a 1:20 dilution, 500 µl of each sample was stored at −80 °C in 20% glycerol and the remaining volume from each sample was used for inoculating experimental plates.

Experimental culturing

Community extracts were added to 96-deep-well plates in triplicate with all media combinations (3 replicates of each source community in each condition), generating a total of 162 microcosms (9 media combinations × 6 source communities × 3 replicates). Cycloheximide, an antifungal agent, was added to each well at 200 µg ml^{-1} (ref. 73) to reach a final OD of 0.1 and volume of 400 µl per well. Plates were then stored in an incubator at 25 °C under constant shaking at 200 RPM. Communities were passaged every 72 h into fresh media at a 1:20 dilution (without cycloheximide) for a final volume of 400 µl, OD_{600} was measured and the remaining volume was stored at −80 °C for DNA extraction.

DNA extraction and sequencing

DNA extraction for the six initial forest soil communities was performed using the PowerSoil DNA extraction (QIAGEN). After adding lysis buffer, samples underwent three cycles of freezing in liquid nitrogen, warming at 55 °F in a water bath and bead-beating for 1 min (PowerLyzer, MoBio), then following the provided protocol for the remainder of the extraction. DNA was extracted from an additional 330 lab-cultured samples from days 3, 6, 9, 12 and 33, where available (Extended Data Table 1). DNA extraction from lab-cultured samples was performed using the PureLink Pro 96 Genomic DNA Kit (ThermoFisher) following the provided protocol except for extending all lysis incubation periods to 2 h. DNA extracts were sent to Quintara Biosciences for library preparation and 16S amplicon sequencing using V4 primers 515 F (GTGYCAGCMGCCGCGGTAA) and 806 R (GGACTACNVGGGTWTCTAAT) on a single Illumina MiSeq run.

Amplicon sequence processing

We received raw sequencing data for our study from Quintara Biosciences and downloaded raw sequencing data from Goldford et al.[8] (SRP144982) and Bittleston et al.[11] (SRP218147) from NCBI. All raw 16S sequencing data for each study were separately processed using BU16S (https://github.com/Boston-University-Microbiome-Initiative/BU16s), a QIIME2[74] pipeline customized to run on Boston University's Shared Computing Cluster. Briefly, BU16S first trims primers and filters out reads of less than 50 base pairs using cutadapt[75], then obtains ASVs using dada2[76] and finally classifies ASVs with 95% or greater sequence identity to the SILVA_132_99 database with VSEARCH[77].

Data analysis

All data analysis was performed in Python version 3.8.11. Pairwise distances between samples were computed with the Aitchison distance because it accounts for the compositional nature of sequencing data, unlike common distance metrics, such as Bray–Curtis, Jenson–Shannon divergence and Unifrac[78,79]. The Aitchison distance, A, between two compositions $\mathbf{x}$ and $\mathbf{y}$ is

$$A(\mathbf{x},\mathbf{y}) = \left\|\mathrm{clr}(\mathbf{x}) - \mathrm{clr}(\mathbf{y})\right\|$$

$$\mathrm{clr}(\mathbf{x}) = (\log(x_1/G(\mathbf{x})), \ldots, \log(x_n/G(\mathbf{x})))$$

$$G(\mathbf{x}) = (x_1 \cdot x_2 \cdot \ldots \cdot x_n)^{1/n},$$

where clr is the centre-log ratio transform and G is the geometric mean. Divergence was computed by calculating the Aitchison distance between all pairs of samples within a condition at each timepoint (see 'Measuring divergence with the Aitchison distance metric' section in Supplementary Information for more details). The divergence for samples at day 33, where we have up to three replicates for each community, was reported as the mean pairwise distance between all replicates in all analyses. The divergence-complexity effect was calculated for each condition type (single- and mixed-metabolite conditions) and timepoint as the slope of a linear least-square regression with metabolic complexity as the independent variable and divergence between all community pairs as the dependent variable. In these linear models, the metabolic complexity of each condition was treated ordinally, increasing from citrate to lignin for single-metabolite conditions and from citrate + glucose to all metabolites for the mixed-metabolite conditions. The significance of the divergence-complexity effect for condition and timepoint was computed with a one-sided Wald test with a null hypothesis of a slope of zero. The mean divergence for each sample in a given condition was computed as the mean pairwise distance from each sample to all other samples in that condition.

Dimensionality reduction of pairwise distances was performed using multidimensional scaling (MDS) in scikit-learn[80]. MDS was computed separately for samples from Goldford et al. and Bittleston et al., while, for our data, MDS was computed jointly on all samples to allow for ease of comparability when viewing community trajectories in separate conditions.

Alpha-diversity was computed by first rarefying (subsampling) all samples to 5,028 reads and dropping 12 samples below this sequencing depth from subsequent alpha-diversity analyses. The Shannon diversity index was calculated as $-\Sigma_i x_i \ln x_i$ and the ecological richness was calculated as $\Sigma_i x_i > 0$ for the abundance, x, of each taxon, i, within a sample.

Condition specificity was calculated for each ASV by calculating the fraction of times each ASV was present in each condition. ASVs with

a condition specificity of 1 were considered 'specialists' since they were found to occur in only a single condition.

Consumer-resource models

We simulated the growth of 168 microcosms (6 communities × 7 conditions × 4 configurations) using microbial CRMs. All model parameters are discussed below and summarized in Extended Data Table 2. With microbial CRMs, the dynamics of species and resources can be modelled with

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = N_i\left(\sum_{\alpha}(1-l)c_{i,\alpha}R_\alpha - m\right)$$

$$\frac{\mathrm{d}R_\alpha}{\mathrm{d}t} = (R_\alpha^0 - R_\alpha) - \sum_j N_j c_{j,\alpha}R_\alpha + \sum_{j,\beta} N_j c_{j,\beta}R_\beta D_{\alpha,\beta}l,$$

where N_i is the abundance of species i, R_α is the concentration of resource α, R_α^0 is the resource supply concentration, l is the leakage fraction (how much of α is 'leaked' or converted into β, where the rest is converted into biomass), m is the consumer maintenance cost, $c_{i,\alpha}$ is the consumer preference matrix and $D_{\alpha,\beta}$ is the resource transformation matrix describing the rate that β turns into α following consumption[19]. We fixed the leakage ($l = 0.8$), maintenance ($m = 1$) and uptake rates ($c_{i,\alpha} = \{0, 1\}$) for all consumers, resulting in an 'physiologically neutral' model.

