Competition for shared resources increases dependence on initial population size during coalescence of gut microbial communities

Doran A. Goldman[1,2,*], Katherine S. Xue[1,2,*,†], Autumn B. Parrott[3], Rashi R. Jeeda[4,‡], Lauryn R. Franzese[3], Jaime G. Lopez[3], Jean C. C. Vila[1], Dmitri A. Petrov[1,5], Benjamin H. Good[1,5,6], David A. Relman[2,7,8], Kerwyn Casey Huang[2,3,5,†]

[1]Department of Biology, Stanford University, Stanford, CA 94305, USA

[2]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA

[3]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA

[4]Department of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

[5]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

[6]Department of Applied Physics, Stanford University, Stanford, CA 94305, USA

[7]Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

[8]Infectious Diseases Section, Veterans Affairs Palo Alto Health Care System, Palo Alto, CA 94304, USA

*These authors contributed equally.

[†]Correspondence: kxue@stanford.edu (K.S.X.) and kchuang@stanford.edu (K.C.H.).

[‡]Current affiliation: Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract

The long-term success of introduced populations depends on their initial size and ability to compete against existing residents, but it remains unclear how these factors collectively shape colonization. Here, we investigate how initial population (propagule) size and resource competition interact during community coalescence by systematically mixing eight pairs of *in vitro* microbial communities at ratios that vary over six orders of magnitude, and we compare our results to a neutral ecological model. Although the composition of the resulting co-cultures deviated substantially from neutral expectations, each co-culture contained species whose relative abundance depended on propagule size even after ~40 generations of growth. Using a consumer-resource model, we show that this dose-dependent colonization can arise when resident and introduced species have high niche overlap and consume shared resources at similar rates. This model predicts that propagule size will have larger, longer-lasting effects in diverse communities in which niche overlap is higher, and we experimentally confirm that strain isolates show stronger dose dependence when introduced into diverse communities than in pairwise co-culture. This work shows how neutral-like colonization dynamics can emerge from non-neutral resource competition and have lasting effects on the outcomes of community coalescence.

Introduction

Colonization by new species plays a major role in shaping community composition and function, but the outcomes of species introductions are notoriously difficult to predict[1–7]. To colonize, a species must disperse from its original population[8–10], overcome the stochastic effects of ecological drift[11,12], and compete successfully against resident species for a niche in the community[13–17]. Models of community assembly place different emphasis on the role of neutral forces—namely, dispersal and ecological drift—compared to competition[15,18,19], and it remains unclear how these neutral and non-neutral forces interact to shape colonization outcomes.

Ecological models differ especially in their predictions about how colonization outcomes depend on the size and frequency of introduced populations, together known as propagule pressure[5,20]. The neutral theory of biodiversity, which assumes that species are ecologically equivalent, predicts that the abundance of an introduced species will remain proportional to its initial propagule size[15,18,21]. In contrast, in classical consumer-resource models, each introduced species reaches an equilibrium abundance determined by its competitive ability, regardless of its initial abundance[19,22]. Investigating the effect of propagule size therefore provides an opportunity to quantify the impact of neutral and non-neutral forces on colonization.

So far, empirical studies have come to contradictory conclusions about the effect of propagule size on colonization. Large propagule size is often associated with successful colonization[23–27], consistent with neutral expectations. However, there are many cases

in which large propagules fail to colonize or small propagules successfully establish[5,20,21,28–34], showing that propagule size is not always the main determinant of colonization success. Many of these studies focus on the introduction of a small number of species[24–26,28,31,32,34], making it difficult to generalize their conclusions across species and community contexts. Resolving these discrepancies requires systematic, large-scale quantification of the effect of propagule size.

To address this gap, we investigated the impact of propagule size on colonization during coalescence of *in vitro* gut microbial communities. Understanding the effect of propagule size in the human gut is important for designing microbiome therapeutics[35,36], but prior studies have come to inconsistent conclusions about how inoculation dose affects the outcome of probiotic introductions and fecal microbiota transplants[37–42]. To investigate propagule size in a more tractable, laboratory setting, we used stool-derived *in vitro* communities, which are stable, diverse, reproducible models of the gut microbiome that can recapitulate *in vivo* responses to perturbation[22,43–47]. We mixed pairs of *in vitro* communities at ratios that varied over six orders of magnitude, allowing us to observe the outcomes of hundreds of species introductions at once. By combining these experiments with consumer-resource modeling, we show how resource competition can amplify the effects of initial population size in complex communities, with lasting effects on the outcomes of community coalescence[16,48].

Results

Diversity and composition of co-culture communities vary across mixture ratios but deviate from neutral predictions

We performed coalescence experiments *in vitro* using a set of diverse, stable communities of gut microbes. Following a previously established protocol[46], we derived these communities from stool samples of eight healthy human subjects and passaged them 15 times with a 1:200 dilution in fresh modified Brain Heart Infusion + mucin (mBHI+mucin) medium every 48 h (**Fig. 1a, Methods**). To assess community composition, we performed 16S rRNA gene amplicon sequencing on multiple passages from each community and tracked the relative abundances of amplicon sequence variants (ASVs, a rough taxonomic equivalent of species). Community diversity decreased in the first few passages after initial laboratory inoculation (**Fig. 1b, Extended Data Fig. 1a,b**), likely because some taxa in the inoculum were nonviable or were outcompeted during *in vitro* passaging. Nonetheless, diversity stabilized after ~3 passages, and most communities thereafter contained ASVs that comprised >50% of the relative abundance of the stool inoculum (**Extended Data Fig. 1c-f**). Thus, these stool-derived *in vitro* communities are diverse and stable models of stool microbiotas that can be used to study ecological interactions among gut microbes.

To systematically assess the effects of initial propagule size on colonization, we mixed eight pairs of *in vitro* “parent” communities at seven mixture ratios ranging from 1000:1 to 1:1000 (**Fig. 1c, Extended Data Fig. 2, Table S1**). We co-cultured mixtures in triplicate for five passages to allow composition to re-equilibrate, and we passaged two

replicates of each parent community alongside to allow for comparison. To assess the impact of propagule size on community composition and the colonization of individual species, we compared our data to the predictions of ecological neutral theory[15,18,21]. We computationally generated a set of theoretical co-cultures whose composition was determined by the propagule size of each species in the initial mixture (**Extended Data Fig. 3a-c**), and we quantified the similarity of our data to these theoretical communities.

Under neutral expectations, co-culture diversity should be maximal at the 1:1 mixture ratio. However, we found that experimental co-cultures consistently maintained diversities similar to or lower than one of the parent communities (**Fig. 1d, Extended Data Fig. 4**). This finding suggests that non-neutral competition occurs during community coalescence, preventing the neutral coexistence of all species. We also compared each experimental co-culture to its corresponding theoretical co-culture using the Jensen-Shannon divergence (JSD, **Methods**). Co-cultures consistently differed from the corresponding neutral theoretical expectation at every mixture ratio (**Fig. 1e, Extended Data Fig. 5**), as we observed with community diversity. Despite these deviations from neutral predictions, the composition of many, though not all, co-cultures varied substantially across mixture ratios (**Fig. 1e, Extended Data Fig. 5**), highlighting the continued impact of propagule size after ~40 generations of growth. These patterns of composition suggest that neither neutral theory nor consumer-resource models alone can fully predict the varied outcomes of community coalescence.