To study the impacts of trophic structure on divergence, we explored four different CRM configurations that varied in whether or not resource transformations or consumer preferences were trophically structured or random (Extended Data Fig. 7). Resources were defined by establishing a resource pool of four resource types, T0, T1, T2 and T3, where each type consisted of 80, 60, 40 and 20 resources, respectively. Trophic resource transformations were parameterized by defining resources of one type to subsequently transform into resources of another type in a unidirectional fashion, with some self-renewal (Extended Data Fig. 7a). For example, T0-type resources mostly transform into T1-type resources and some T0-type resources (Extended Data Fig. 7a). Random resource transformations were defined by allowing each resource transform into any other resource with uniform probability (Extended Data Fig. 7b). Transformation profiles for each resource in both configurations were sampled from Dirichlet distributions.

Consumers were defined by establishing a metacommunity of four 'families', F0, F1, F2 and G, where each type consisted of 500, 300, 100 and 100 consumers, respectively. Trophic consumer preferences were defined by allowing consumers of each family to utilize a total of 35 sampled resources from their associated type (that is F0 consumers could utilize T0 resources) and a common resource type. The skewed distribution of consumer family size was chosen to model our experimental results where the number of specialists correlated with metabolite complexity. Consumers belonging to the G (generalist) family could consume resources of any type (Extended Data Fig. 7c). Random consumer preferences were defined by allowing each consumer to utilize 35 random resources from any type (Extended Data Fig. 7d). Code for sampling trophic resource transformations and trophic consumer preferences can be found via GitHub at https://github.com/michaelsilverstein/ms_tools/blob/main/ms_tools/crm.py.

Initial conditions and source communities were defined to mimic our experimental design. Seven conditions were defined by sampling 20 resources each with an initial abundance of 50 from each resource type for single-metabolite conditions (T0, T1, T2 and T3) and from mixtures of resource types for mixed-metabolite conditions (T3 + T2, T3 + T2 + T1 and T3 + T2 + T1 + T0; Extended Data Fig. 7e). Six source communities were defined by sampling 200 consumers from the metacommunity each with an initial abundance of 1 (Extended Data Fig. 7f).

The dynamics of each source community was then simulated in each condition using all four parameter configurations (trophic transformations and preferences, trophic transformations and random preferences, random transformations and trophic preferences, and random transformations and preferences), and the divergence was computed for each condition and configuration. Simulations of community assembly were performed by passing model parameters (D matrix, c matrix and initial conditions) to the Community Simulator package[21], which provides utility functions for constructing and solving the system of ordinary differential equations. To appropriately compare our simulation results, which simulates actual abundances of each consumer, with our experimental results, which reports relative abundance of each ASV, we rescaled the abundance of all communities to the same range to simulate the process of sequencing. Divergence was then calculated on the rescaled simulated community composition profiles in the same way as with our experimental data (using the Aitchison distance).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All raw 16S sequencing data and associated metadata generated for this study can be accessed with the NCBI BioProject accession PRJNA1074799 (https://www.ncbi.nlm.nih.gov/sra/PRJNA1074799). All processed data used in our analyses can be found on GitHub at https://github.com/segrelab/MetabolicComplexityDivergence. Source data are provided with this paper.

Code availability

All analysis and simulation code can be found on GitHub at https://github.com/segrelab/MetabolicComplexityDivergence.

References

1. Gilbert, J. A. et al. Current understanding of the human microbiome. *Nat. Med.* **24**, 392–400 (2018).
2. Bowman, K. A., Broussard, E. K. & Surawicz, C. M. Fecal microbiota transplantation: current clinical efficacy and future prospects. *Clin. Exp. Gastroenterol.* **8**, 285–291 (2015).
3. Averill, C. et al. Defending Earth's terrestrial microbiome. *Nat. Microbiol.* **7**, 1717–1725 (2022).
4. Jansson, J. K. & Hofmockel, K. S. Soil microbiomes and climate change. *Nat. Rev. Microbiol.* **18**, 35–46 (2020).
5. Silverstein, M. R., Segrè, D. & Bhatnagar, J. M. Environmental microbiome engineering for the mitigation of climate change. *Glob. Change Biol.* https://doi.org/10.1111/gcb.16609 (2023).
6. Martiny, J. B. H. et al. Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.* **4**, 102–112 (2006).
7. Martiny, J. B. H., Eisen, J. A., Penn, K., Allison, S. D. & Horner-Devine, M. C. Drivers of bacterial β-diversity depend on spatial scale. *Proc. Natl Acad. Sci. USA* **108**, 7850–7854 (2011).
8. Goldford, J. E. et al. Emergent simplicity in microbial community assembly. *Science* **361**, 469–474 (2018).
9. Pacheco, A. R., Osborne, M. L. & Segrè, D. Non-additive microbial community responses to environmental complexity. *Nat. Commun.* **12**, 2365 (2021).
10. Enke, T. N. et al. Modular assembly of polysaccharide-degrading marine microbial communities. *Curr. Biol.* **29**, 1528–1535.e6 (2019).
11. Bittleston, L. S., Gralka, M., Leventhal, G. E., Mizrahi, I. & Cordero, O. X. Context-dependent dynamics lead to the assembly of functionally distinct microbial communities. *Nat. Commun.* **11**, 1440 (2020).
12. Čaušević, S., Tackmann, J., Sentchilo, V., von Mering, C. & van der Meer, J. R. Reproducible propagation of species-rich soil bacterial communities suggests robust underlying deterministic principles of community formation. *mSystems* **7**, e00160–22 (2022).

13. Rivett, D. W. & Bell, T. Abundance determines the functional role of bacterial phylotypes in complex communities. *Nat. Microbiol.* **3**, 767–772 (2018).
14. Datta, M. S., Sliwerska, E., Gore, J., Polz, M. F. & Cordero, O. X. Microbial interactions lead to rapid micro-scale successions on model marine particles. *Nat. Commun.* **7**, 11965 (2016).
15. Estrela, S. et al. Functional attractors in microbial community assembly. *Cell Syst.* **13**, 29–42.e7 (2022).
16. Grilli, J. Macroecological laws describe variation and diversity in microbial communities. *Nat. Commun.* **11**, 4743 (2020).
17. Lawson, C. E. et al. Common principles and best practices for engineering microbiomes. *Nat. Rev. Microbiol.* **17**, 725–741 (2019).
18. Dal Bello, M., Lee, H., Goyal, A. & Gore, J. Resource–diversity relationships in bacterial communities reflect the network structure of microbial metabolism. *Nat. Ecol. Evol.* **5**, 1424–1434 (2021).
19. Marsland, R. et al. Available energy fluxes drive a transition in the diversity, stability, and functional structure of microbial communities. *PLoS Comput. Biol.* **15**, e1006793 (2019).
20. Marsland, R., Cui, W. & Mehta, P. A minimal model for microbial biodiversity can reproduce experimentally observed ecological patterns. *Sci. Rep.* **10**, 3308 (2020).
21. Marsland, R., Cui, W., Goldford, J. & Mehta, P. The Community Simulator: a Python package for microbial ecology. *PLoS ONE* **15**, e0230430 (2020).
22. Gralka, M., Szabo, R., Stocker, R. & Cordero, O. X. Trophic interactions and the drivers of microbial community assembly. *Curr. Biol.* **30**, R1176–R1188 (2020).
23. Thompson, L. R. et al. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457–463 (2017).
24. Sposito, G. *The Chemistry of Soils* (Oxford Univ. Press, 2016).
25. Atiwesh, G., Parrish, C. C., Banoub, J. & Le, T. T. Lignin degradation by microorganisms: a review. *Biotechnol. Prog.* **38**, e3226 (2022).
26. Weng, C., Peng, X. & Han, Y. Depolymerization and conversion of lignin to value-added bioproducts by microbial and enzymatic catalysis. *Biotechnol. Biofuels* **14**, 84 (2021).
27. Xu, Z., Lei, P., Zhai, R., Wen, Z. & Jin, M. Recent advances in lignin valorization with bacterial cultures: microorganisms, metabolic pathways and bio-products. *Biotechnol. Biofuels* **12**, 32 (2019).
28. Iram, A., Berenjian, A. & Demirci, A. A review on the utilization of lignin as a fermentation substrate to produce lignin-modifying enzymes and other value-added products. *Molecules* **26**, 2960 (2021).
29. Li, K. et al. Investigating lignin-derived monomers and oligomers in low-molecular-weight fractions separated from depolymerized black liquor retentate by membrane filtration. *Molecules* **26**, 2887 (2021).
30. Gilkes, N. R., Kilburn, D. G., Miller, R. C. & Warren, R. A. J. Bacterial cellulases. *Bioresour. Technol.* **36**, 21–35 (1991).
31. Bhardwaj, N., Kumar, B., Agrawal, K. & Verma, P. Current perspective on production and applications of microbial cellulases: a review. *Bioresour. Bioprocess.* **8**, 95 (2021).
32. Jones, D. L. Organic acids in the rhizosphere—a critical review. *Plant Soil* **205**, 25–44 (1998).
33. Fischer, Z., Blažka, P. & Dubis, L. Respiration rates of organic soil depending on changes of moisture and aeration. *Open J. Soil Sci.* **7**, 101–110 (2017).
34. Ramonell, K. M. et al. Influence of atmospheric oxygen on leaf structure and starch deposition in *Arabidopsis thaliana*: low oxygen effects on leaf development in *Arabidopsis*. *Plant Cell Environ.* **24**, 419–428 (2001).
35. Petersen, S. O., Nielsen, T. H., Frostegård, Å. & Olesen, T. O_2 uptake, C metabolism and denitrification associated with manure hot-spots. *Soil Biol. Biochem.* **28**, 341–349 (1996).
36. Sierra, J. & Renault, P. Oxygen consumption by soil microorganisms as affected by oxygen and carbon dioxide levels. *Appl. Soil Ecol.* **2**, 175–184 (1995).
37. Parisutham, V., Chandran, S.-P., Mukhopadhyay, A., Lee, S. K. & Keasling, J. D. Intracellular cellobiose metabolism and its applications in lignocellulose-based biorefineries. *Bioresour. Technol.* **239**, 496–506 (2017).
38. Blount, Z. D., Borland, C. Z. & Lenski, R. E. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **105**, 7899–7906 (2008).
39. Ibarra, R. U., Edwards, J. S. & Palsson, B. O. *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* **420**, 186–189 (2002).
40. Goyal, A., Bittleston, L. S., Leventhal, G. E., Lu, L. & Cordero, O. X. Interactions between strains govern the eco-evolutionary dynamics of microbial communities. *eLife* **11**, e74987 (2022).
41. Dragosits, M. & Mattanovich, D. Adaptive laboratory evolution—principles and applications for biotechnology. *Microb. Cell Fact.* **12**, 64 (2013).
42. Debray, R. et al. Priority effects in microbiome assembly. *Nat. Rev. Microbiol.* **20**, 109–121 (2022).
43. Estrela, S., Sanchez-Gorostiaga, A., Vila, J. C. & Sanchez, A. Nutrient dominance governs the assembly of microbial communities in mixed nutrient environments. *eLife* **10**, e65948 (2021).
44. Louca, S. et al. Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.* **2**, 936–943 (2018).
45. Allison, S. D. & Martiny, J. B. H. Resistance, resilience, and redundancy in microbial communities. *Proc. Natl Acad. Sci. USA* **105**, 11512–11519 (2008).
46. Louca, S. et al. High taxonomic variability despite stable functional structure across microbial communities. *Nat. Ecol. Evol.* **1**, 0015 (2017).
47. Dubinkina, V., Fridman, Y., Pandey, P. P. & Maslov, S. Multistability and regime shifts in microbial communities explained by competition for essential nutrients. *eLife* **8**, e49720 (2019).
48. Johnson, D. R., Goldschmidt, F., Lilja, E. E. & Ackermann, M. Metabolic specialization and the assembly of microbial communities. *ISME J.* **6**, 1985–1991 (2012).
49. Berlemont, R. & Martiny, A. C. Phylogenetic distribution of potential cellulases in bacteria. *Appl. Environ. Microbiol.* **79**, 1545–1554 (2013).
50. Kost, C., Patil, K. R., Friedman, J., Garcia, S. L. & Ralser, M. Metabolic exchanges are ubiquitous in natural microbial communities. *Nat. Microbiol.* **8**, 2244–2252 (2023).
51. Speck, E. L. & Freese, E. Control of metabolite secretion in *Bacillus subtilis*. *J. Gen. Microbiol.* **78**, 261–275 (1973).
52. Hannya, A., Nishimura, T., Matsushita, I., Tsubota, J. & Kawata, Y. Efficient production and secretion of oxaloacetate from *Halomonas* sp. KM-1 under aerobic conditions. *AMB Express* **7**, 209 (2017).
53. Beg, Q. K. et al. Detection of transcriptional triggers in the dynamics of microbial growth: application to the respiratorily versatile bacterium *Shewanella oneidensis*. *Nucleic Acids Res.* **40**, 7132–7149 (2012).
54. Fernández-Veledo, S. & Vendrell, J. Gut microbiota-derived succinate: friend or foe in human metabolic diseases? *Rev. Endocr. Metab. Disord.* **20**, 439–447 (2019).
55. Balado, M. et al. Secreted citrate serves as iron carrier for the marine pathogen *Photobacterium damselae* subsp *damselae*. *Front. Cell. Infect. Microbiol.* **7**, 361 (2017).