131 **Dose-dependent colonizers are present in every community mixture**

132 To investigate how the behaviors of individual species give rise to non-neutral but
133 propagule size-dependent patterns of community composition, we examined
134 colonization outcomes at the ASV level. We compared the relative abundance of each
135 ASV in experimental co-cultures after five passages to the neutral theoretical prediction
136 determined solely by its propagule size in the initial mixture (**Extended Data Fig. 6**). In
137 contrast to neutral predictions, many ASVs that started at higher abundance in one
138 parent community maintained a consistently high or low relative abundance across
139 mixture ratios in experimental co-cultures (*n*=32 (7.8%) of the 410 ASVs above the limit
140 of detection, **Methods**; **Fig. 2a, Extended Data Fig. 6, 7a,b**). We classified these dose-
141 independent ASVs as strong or weak colonizers depending on whether their
142 experimental relative abundance was greater or less than the neutral prediction,
143 respectively. For these strong and weak colonizers, deviations from neutral predictions
144 likely reflect competitive advantages or disadvantages that ultimately outweighed the
145 effects of propagule size.

146

147 In contrast, many of the remaining ASVs exhibited dose-dependent colonization, with
148 relative abundances that varied systematically across mixture ratios (*n*=43/410, 10.5%,
149 **Fig. 2b, Extended Data Fig. 6, 7c**). Dose-dependent colonizers were present in every
150 community mixture (**Fig. 2c**) and accounted for up to 20% of co-culture relative
151 abundance at the 1:1 mixture ratio (**Fig. 2d**). Most of the remaining abundance was
152 composed of resident ASVs, which were present at approximately the same relative
153 abundance in both parent communities and across all mixture ratios (*n*=175/410, 42.7%,

154 **Fig. 2e, Extended Data Fig. 6, 8a**), or by noisy ASVs whose relative abundance
155 changed non-monotonically across mixture ratios (*n*=160/410, 39.0%, **Fig. 2f,**
156 **Extended Data Fig. 6, 8b**). The prevalence of dose dependence among these varied
157 colonization behaviors demonstrates that propagule size can influence the colonization
158 of a sizeable fraction of species in each community during coalescence, even when
159 communities show substantial deviations from neutral composition (**Fig. 1e, Fig. 2c,d**).
160 Intriguingly, most ASVs exhibited distinct colonization patterns across co-cultures and
161 relative to closely related taxa (**Fig. 2g, Methods**), suggesting that these patterns were
162 influenced more strongly by community context than by species identity or phylogeny.
163 Taken together, these behaviors reveal that species in the same co-cultures can exhibit
164 varying levels of dependence on initial conditions.
165

166 **Transient dose dependence arises from high niche overlap in a consumer-**
167 **resource model**

168 We next sought to determine how the combination of neutral and non-neutral forces in
169 each mixture leads to dose-dependent colonization patterns. In the D1/D2 co-cultures,
170 we identified three dose-dependent *Lactococcus* and *Enterococcus* colonizers from the
171 D1 parent community. The relative abundances of these ASVs were approximately
172 anticorrelated with those of another dose-dependent *Lactococcus* from the D2
173 community (**Fig. 3a**). This observation led us to hypothesize that dose dependence
174 could arise from competition between phylogenetically related taxa that occupy
175 overlapping niches.
176

We used a standard consumer-resource model[19] to investigate the conditions under which competition for shared resources could give rise to neutral-like colonization (**Methods**). We first studied a system in which two species compete for a set of three common resources: μ_1, μ_2, and μ_{12} (**Fig. 3b**). Two of these resources, μ_1 and μ_2, are consumed exclusively by species 1 and 2, respectively, and the third resource, μ_{12}, is consumed by both species at equal rates ($R_{1,12} = R_{2,12}$). We investigated how the niche overlap of species 1 with species 2, defined as $\gamma_{12} = \frac{\mu_{12}}{\mu_1+\mu_{12}}$, would affect the colonization behavior of these two species. We varied γ_{12} from 0.025 to 1 while maintaining the same amount of unique resources for each species ($\mu_1 = \mu_2$). We initialized the abundances of the two species at ratios from 1000:1 to 1:1000, as in our community coalescence experiments, and we simulated the model to test how species abundances would change over five passages.

In this system, the dose dependence of both species increased with the amount of overlap between their niches (**Figure 3c**). For example, when niche overlap was low ($\gamma_{12} = 0.75$), the two species rapidly reached consistent relative abundances across mixture ratios (**Figure 3c**). By contrast, when niche overlap was high ($\gamma_{12} = 0.975$), the relative abundance of both species remained dose dependent, varying by ~20-fold between the 1000:1 and 1:1000 mixture ratios after three passages (**Figure 3c**), consistent with the magnitude of dose dependence that we often observed during community coalescence (**Fig. 2b, Fig. 3a**). Similar levels of dose dependence were present even when the consumption rate of the shared resource varied by ~10% relative to that of the unique resource (**Extended Data Fig. 9a**). These results show that

200 neutral-like colonization can arise when the growth of two species is dominated by their
201 consumption of shared resources.
202
203 Although these species showed dose dependence when niche overlap was high, the
204 magnitude of this dose dependence declined over time (**Fig. 3c,d**). For example, when
205 $\gamma_{12} = 0.975$, the dose dependence of the two species declined from ~20-fold to ~10-fold
206 from passage three to passage five (**Fig. 3c**). This declining dose dependence was
207 consistent with decreases in dose dependence over passages that we observed
208 experimentally for some ASVs during community coalescence (**Extended Data Fig.**
209 **9b**). Taken together, our model suggests that high niche overlap can slow a
210 community's convergence to equilibrium, creating transient, neutral-like colonization
211 dynamics.
212
213 **Strain isolates show less dose dependence in pairwise mixtures than during**
214 **community coalescence**
215 Based on the predictions of our consumer-resource model, we hypothesized that the
216 *Lactococcus* and *Enterococcus* species that showed dose-dependent colonization in the
217 D1/D2 co-cultures had high levels of niche overlap that would lead them to show
218 neutral-like colonization in pairwise mixtures as well. To test this prediction, we isolated
219 representatives of these ASVs (**Methods**). We performed Sanger sequencing of the
220 16S rRNA gene and whole-genome sequencing (**Methods**) to identify the strains from
221 the D1 community as *Lactococcus garvieae*, *Enterococcus faecalis*, and *Enterococcus*
222 *casseliflavus*, and the strain from the D2 community as *Lactococcus lactis* (**Fig. 4a**).

These strains exhibited similar maximum growth rates in monoculture, although small differences in growth dynamics suggested the potential for distinct profiles of resource utilization (**Fig. 4a**).

We performed a set of pairwise mixtures of *L. lactis* (from community D2) with the strains isolated from community D1 at ratios from 1000:1 to 1:1000 (**Fig. 4b-d**). Most strains exhibited some degree of dose-dependent colonization after three passages, suggesting that their niches overlapped to some degree (**Fig. 4b-d**). However, after five passages, each strain reached a constant relative abundance across mixture ratios (**Fig. 4b-d**). This lack of dose dependence contrasted with their behavior in the D1/D2 co-cultures, in which dose dependence was maintained even after five passages (**Fig. 3a**). This discrepancy suggests that pairwise strain mixtures are unlikely to exhibit the same magnitude and duration of neutral-like colonization that we observed during coalescence of more complex communities.

Dose dependence is enhanced in mixtures of more diverse communities

We used our consumer-resource model to investigate how colonization dynamics would be affected by the addition of a third species with overlapping niches. We introduced a third species into the model that consumes a unique resource initialized each passage at the same level as the other resources (i.e., $\mu_3 = \mu_2$, **Fig. 5a**). Species 3 also competes with species 1 for a fraction β of resource μ_1, which it consumes at relative rate r compared to species 1. We varied β while keeping r constant (**Fig. 5b**) and vice versa (**Extended Data Fig. 9c**), initializing the abundances of species 1 and 2 at a

range of ratios from 1000:1 to 1:1000 as before, and we set the abundance of species 3 equal to that of species 2 as if they were members of the same community. We simulated the model to investigate how species abundances changed over five passages.

In this three-species model, species 1 showed higher dose dependence as its niche overlap with species 3 increased (**Fig. 5b**). When species 3 had minimal niche overlap with species 1 ($\beta = 0.1$, **Fig. 5b**) or was a weak competitor ($r \ll 1$, **Extended Data Fig. 9c**), species 1 showed similar levels of dose dependence as when species 3 was absent (**Fig. 3c**), as expected. However, when niche overlap was high ($\beta > 0.5$) and species 3 was a strong competitor ($r \gg 1$), species 3 effectively outcompeted species 1 for the use of their shared resources (**Fig. 5b, Extended Data Fig. 9c**). Under these conditions, the dose dependence of species 1 increased substantially due to its increased reliance on the resource shared with species 2 (**Fig. 5b**, **Extended Data Fig. 9c**). At the same time, the dose dependence of species 2 was essentially unaffected by the amount of niche overlap between species 1 and 3, showing that the addition of new species does not necessarily alter the colonization behavior of all community members (**Fig. 5b, Extended Data Fig. 9c**). Together, these results predict that neutral-like colonization dynamics should be longer-lasting and more pronounced in diverse communities, in which competition with multiple species limits the unique resources available for any species.

268 To test this prediction, we introduced the *L. lactis* strain into an equal mixture of *E.*
269 *faecalis* and *E. casseliflavus* at ratios from 1000:1 to 1:1000. Consistent with our
270 prediction, after three passages, the degree of dose dependence exhibited by *L. lactis*
271 increased to ~200-fold (**Fig. 5c**) from ~60-fold in pairwise co-cultures (**Fig. 4b,c**), while
272 neither *E. faecalis* nor *E. casselifavus* exhibited dose-dependent colonization (**Fig. 5c**).
273 After five passages, none of the three strains exhibited dose dependence (**Fig. 5c**).
274 These results support our model's prediction that the addition of a new species can
275 create more pronounced neutral-like colonization dynamics, even if dose dependence
276 still declines more rapidly than during coalescence of complex communities.
277
278 To test the prediction that introducing strains into a more diverse community would
279 increase the magnitude and duration of dose dependence, we performed a set of strain-
280 community mixtures in which we mixed each of the two *Enterococcus* and *Lactococcus*
281 strains into the D1 or D2 communities at ratios from 1000:1 to 1:1000 (**Fig. 5d,e,**
282 **Extended Data Fig. 10**). In striking contrast to the pairwise and three-strain mixtures, a
283 subset of strains in each strain-community mixture showed dose dependence after five
284 passages (**Fig. 5d,e, Extended Data Fig. 10**). For instance, *E. casseliflavus* and *L.*
285 *lactis* were no longer dose dependent after the fifth passage in pairwise co-cultures
286 (**Fig. 4c**). However, when *L. lactis* was introduced into the D1 community, both strains
287 exhibited dose dependence after five passages, varying by ~10-fold in relative
288 abundance across all mixture ratios (**Fig. 5d**). Similarly, when *L. garvieae* was
289 introduced into the D2 community, it showed complementary patterns of dose
290 dependence with *L. lactis*, even though *E. casseliflavus* remained at constant relative

291 abundance across mixture ratios (**Fig. 5e, Extended Data Fig. 10c**). Together, these
292 results demonstrate that dose dependence can increase with community diversity.
293 These findings are consistent with the model prediction that the presence of additional
294 species increases niche overlap, prolonging the effect of initial propagule size and
295 creating neutral-like colonization dynamics.

Discussion

In this study, we quantified the effect of propagule size on hundreds of simultaneous colonization events during community coalescence, and we compared the resulting co-cultures to neutral ecological predictions. Community composition deviated systematically from neutral predictions, but initial population size had a strong effect on colonization for a subset of species in each mixture, even after ~40 generations of competition. Using a consumer-resource model, we showed that this dose-dependent colonization can arise from non-neutral competition for shared resources. These neutral-like dynamics frequently emerge in diverse communities due to high niche overlap among community members, even in the absence of more complex phenomena like resource fluctuations[49,50], spatial structure[51,52], and environmental modification[7,14].

Our model provides an explanation for how neutral-like colonization can occur in a community dominated by non-neutral competition. When considering either neutral theory[18,53] or resource competition alone[49,54], all species in a community are expected to show the same colonization dynamics. However, we observed both neutral-like and non-neutral colonization dynamics in each community mixture. Our model explains this co-occurrence by showing that neutral-like behavior is neither an intrinsic property of a species, nor of the community as a whole. Instead, neutral-like dynamics are emergent phenomena[55–58] that arise from non-neutral competitive interactions among subsets of community members.

Another feature of our model is that neutral-like colonization dynamics are transient: communities eventually converge to a stable equilibrium, as in classical consumer-resource models[19]. However, both our model and experiments show that neutral-like dynamics can persist for ~40 generations. The timescales of these transient states[59–61] may be long enough for introduced species to affect community composition and function[60,62,63], induce shifts between alternative states[50,64], and alter the outcomes of subsequent species introductions[14,65–68]—even if the initial species introductions are ultimately unsuccessful.

Our findings have several implications for the design of ecological interventions like microbiome therapeutics[35] and agricultural inoculants[69]. We showed that introduced species often have higher levels of dose dependence in diverse communities, meaning that experiments with small, model communities may systematically underestimate the effects of propagule size in more diverse, natural contexts. Our model also explains why the same species can show varying levels of dose dependence in different community contexts, which may explain why previous studies have come to contradictory conclusions about how propagule size affects colonization success[5,20,26,30,33,34,37].