56. Pacheco, A. R., Moel, M. & Segrè, D. Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. *Nat. Commun.* **10**, 103 (2019).
57. Tiedje, J., Sexstone, A., Parkin, T. & Revsbech, N. Anaerobic processes in soil. *Plant Soil* **76**, 197–212 (1984).
58. Fritts, R. K., McCully, A. L. & McKinlay, J. B. Extracellular metabolism sets the table for microbial cross-feeding. *Microbiol. Mol. Biol. Rev.* **85**, e00135–20 (2021).
59. Amarnath, K. et al. Stress-induced metabolic exchanges between complementary bacterial types underly a dynamic mechanism of inter-species stress resistance. *Nat. Commun.* **14**, 3165 (2023).
60. Cui, W., Marsland, R. & Mehta, P. Diverse communities behave like typical random ecosystems. *Phys. Rev. E* **104**, 034416 (2021).
61. Al-Ani, A. et al. Oxygenation in cell culture: critical parameters for reproducibility are routinely not reported. *PLoS ONE* **13**, e0204269 (2018).
62. Rocca, J. D. et al. The microbiome stress project: toward a global meta-analysis of environmental stressors and their effects on microbial communities. *Front. Microbiol.* **9**, 3272 (2019).
63. Mandakovic, D. et al. Structure and co-occurrence patterns in microbial communities under acute environmental stress reveal ecological factors fostering resilience. *Sci. Rep.* **8**, 5875 (2018).
64. Dal Co, A., van Vliet, S., Kiviet, D. J., Schlegel, S. & Ackermann, M. Short-range interactions govern the dynamics and functions of microbial communities. *Nat. Ecol. Evol.* **4**, 366–375 (2020).
65. Carthew, R. W. Gene regulation and cellular metabolism: an essential partnership. *Trends Genet.* **37**, 389–400 (2021).
66. Ramin, K. I. & Allison, S. D. Bacterial tradeoffs in growth rate and extracellular enzymes. *Front. Microbiol.* **10**, 2956 (2019).
67. Malik, A. A. et al. Defining trait-based microbial strategies with consequences for soil carbon cycling under climate change. *ISME J.* **14**, 1–9 (2020).
68. Stokes, J. M., Lopatkin, A. J., Lobritz, M. A. & Collins, J. J. Bacterial metabolism and antibiotic efficacy. *Cell Metab.* **30**, 251–259 (2019).
69. van den Berg, N. I. et al. Ecological modelling approaches for predicting emergent properties in microbial communities. *Nat. Ecol. Evol.* **6**, 855–865 (2022).
70. Sánchez, Á. et al. Directed evolution of microbial communities. *Annu. Rev. Biophys.* **50**, 323–341 (2021).
71. M9 minimal medium (modified). *Cold Spring Harb. Protoc.* pdb. rec12296 (2010).
72. Liu, J., Li, J., Feng, L., Cao, H. & Cui, Z. An improved method for extracting bacteria from soil for high molecular weight DNA recovery and BAC library construction. *J. Microbiol.* **48**, 728–733 (2010).
73. Syuhada, N. H. et al. Strong and widespread cycloheximide resistance in *Stichococcus*-like eukaryotic algal taxa. *Sci. Rep.* **12**, 1080 (2022).
74. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
75. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10 (2011).
76. Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
77. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
78. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome datasets are compositional: and this is not optional. *Front. Microbiol.* **8**, 2224 (2017).
79. Quinn, T. P., Erb, I., Richardson, M. F. & Crowley, T. M. Understanding sequencing data as compositions: an outlook and review. *Bioinformatics* **34**, 2870–2878 (2018).
80. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

Acknowledgements

We thank M. Dal Bello, A. Goyal, M. Gralka, Z. Werbin, J. Zhang, K. Herbst, J. Goldford, H. Scott, S. Bald and the members of the Bhatnagar and Segrè labs for their guidance and insight. Additionally, we thank C. Vietorisz for assistance with soil community sampling. M.R.S. was supported by a synthetic biology NIH-funded predoctoral training fellowship (T32GM130546), a bioinformatics NIH-funded predoctoral training fellowship (T32GM100842), the Biological Design Center Kilachand Multicellular Design Program Graduate Fellowship and the Hariri Graduate Student Fellows Program. J.M.B. acknowledges support from DOE DE-SC0020403 and DOE DE-SC0012704. D.S. acknowledges support from the National Institutes of Health (National Institute of General Medical Sciences, award R01GM121950; National Institute on Aging, award number UH2AG064704), the US Department of Energy, Office of Science, Office of Biological & Environmental Research through the Microbial Community Analysis and Functional Evaluation in Soils SFA Program (m-CAFEs) under contract number DE-AC02-05CH11231 to Lawrence Berkeley National Laboratory, the National Science Foundation (grants 1457695, NSFOCE-BSF 1635070 and NSF-BSF 2246707; and the NSF Center for Chemical Currencies of a Microbial Planet, publication #044) and the Human Frontiers Science Program (RGP0020/2016 and RGP0060/2021). Figures were created with BioRender.com.