This work identifies niche overlap as the main reason propagule size has stronger impacts on colonization in more diverse communities. While niche overlap is challenging to measure, it is often higher among closely related species, which tend to consume more similar resources[70–72]. We might therefore expect higher levels of dose dependence when introduced species are taxonomically related to existing residents, as

341 our *Lactococcus* and *Enterococcus* example illustrates. However, more precise
342 methods to quantify resource utilization, such as metabolomics[22,71,73,74], may be
343 necessary to accurately predict niche overlap in microbial communities. Future studies
344 that combine controlled colonization experiments with metabolomics of microbial
345 isolates will improve our ability to predict how neutral forces like propagule size impact
346 colonization.

347 **Materials and Methods**

348

349 **Data availability**

350 Sequencing data will be made publicly available from the NCBI SRA upon manuscript

351 publication.

352

353 **Code availability**

354 The computer code that performs the analysis is available at

355 https://github.com/DoranG1/dose-dependent-colonization.

356

357 **Collection of stool samples from human subjects**

358 Stool samples were collected as a part of a study[75] approved by the Stanford University

359 Institutional Review Board under Protocol 54715. Informed written consent was

360 obtained from all participants. Samples were provided by healthy, non-pregnant adults

361 living with at least one roommate or housemate. At least two participants were recruited

362 to provide samples from each household, one of whom took a five-day course of the

363 antibiotic ciprofloxacin during sampling. Subjects collected stool samples in sterile

364 plastic tubes and froze them immediately at -20°C inside multiple layers of insulating

365 protection. Samples were transported to the laboratory on dry ice within six weeks of

366 collection and were stored at -80°C. For this study, we derived *in vitro* communities from

367 the last pre-antibiotic stool sample collected by eight study subjects from four household

368 pairs. These eight subjects lived in the San Francisco Bay Area and ranged from 25-30

369 years old.

370

371 **Derivation of top-down *in vitro* communities**

372 To obtain stool fragments for inoculation of *in vitro* communities, we used a Basque
373 Engineering CXT 353 Benchtop Frozen Aliquotter to drill cores from the frozen stool
374 samples of eight study subjects: A1, A2, B1, B2, C1, C2, D1, and D2 (**Table S1**). Cores
375 were kept frozen by immersing the drill area in liquid nitrogen and were stored at -80 °C
376 after drilling.

377

378 We generated *in vitro* communities following methods established in our previous
379 studies[46,47]. All steps of community inoculation and passaging were performed in an
380 anaerobic chamber (Coy Instruments), using sterilized pipette tips. First, a 50-mg
381 fragment of each stool core was resuspended in 500 µL of filter-sterilized PBS. Next, 20
382 µL of the stool resuspension were added to 180 µL of Brain Heart Infusion (BHI, BD
383 Biosciences 237200) supplemented with 0.2 mg/mL L-tryptophan, 1 mg/mL L-arginine,
384 0.5 mg/mL L-cysteine, 5 µg/mL vitamin K, 1 µg/mL haemin, and 5 g/mL mucin (referred
385 to here as mBHI+mucin) in a flat-bottom 96-well plate (Greiner Bio-One 655161). Each
386 sample was inoculated as three replicates in separate wells. All media were autoclaved
387 before inoculation to ensure sterility.

388

389 After inoculation, communities were grown at 37 °C in the anaerobic chamber and
390 passaged every 48 h by transferring 1 µL of saturated culture into a new 96-well plate
391 containing 199 µL of fresh mBHI+mucin. We passaged each community 15 times and

392 stored each plate of saturated cultures at -80 °C after passaging by mixing 100 μL of
393 saturated culture with 100 μL of 50% filter-sterilized glycerol.
394
395 **Whole community mixtures**
396 We thawed glycerol stocks of one replicate of each community from passage 15 and
397 reinoculated these cultures in an anaerobic chamber (Coy Instruments) by adding 3 μL
398 of glycerol stock to 197 μL of mBHI+mucin. These communities were grown at 37 °C for
399 48 h.
400
401 To create community mixtures, we first generated a dilution series of each of the
402 unmixed parent communities. Fresh mBHI+mucin was added to each saturated culture
403 to create communities with concentrations 1/10, 1/100, and 1/1000 of the parent
404 community. We generated each dilution in triplicate from the single replicate of each
405 parent community, resulting in 24 communities for each dilution factor from the eight
406 parent communities, and 72 dilution communities in total across all three dilution factors.
407
408 To generate each pairwise mixture, we added 1 μL of each dilution community to 1 μL
409 of another parent community and 198 μL of fresh mBHI+mucin. We mixed communities
410 in eight combinations: A1/A2, B1/B2, C1/C2, D1/D2, A1/B1, B1/C1, C1/D1, and D1/A1.
411 Four pairs of communities (A1/A2, B1/B2, C1/C2, D1/D2) were derived from subjects
412 from the same household, while the other four mixtures (A1/B1, B1/C1, C1/D1, D1/A1)
413 involved communities from different households. Each pair of communities was mixed
414 at seven ratios: 1000:1, 100:1, 10:1, 1:1, 1:10, 1:100, and 1:1000. Each mixture was

415 passaged in triplicate. This protocol resulted in 168 co-culture communities (8 mixtures
416 × 7 ratios × 3 replicates), which were passaged in two flat-bottom 96 well plates. We
417 also inoculated 2 μL of each undiluted parent community in 198 μL of mBHI+mucin in
418 duplicate, comprising 16 unmixed control parent communities. The four corners of each
419 96-well plate were filled with 200 μL of mBHI+mucin as blank controls to monitor
420 contamination.
421
422 Both 96-well plates containing mixture co-cultures and unmixed control communities
423 were grown at 37 °C and passaged every 48 h five times by adding 1 μL of saturated
424 culture to 199 μL of mBHI+mucin. Saturated cultures from each passage were stored at
425 -80 °C.
426
427 **16S rRNA gene sequencing**
428 DNA was extracted from 50 μL of each community from each mixture using the Qiagen
429 DNeasy Ultraclean 96 Microbial Kit (Qiagen 10196-4) according to the manufacturer's
430 instructions. Extracted DNA was stored at -20 °C.
431
432 We performed 16S rRNA gene amplicon sequencing targeted to the V4 region.
433 Extracted DNA was amplified using primers modified from the Earth Microbiome Project
434 spanning 515F–806R, with gene-specific sequences GTGYCAGCMGCCGCGGTAA for
435 the forward primer and GGACTACNVGGGTWTCTAAT for the reverse primer[76,77]. The
436 16S region was amplified using 27 cycles of PCR with an annealing temperature of 50
437 °C, and successful amplification of the 16S region was confirmed using gel

438 electrophoresis. Sequencing primers were attached using 10 cycles of PCR with an
439 annealing temperature of 54 °C. Excess primer was removed using a 0.8X Ampure
440 bead cleanup. Libraries were pooled to contain equal volume from each sample and
441 sequenced using an Illumina NovaSeq SP at the DNA Services Lab, Roy J. Carver
442 Biotechnology Center, University of Illinois at Urbana-Champaign.
443
444 Raw reads were annotated and demultiplexed using UMI-tools[78], and primer and
445 adapter sequences were trimmed using cutadapt[79]. To remove the blank media
446 controls, samples with <100 reads were excluded. DADA2[80] was used to filter and
447 truncate reads and to assign amplicon sequence variant (ASV) taxonomy based on the
448 SILVA (release 138) reference database[81], and to obtain a phylogeny of all ASVs.
449
450 Our final community coalescence dataset contained 91,504,441 reads with a median of
451 201,425 reads (standard deviation 69,902) per non-blank sample. Two samples in our
452 dataset (A1 replicate 1 from passage 3, and D1/A1 100:1 replicate 1 from passage 5)
453 received zero sequencing coverage, so these were excluded prior to subsequent
454 analyses. 1715 unique ASVs were detected across all samples, and 198 of these were
455 present at $\geq 10^{-3}$ relative abundance in at least one sample, which we estimated as the
456 limit of detection. Although subjects from the same households often shared a higher
457 proportion of gut microbial strains compared to subjects from different households[75], we
458 observed no differences in the proportion of ASVs shared in mixtures of communities
459 from the same versus different households (**Extended Data Fig. 3**). We used the
460 relative abundance of each ASV to calculate Jensen-Shannon divergence.

461

462 In addition to our community coalescence dataset, we obtained a second dataset
463 representing the *L. garvieae*/*L. lactis* co-cultures after five passages and the D2/*L. lactis*
464 co-cultures after five passages. This dataset contained 48,347,562 reads with a median
465 of 683,158 reads (standard deviation 370,504) per non-blank sample. We also obtained
466 a third dataset representing the *E. faecalis*/*L. lactis*, *E. casseliflavus*/*L. lactis*, three-
467 strain, and all strain-community co-cultures after three and five passages. This third
468 dataset contained 9,429,310 reads with a median of 12,927 reads (standard deviation
469 12,375) per non-blank sample. We excluded one replicate each from the *L. garvieae*/*L.*
470 *lactis*, D1/*L. lactis, L. garvieae*/D2, and *E. casseliflavus*/D2 (passage 5) mixtures based
471 on poor sequencing coverage at one or more mixture ratios in those replicates.

472

473 **Estimation of theoretical community compositions**

474 To assess the effect of propagule size in our community mixtures, we computationally
475 generated a set of theoretical community mixtures according to the predictions of
476 ecological neutral theory[18]. The expected abundance of each ASV across mixture ratios
477 was calculated based on a weighted average of the relative abundance in the parent
478 communities: $\frac{x_1}{A} + \frac{x_2}{B}$, where x_1 and x_2 represent the relative abundance of the ASV in
479 parent community 1 and 2, respectively, and A and B represent the dilution factor of
480 parent community 1 and 2, respectively, at the given mixture ratio. For example, the
481 predicted abundance of an ASV in a 1:100 mixture of A1 and A2 in which the propagule
482 size of A2 is 100 times greater than that of A1 is $\frac{x_{A1}}{100} + x_{A2}$. We normalized these
483 weighted averages to calculate a theoretical relative abundance for each ASV in a

484 mixture, and we ignored ASVs with predicted relative abundance $<10^{-3}$, our sequencing
485 limit of detection. We generated two theoretical communities for each pair of parent
486 communities based on the two replicates of each parent community in our experiments.
487 We generated separate sets of theoretical communities from passages 3 and 5 of the
488 parent communities to account for changes to parent community composition that may
489 have occurred during passaging.
490

491 **Categorization of ASVs by colonization behavior**

492 We developed a set of summary statistics for each ASV based on its average change in
493 relative abundance across mixture ratios and the deviation from predicted neutral
494 relative abundances, which enabled us to systematically categorize each ASV into one
495 of six distinct patterns: low abundance, noisy, resident, dose-dependent, strong
496 colonizer, or weak colonizer (**Extended Data Fig. 5**).
497

498 For each ASV in each mixture, we calculated four measures. The average $\log_{10}$ fold
499 change in relative abundance across mixture ratios (net dose difference) and average
500 absolute value of $\log_{10}$ fold change in relative abundance across mixture ratios
501 (magnitude dose difference) were used to quantify each ASV's level of dose
502 dependence. The average $\log_{10}$ fold difference between experimental and theoretical
503 relative abundances across mixture ratios (net theoretical difference) and average
504 absolute value of $\log_{10}$ fold difference between experimental and theoretical relative
505 abundances across mixture ratios (magnitude theoretical difference) were used to

506 quantify each ASV's deviation from neutral predictions and distinguish between
507 resident, strong colonizing, and weak colonizing ASVs.
508
509 Net dose difference was calculated as $\frac{\sum(x_i - x_{i-1})}{n}$, where x_i represents the $\log_{10}$(relative
510 abundance) of an ASV at a single mixture ratio (e.g., 100:1), x_{i-1} represents the
511 $\log_{10}$(relative abundance) at the preceding mixture ratio (e.g., 1000:1), and $n = 7$
512 represents the total number of mixture ratios. Magnitude dose difference was calculated
513 as $\frac{\sum|x_i - x_{i-1}|}{n}$. Both net dose difference and magnitude dose difference were first
514 calculated individually for each replicate of the ASV and then averaged across
515 replicates.
516
517 Net theoretical difference was calculated as $\frac{\sum(\bar{x}_a - \bar{x}_t)}{n_{\text{ratio}}}$, where $\bar{x}_a$ and $\bar{x}_t$ represent the
518 average experimental and theoretical, respectively, $\log_{10}$(relative abundance) across
519 replicates at a single mixture ratio. Magnitude theoretical difference was calculated as
520 $\frac{\sum|\bar{x}_a - \bar{x}_t|}{n_{\text{ratio}}}$. For ASVs that were not detected in the co-cultures at some mixture ratios
521 and/or parent communities, we set the relative abundance to 10^{-3}.
522
523 To categorize ASVs, we used five binary decisions based on our summary statistics and
524 other measures (**Extended Data Fig. 5**). First, ASVs were classified as "low
525 abundance" if their relative abundance was $<10^{-2.6}$, a threshold slightly above our limit of
526 detection, in both parent communities. Next, ASVs were classified as "noisy" if the ratio
527 of magnitude dose difference to the absolute value of net dose difference was >2.25,