Author contributions

M.R.S., J.M.B. and D.S. conceived the study. M.R.S. acquired samples, conducted the experiments, performed the simulations and analysed the data. M.R.S., J.M.B. and D.S. interpreted results and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

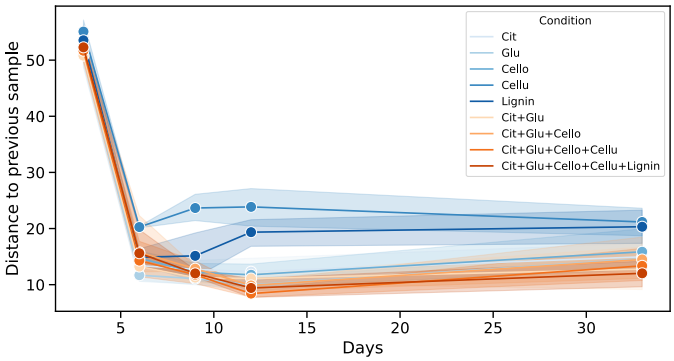
Extended data is available for this paper at https://doi.org/10.1038/s41559-024-02440-6.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-024-02440-6.

Correspondence and requests for materials should be addressed to Daniel Segrè.

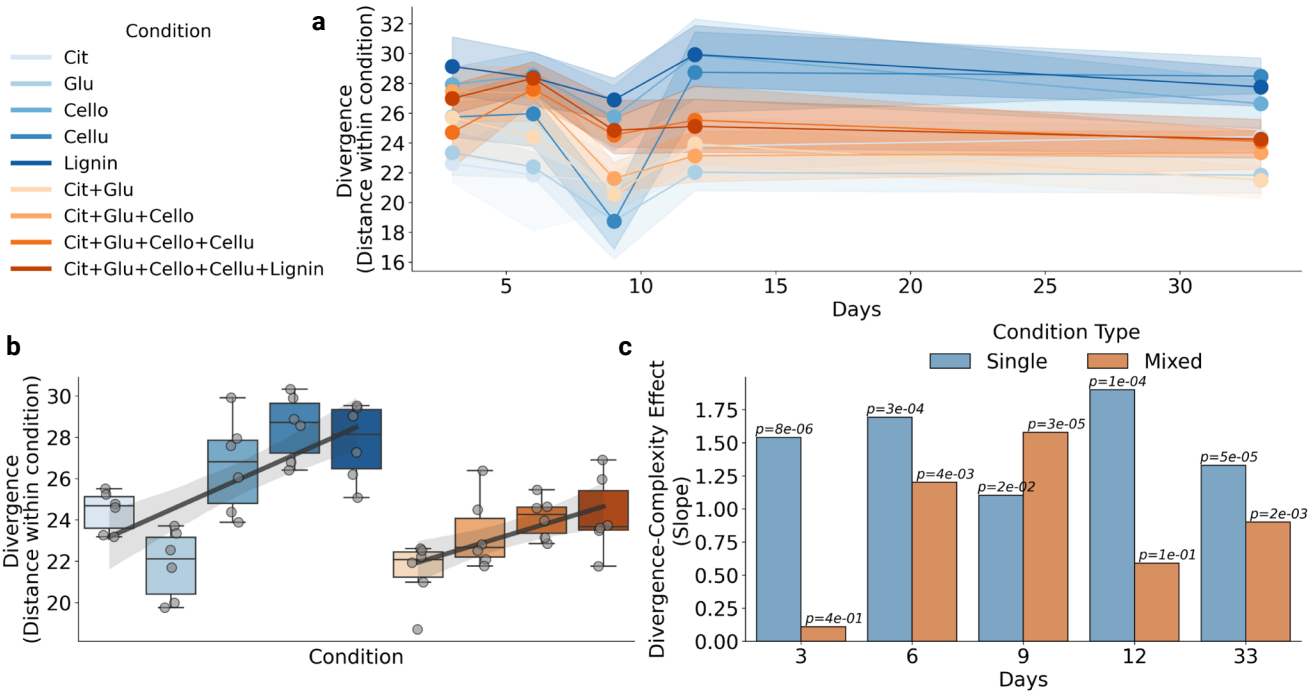
Peer review information *Nature Ecology & Evolution* thanks Otto Cordero and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

https://doi.org/10.1038/s41559-024-02440-6



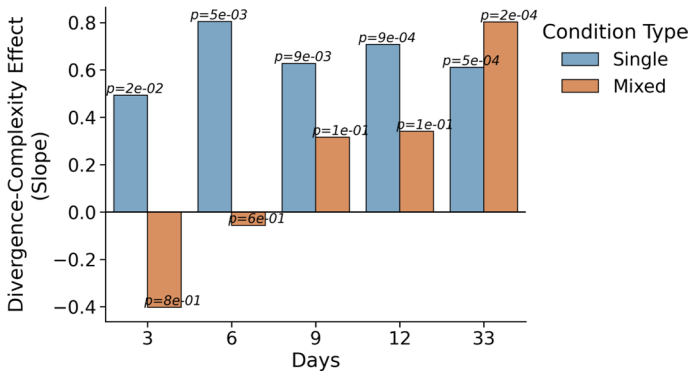
Extended Data Fig. 1 | Microbial communities stabilize. Each point is the mean distance to a community's previous timepoint within each condition and the shaded region represents the 95% confidence interval. On day 3, communities are substantially different from their initial state and then continue to change in composition, but by approximately the same amount.

https://doi.org/10.1038/s41559-024-02440-6



Extended Data Fig. 2 | Divergence of a subset of communities which are available for all days. A reproduction of Fig. 1d–f, but excluding all samples from communities HF1P and HF3H, which are missing data for days 3, 6, and 9 (Extended Data Table 1). **a**, Divergence of communities within each condition over time from day 3 onwards. Points on each line represent the mean divergence and the shaded region represents the 95% confidence interval for pairwise distances between all four included communities within each condition. **b**, Distribution of divergence for the final time point. Boxes are bound by the interquartile range, divided by the median, and whiskers extend to a maximum of 1.5 times the interquartile range. **c**, Divergence-complexity effect by condition type for all time points (one-sided Wald Test on effect (slope) > 0, p-values shown in figure for each day and condition type.