528 indicating large, non-monotonic changes in relative abundance over mixture ratios. To
529 prevent ASVs with extremely low change over mixture ratios (absolute value net dose
530 difference <0.1) from artificially inflating the number of noisy ASVs, we replaced these
531 low values with 0.1 as a pseudocount.
532
533 If ASVs were not classified as low abundance or noisy, they were classified as "dose-
534 dependent colonizers" if the absolute value of the net dose difference was >0.09,
535 meaning that they exhibited large monotonic changes in relative abundance over
536 mixture ratios. Of the remaining ASVs, if the magnitude theoretical difference was >0.3
537 and the absolute value of the net theoretical difference was ≥0.2, indicating a strong
538 deviation from neutral predictions, they were classified as "strong colonizers" or "weak
539 colonizers" based on whether the sign of net theoretical difference was positive or
540 negative, respectively. All other ASVs, which showed little dose dependence or
541 deviation from neutral predictions, were classified as "resident". All numeric cutoffs
542 between categories were chosen to minimize the number of ASVs classified as low
543 abundance, noisy, resident, or low abundance, while ensuring that dose-dependent,
544 strong, and weak colonizers could be readily distinguished by inspection.
545
546 **Consumer-resource model**
547 We implemented a standard consumer-resource model[19]:
548

$$\frac{dX_i}{dt} = X_i \sum_{\mu=1}^{M} R_{i,\mu} Y_\mu$$

$$\frac{dY_\mu}{dt} = -Y_\mu \sum_{i=1}^{N} R_{i,\mu} X_i$$

where X_i denotes the absolute abundance of species i, Y_μ the amount of resource μ, and $R_{i,\mu}$ the consumption rate of resource μ by species i. Species abundances were simulated in Matlab (Mathworks, Inc.) using the ode45 solver until all resources were depleted. After each such passage, the abundances were diluted 1:200, and resource concentrations were refreshed. We initialized the abundances of the two species at a range of ratios from 1000:1 to 1:1000, as in our community coalescence experiments, and we simulated five passages for each community. We defined dose dependence as the ratio of a species' relative abundances at starting ratios of 1000:1 and 1:1000.

We first investigated a simple system in which two species, denoted 1 and 2, compete to consume use a set of three common resources: μ_1, μ_2, and μ_{12} (**Fig. 3b**). Two of these resources, μ_1 and μ_2, are consumed exclusively by species 1 and 2 respectively, and the third resource, μ_{12}, is consumed by both species at equal rates ($R_{1,12} = R_{2,12} = 1$). To investigate how the niche overlap of species 1 with species 2, defined as $\gamma_{12} = \frac{\mu_{12}}{\mu_1+\mu_{12}}$, would affect the colonization behavior of these two species (**Fig. 3c**), we varied γ_{12} from 0.025 to 1 while maintaining the same amount of unique resources for each species $\mu_1 = \mu_2 = 1$. To test how varying the rates of shared resource consumption affects colonization behavior, we varied $R_{1,12}$ from 0.9 to 1.1 while maintaining a constant $R_{2,12} = 1$ (**Extended Data Fig. 8a**).

570 We investigated how dose dependence would be affected by the addition of a third
571 species. We introduced a third species into the model that consumes a unique
572 resource, $\mu_3 = \mu_2 = 1$. Species 3 also competes with species 1 for a fraction β of
573 resource μ_1, which it consumes at relative rate $r = \frac{R_{3,13}}{R_{1,13}}$ compared to species 1. To
574 investigate how the presence of this new species affects colonization, we varied β from
575 0 to 0.9 while maintaining a constant $r = 10$ (**Fig. 5b**). As before, we initialized the
576 abundances of species 1 and 2 at a range of ratios from 1000:1 to 1:1000, and we set
577 the abundance of species 3 equal to that of species 2. To investigate how the
578 consumption rate of this shared resource affects colonization behavior, we also tested
579 the effect of varying r from 10^{-2} to 10^{2} while maintaining a constant $\beta = 0.5$ (**Extended**
580 **Data Fig. 8d**).
581

582 **Strain isolation**

583 *Lactococcus garvieae*, *Lactococcus lactis*, *Enterococcus faecalis*, and *Enterococcus*
584 *casseliflavus* strains representing dose-dependent ASVs were isolated by sorting single
585 cells from communities D1 and D2 into mBHI-filled wells of a 96-well plate using a
586 fluorescence-activated cell sorter (FACS)[82]. To do so, communities were transported
587 from the anaerobic chamber to the FACS facility in an anaerobic container (BD
588 Biosciences) to ensure minimal oxygen exposure. We confirmed the purity of each
589 strain isolate by streaking multiple times on BHI agar plates supplemented with 5%
590 horse blood and submitted individual colonies for Sanger and whole-genome
591 sequencing. Colonies verified as pure were grown for 48 h in 200 µL of mBHI+mucin in

592 a 96-well plate, and 75 µL of each saturated culture were mixed with 75 µL of 50%
593 glycerol and stocked in plastic crimp vials (ThermoFisher C4011-11) for future use.
594
595 To confirm isolate purity, we performed Sanger sequencing targeted to the V4 region of
596 the 16S rRNA gene. We amplified the 16S gene with 35 cycles of PCR with an
597 annealing temperature of 55 °C using the previously described primers. The resulting
598 PCR products were sent to Elim Biopharmaceuticals for sequencing.
599
600 We determined the species identity of strain isolates using whole-genome sequencing.
601 DNA was extracted from isolate cultures using the DNeasy Ultraclean 96 Microbial Kit
602 (Qiagen 10196-4) according to the manufacturer's instructions. Libraries were prepared
603 for whole-genome sequencing using the Nextera DNA Flex Library Prep Kit (Illumina
604 20018705) according to the manufacturer's instructions. The DNA concentration of each
605 library was quantified using the Quant-iT High-Sensitivity dsDNA Assay Kit (Invitrogen
606 Q33130) and a Synergy HT Microplate Reader (Biotek Instruments), and libraries were
607 pooled at equal DNA input. Pooled libraries were sequenced at the DNA Services Lab,
608 Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign.
609
610 **Co-cultures of strain isolates and strain-community mixtures**
611 We performed a set of four mixtures of two or three strains (*L. garvieae*/*L. lactis*, *E.*
612 *faecalis*/*L. lactis*, *E. casseliflavus*/*L. lactis*, and *E. faecalis*/*E. casseliflavus*/*L. lactis*) and
613 five strain-community mixtures (D1/*L. lactis*, *L. garvieae*/D2, *E. faecalis*/D2, *E.*
614 *casseliflavus*/D2, *E. faecalis*/*E. casseliflavus*/D2) using the glycerol stocks of our four

615 strain isolates and of the D1 and D2 parent communities. We performed mixtures at
616 ratios from 1000:1 to 1:1000 in triplicate as described above for whole community
617 mixtures. For mixtures that included both *E. faecalis* and *E. casseliflavus*, 0.5 μL of each
618 strain culture were used. Mixtures were passaged in mBHI+mucin for 48 h at 37 °C five
619 times.
620

621 **Analysis of pairwise mixture composition using Sanger sequencing data**

622 To analyze the composition of co-cultures from the *L. garvieae*/*L. lactis* mixture after
623 three passages, we performed Sanger sequencing of each co-culture as described
624 above and used the CASEU R package[83] to quantify the relative abundance of both
625 strains at each mixture ratio.
626

627 **Quantification of strain isolate growth**

628 Growth curves were obtained using an Epoch 2 Microplate Spectrophotometer (Biotek
629 Instruments). Strains were grown overnight and then diluted 1:10, 1:100, or 1:1000 into
630 fresh BHI+mucin in 96-well plates. We measured the optical density at 600 nm (OD_{600})
631 of 10 replicates of each strain over 48 h of anaerobic growth at 37 °C with continuous
632 orbital shaking. Instantaneous growth rate was quantified as $d\ \ln(OD_{600})/dt$ using the
633 gcplyr R package[84].