https://doi.org/10.1038/s41559-024-02440-6

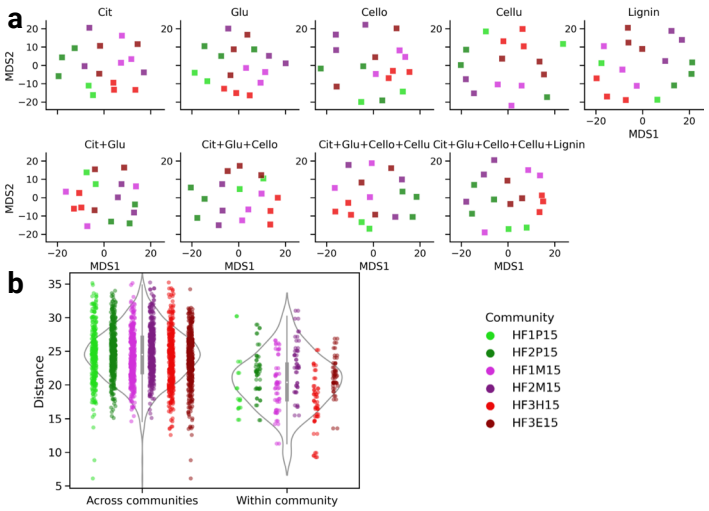


Extended Data Fig. 3 | Divergence-complexity effect at Family level. A reproduction of divergence-complexity effect calculation in Fig. 2f except computing divergence at the Family level instead of at the ASV level (one-sided Wald Test on effect (slope) > 0, p-values shown in figure for each day and condition type). Unlike our results at the ASV level, the divergence-complexity effect for our mixed-metabolite conditions at the Family level was only significant for day 33. The lack of significance at other points may be related to their reduced sample size (for days 3, 6, and 9; see Extended Data Table 1), but may in principle reflect additional effects emerging upon taxonomic coarse graining.

https://doi.org/10.1038/s41559-024-02440-6

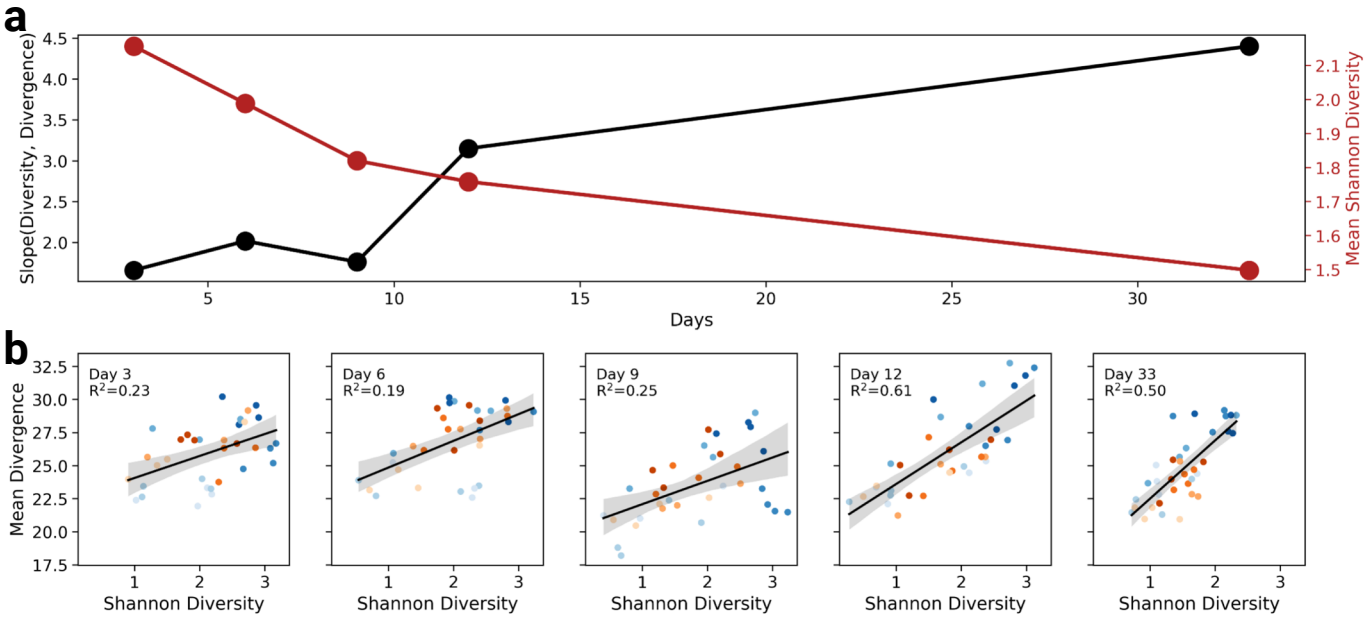
Extended Data Fig. 4 | Family composition over time. Each subplot shows the Family composition over time for a microcosm. Colors for each Family are defined by Phylum where blue represents Actinobacteria, yellow represents Bacteroidetes, green represents Firmicutes, and red represents Proteobacteria. Each column of subplots pertains to one source community and each row pertains to a condition (increasing in complexity from top to bottom, with single- and then mixed-metabolite conditions).