634 **Author contributions**

635 D.A.G., K.S.X., D.A.R., and K.C.H. designed the research; D.A.G., K.S.X., A.B.P.,
636 R.R.J., and L.R.F. performed the research. D.A.G., K.S.X., J.G.L., J.C.C.V., D.A.P.,
637 B.H.G., D.A.R., and K.C.H. analyzed the data; D.A.G., K.S.X., D.A.R., and K.C.H. wrote
638 the paper; and all authors reviewed it before submission.

Acknowledgements

We thank Mark Bitter, Tadashi Fukami, and members of the Huang, Relman, Good, and Petrov labs for helpful discussions. Sequencing support for this project was provided by the DNA Services Lab, Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign. This work was funded by Stanford Bio-X Undergraduate Summer Research Program fellowships (to D.A.G. and A.B.P.), Stanford Vice Provost for Undergraduate Education Small Grants (to D.A.G. and A.B.P.), a Carol Carmichael Summer Undergraduate Research Fellowship (to R.R.J.), NSF Awards IOS-2032985 and EF-2125383 (to K.C.H.), NIH Awards R01 AI147023 (to D.A.R. and K.C.H.) and RM1 GM135102 (to K.C.H.), the Thomas C. and Joan M. Merigan Endowment at Stanford University (to D.A.R.), and NIH/NIAID Award R21-AI168860 (to D.A.R.). J.G.L. is supported by the PRISM Baker Fellowship. K.S.X. has been supported by a James McDonnell Foundation Postdoctoral Fellowship in Understanding Dynamic and Multi-Scale Systems and a Jane Coffin Childs Memorial Fund Postdoctoral Fellowship. K.C.H., B.H.G., and D.A.P. are Chan Zuckerberg Biohub Investigators.

Figures

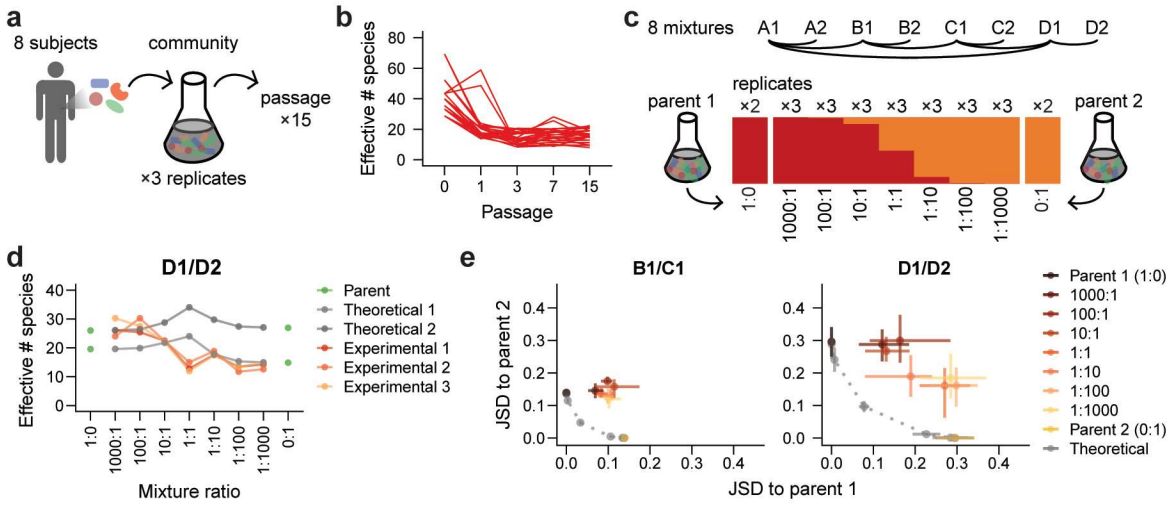


Figure 1: Diversity and composition of co-cultures vary by mixture ratio and deviate from neutral predictions.

a) *In vitro* communities were inoculated in triplicate from stool samples collected from eight healthy human subjects and passaged 15 times to reach stability.

b) The diversity of each community over time, quantified as the effective number of species ($e^{H'}$) calculated from the Shannon diversity index (H'), initially decreased and then plateaued by passage 3.

c) Experimental design: eight pairs of *in vitro* communities were mixed in triplicate at mixture ratios ranging from 1000:1 to 1:1000. The resulting co-cultures were passaged five times.

d) The diversity of co-culture communities was similar to or lower than the diversity of parent communities across mixture ratios. Gray lines represent both replicates of predicted neutral mixtures for an example mixture based on the diversity of replicate parent communities (1:0 and 0:1, shown in green), and orange lines represent three

replicates of experimental community mixtures. Diversity was quantified by the effective number of species ($e^{H'}$) calculated from the Shannon diversity index (H').

e) Variation of community composition across example mixtures and mixture ratios. Each plot shows the Jensen-Shannon divergence (JSD) of all co-culture and theoretical communities relative to both parent communities (brown and gold) for a single mixture. Grey points along the dotted line represent theoretical mixtures, and colored points represent data from experimental co-culture communities. Points show averages and error bars show full range of values across inoculation replicates.