https://doi.org/10.1038/s41559-024-02440-6



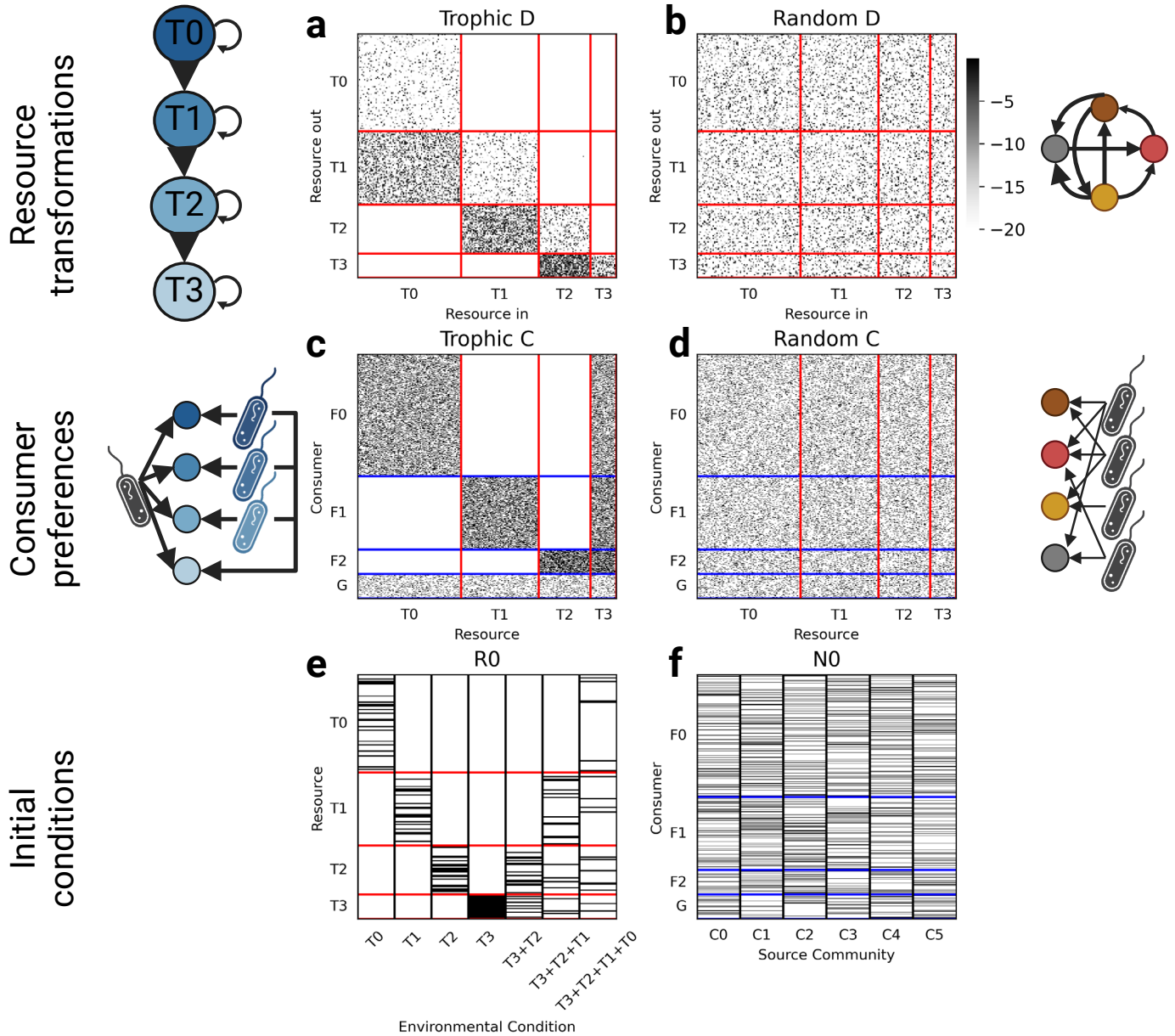
Extended Data Fig. 5 | Microcosm replicates cluster. a, The final state at day 33 of two replicates for community HF1P and three replicates for all other communities (see Extended Data Table 1) projected separately for each condition using MDS. **b**, The distribution of distances across communities (N=1,066 pairwise distances) and between replicates within the same community (N=142 pairwise distances). Each violin outlines the kernel density estimate and contains a box which is bound by the interquartile range with an open circle at the media and whiskers that extend up to 1.5 times the interquartile range. Communities are colored by their source and squares represent the replicate used in the main text. Independent of condition, communities from the same replicate are more similar to each other than communities from separate replicates.

https://doi.org/10.1038/s41559-024-02440-6



Extended Data Fig. 6 | Divergence and diversity on a subset of communities which are available for all days. A reproduction of Fig. 3, but excluding all samples from communities HF1P and HF3H, which are missing data for days 3, 6, and 9 (Extended Data Table 1). **a**, The slope of the relationship between community alpha diversity and divergence (black) and the mean community alpha diversity (red) over time. Shaded areas around each regression line represents the 95% confidence interval. **b**, The data underlying the relationship in **a** over time. Each point is the diversity of a community in a condition (x-axis) and the mean divergence of that community from all others within a condition (y-axis).

https://doi.org/10.1038/s41559-024-02440-6



Extended Data Fig. 7 | Consumer resource parameterizations. *D* Matrices for trophic (**a**) and random (**b**) resource transformations. Each matrix column defines which resources are generated from a given input resource. **a**, Trophic resource transformations were parameterized by defining resource types (T0, T1, T2, and T3; red lines) which mostly transform into resources of the subsequent type with some self-renewal. **b**, Random resource transformations were parameterized by defining any resource to transform into any other resource with uniform probability. The colorbar indicates $\log_{10}$ resource transformation rate for **a** and **b**. *C* matrices for trophic (**c**) and random (**d**) consumer preferences. Each matrix row defines which resources can be utilized by each consumer. **c**, Trophic consumer preferences were parameterized by defining consumer families (F0, F1, F2, and G; blue lines) which consume a total of 35 resources of their associated type (for example, F0 consumers utilize T0 resources) and a common resource, T3. G consumers (generalists) consume resources of any type. **d**, Random consumer preferences were parameterized by defining consumers to utilize any resource with uniform probability. **e**, Seven initial environmental conditions were defined with 20 resources sampled from (1) T0, (2) T1, (3) T2, (4) T3, (5) T3+T2, (6) T3+T2+T1, and (7) T3+T2+T1+T0. **f**, Six source communities were defined by sampling 200 consumers.

https://doi.org/10.1038/s41559-024-02440-6

Extended Data Table 1 | Availability of each sample

Day	Condition	HF1P	HF2P	HF3H	HF3E	HF1M	HF2M	Total
0		1/1	1/1	1/1	1/1	1/1	1/1	6
	Citrate	**0/1**	1/1	**0/1**	1/1	1/1	1/1	4
	Glucose	**0/1**	1/1	**0/1**	1/1	1/1	1/1	4
	Cellobiose	**0/1**	1/1	**0/1**	1/1	1/1	1/1	4
	Cellulose	**0/1**	1/1	**0/1**	1/1	1/1	1/1	4
3	Lignin	**0/1**	1/1	**0/1**	1/1	1/1	1/1	4
	C+G	**0/1**	1/1	**0/1**	1/1	1/1	1/1	4
	C+G+C	**0/1**	1/1	**0/1**	1/1	1/1	1/1	4
	C+G+C+C	**0/1**	1/1	**0/1**	1/1	1/1	1/1	4
	C+G+C+C+L	**0/1**	1/1	**0/1**	1/1	1/1	1/1	4
	Citrate	**0/1**	1/1	1/1	1/1	**0/1**	1/1	4
	Glucose	**0/1**	1/1	1/1	1/1	1/1	1/1	5
	Cellobiose	**0/1**	1/1	1/1	1/1	1/1	1/1	5
	Cellulose	**0/1**	1/1	1/1	1/1	**0/1**	1/1	4
6	Lignin	**0/1**	1/1	1/1	1/1	1/1	1/1	5
	C+G	**0/1**	1/1	1/1	1/1	1/1	1/1	5
	C+G+C	**0/1**	1/1	1/1	1/1	1/1	1/1	5
	C+G+C+C	**0/1**	1/1	1/1	1/1	1/1	1/1	5
	C+G+C+C+L	**0/1**	1/1	1/1	1/1	1/1	1/1	5
	Citrate	1/1	1/1	**0/1**	1/1	1/1	1/1	5
	Glucose	1/1	1/1	**0/1**	1/1	1/1	1/1	5
	Cellobiose	1/1	1/1	**0/1**	1/1	1/1	1/1	5
	Cellulose	1/1	1/1	**0/1**	1/1	1/1	1/1	5
9	Lignin	1/1	1/1	**0/1**	1/1	1/1	1/1	5
	C+G	1/1	1/1	**0/1**	1/1	1/1	1/1	5
	C+G+C	1/1	1/1	**0/1**	1/1	1/1	1/1	5
	C+G+C+C	1/1	1/1	**0/1**	1/1	1/1	1/1	5
	C+G+C+C+L	1/1	1/1	**0/1**	1/1	1/1	1/1	5
	Citrate	1/1	1/1	1/1	1/1	1/1	1/1	6
	Glucose	1/1	1/1	1/1	1/1	1/1	1/1	6
	Cellobiose	1/1	1/1	1/1	1/1	1/1	1/1	6
	Cellulose	1/1	1/1	1/1	1/1	1/1	1/1	6
12	Lignin	1/1	1/1	1/1	1/1	1/1	1/1	6
	C+G	1/1	1/1	1/1	1/1	1/1	1/1	6
	C+G+C	1/1	1/1	1/1	1/1	1/1	1/1	6
	C+G+C+C	1/1	1/1	1/1	1/1	1/1	1/1	6
	C+G+C+C+L	1/1	1/1	1/1	1/1	1/1	1/1	6
	Citrate	**2/3**	3/3	3/3	3/3	3/3	3/3	17
	Glucose	**2/3**	3/3	3/3	3/3	3/3	3/3	17
	Cellobiose	**2/3**	3/3	3/3	3/3	3/3	3/3	17
	Cellulose	**2/3**	3/3	3/3	3/3	3/3	**2/3**	16
33	Lignin	**2/3**	3/3	3/3	3/3	3/3	3/3	17
	C+G	**2/3**	3/3	3/3	3/3	3/3	3/3	17
	C+G+C	**2/3**	3/3	3/3	3/3	3/3	3/3	17
	C+G+C+C	**2/3**	3/3	3/3	3/3	3/3	3/3	17
	C+G+C+C+L	**2/3**	3/3	3/3	3/3	3/3	3/3	17
	Total	37	64	46	64	62	63	336

When available, one sample was sequenced from each microcosm for days 0, 3, 6, 9, and 12. Three samples were sequenced for day 33. Each column (HF1P, HF2P, HF3H, HF3E, HF1M, HF2M) corresponds to a source community. Bolded text indicates unsequenced conditions.

https://doi.org/10.1038/s41559-024-02440-6

Extended Data Table 2 | Consumer-resource model parameters

Parameter	Description	Value
i	Consumer (species)	
N_i	Abundance of consumer i	
α	Resource	
R_α	Amount of resource α	
R_α^0	Resource supply concentration	
$D_{\alpha,\beta}$	Resource transformation matrix, indicating which resources β are transformed into α following the consumption.	
$\sigma_{i,\alpha}$	Consumer preference matrix, indicating which resources α are consumed by consumer i.	{0, 1}
l	Leakage fraction, or the fraction of β that becomes α, where the remainder becomes biomass.	0.8
m	Maintenance cost	1
	Number of resources	200 total (T0: 80, T1: 60, T2: 40, T3: 20)
	Number of consumers	1000 total (F0: 500, F1: 300, F2: 100, G: 100)
	Number of resource preferences per consumer	35
	Number of communities	6
	Initial number of consumers per community	200
	Initial abundance of consumers	1
	Number of conditions	7
	Number of resources per condition	20
	Initial abundance of resources	50

Description of each consumer-resource model parameter for all parameterizations. All quantities are unitless. See Extended Data Fig. 7 for a visual representation of these parameters.

nature portfolio

Corresponding author(s): Daniel Segrè

Last updated by author(s): Feb 14, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All raw 16S sequencing data for each study was separately processed using BU16S (https://github.com/Boston-University-Microbiome-Initiative/BU16s), a QIIME2 pipeline customized to run on Boston University's Shared Computing Cluster. Consumer-resource model simulations were performed using the Community Simulator package (https://github.com/Emergent-Behaviors-in-Biology/community-simulator) and simulated communities with "trophic" structure were sampled using custom code available at: https://github.com/michaelsilverstein/ms_tools/blob/main/ms_tools/crm.py.
Data analysis	All data analysis was performed in Python version 3.8.11 and code can be found at: https://github.com/segrelab/MetabolicComplexityDivergence.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All raw 16S sequencing data and associated metadata generated for this study can be accessed with the NCBI BioProject accession PRJNA1074799 (https://www.ncbi.nlm.nih.gov/sra/PRJNA1074799). Processed data can be found at: https://github.com/segrelab/MetabolicComplexityDivergence.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used.* *Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected.* *Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Reporting on race, ethnicity, or other socially relevant groupings	*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status).* *Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.)* *Please provide details about how you controlled for confounding variables in your analyses.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Six soil microbial communities were grown on nine different conditions and sequencing was performed on five days of samples (days 3, 6, 9, 12, and 33). Our main metric, divergence, is the pairwise distance between all samples from the same condition and time point. When all samples were available for a given time point and condition, we had 15 (6 choose 2) for that time point and condition.
Data exclusions	All samples with sufficient DNA concentrations were included in our study. Samples without sufficient DNA concentrations were not able to be sequenced and are indicated in Table 1.
Replication	All conditions were cultured in triplicate, however only replicates at day 33 (the final time point) were sequenced and analyzed.
Randomization	Soil microbial communities were grown on different minimal media conditions, each with one carbon source. No carbon source media controls were used, but not reported since those communities did not survive.
Blinding	Our analysis was performed on sequencing data, which was not generated until the experiment concluded.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.