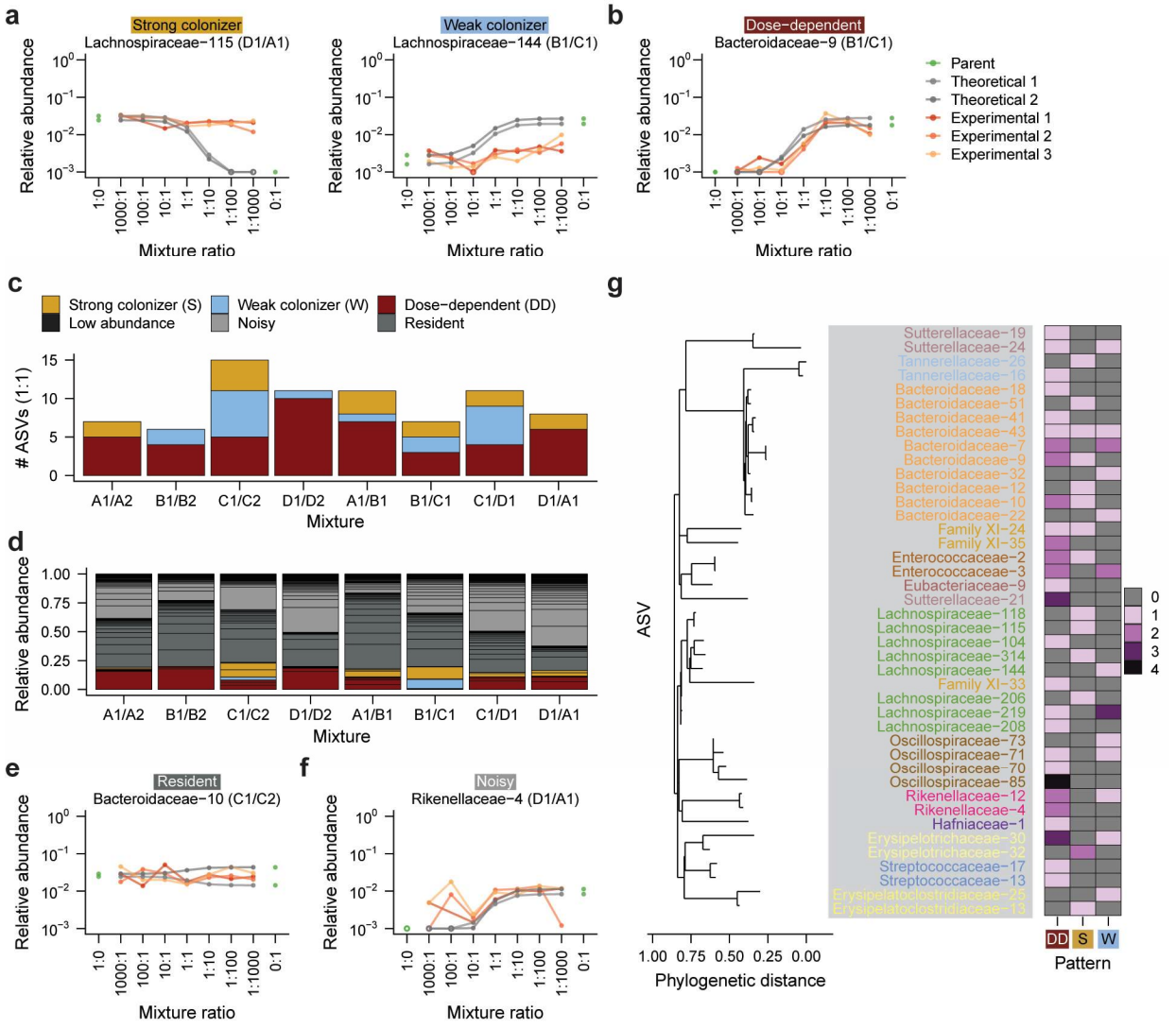


Figure 2: Colonization behaviors of individual ASVs influence community composition.

a) Representative examples of strong and weak colonizers. Each panel shows an individual ASV from one mixture. Gray lines represent both replicates of predicted neutral relative abundances for each ASV, and orange lines represent three replicates of experimental relative abundances.

b) Dose-dependent colonizers exhibit large changes in relative abundance across mixture ratios. Lines are colored as in (**a**).

688 c) The number of ASVs that displayed a strong, weak, or dose-dependent colonization
689 pattern in each set of mixtures. ASVs were counted as present if they were detected
690 in the 1:1 mixture at the fifth passage.
691 d) Total relative abundance of each type of colonizer in each set of mixtures. Only the
692 1:1 mixture is shown. Black lines indicate separate ASVs. Relative abundances are
693 averaged across the three inoculation replicates.
694 e) A representative example of a resident ASV that was present at similar relative
695 abundances in both parent communities. Lines are colored as in (**a**).
696 f) A representative example of a noisy ASV whose relative abundance showed large,
697 non-monotonic fluctuations between adjacent mixture ratios. Lines are colored as in
698 (**a**).
699 g) Colonization behavior is not associated with ASV phylogeny. The heatmap shows
700 the number of times each ASV exhibited a dose-dependent (DD), strong (S), or
701 weak (W) colonization pattern in any mixture.

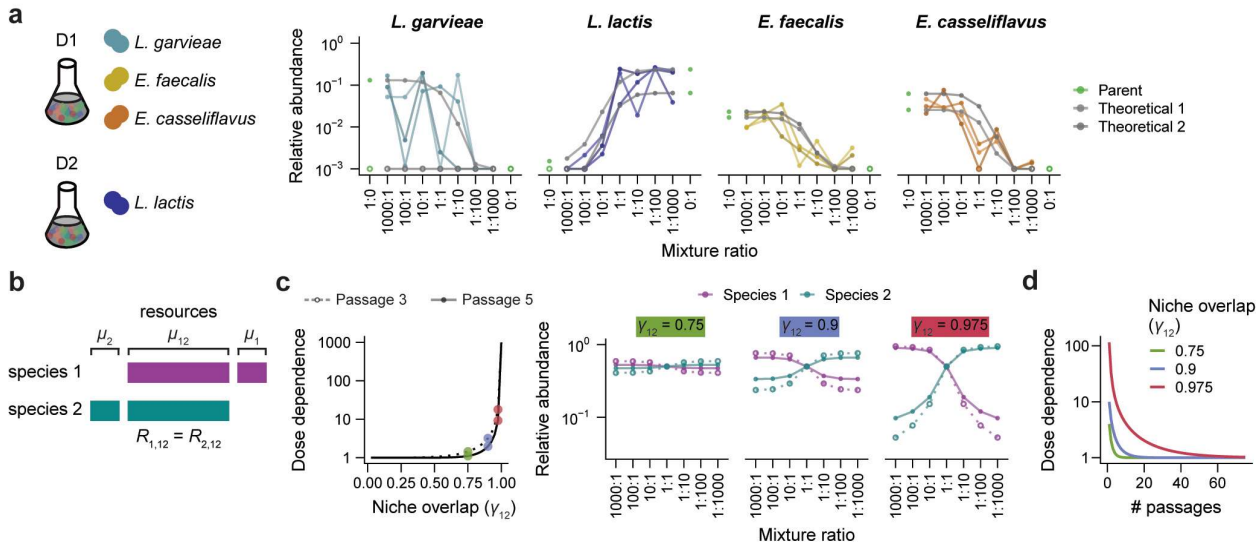


Figure 3: A consumer-resource model predicts transient dose dependence.

a) ASVs from the *Enterococcus* and *Lactococcus* genera showed dose dependence during community coalescence. *Enterococcus faecalis*, *Enterococcus casseliflavus*, and *Lactococcus lactis* were isolated from the D1 community, and *Lactococcus lactis* was isolated from the D2 community. Gray lines represent both replicates of predicted neutral relative abundances for each ASV after five passages, and colored lines represent three replicates of experimental relative abundances after five passages.

b) Schematic of a simple consumer-resource model in which two species compete neutrally (equal consumption rates $R_{1,12} = R_{2,12} = 1$) for resource μ_{12}, and resources μ_1 and μ_2 are unique to species 1 and 2, respectively (**Methods**).

c) The dose dependence of both species in the consumer-resource model increases as the niche overlap of species 1 with species 2 (γ_{12}) increases. Dose dependence also decreases from passage 3 (dotted line) to passage 5 (solid line). Colored points highlight three values of γ_{12} for which the relative abundances over mixture ratios of

each species are shown at right. Dose dependence is defined as the ratio of relative abundances after passaging starting from a 1000:1 versus 1:1000 mixture ratio.

d) Dose dependence declines over passages in the two-species consumer resource model, even when niche overlap (γ_{12}) is high.

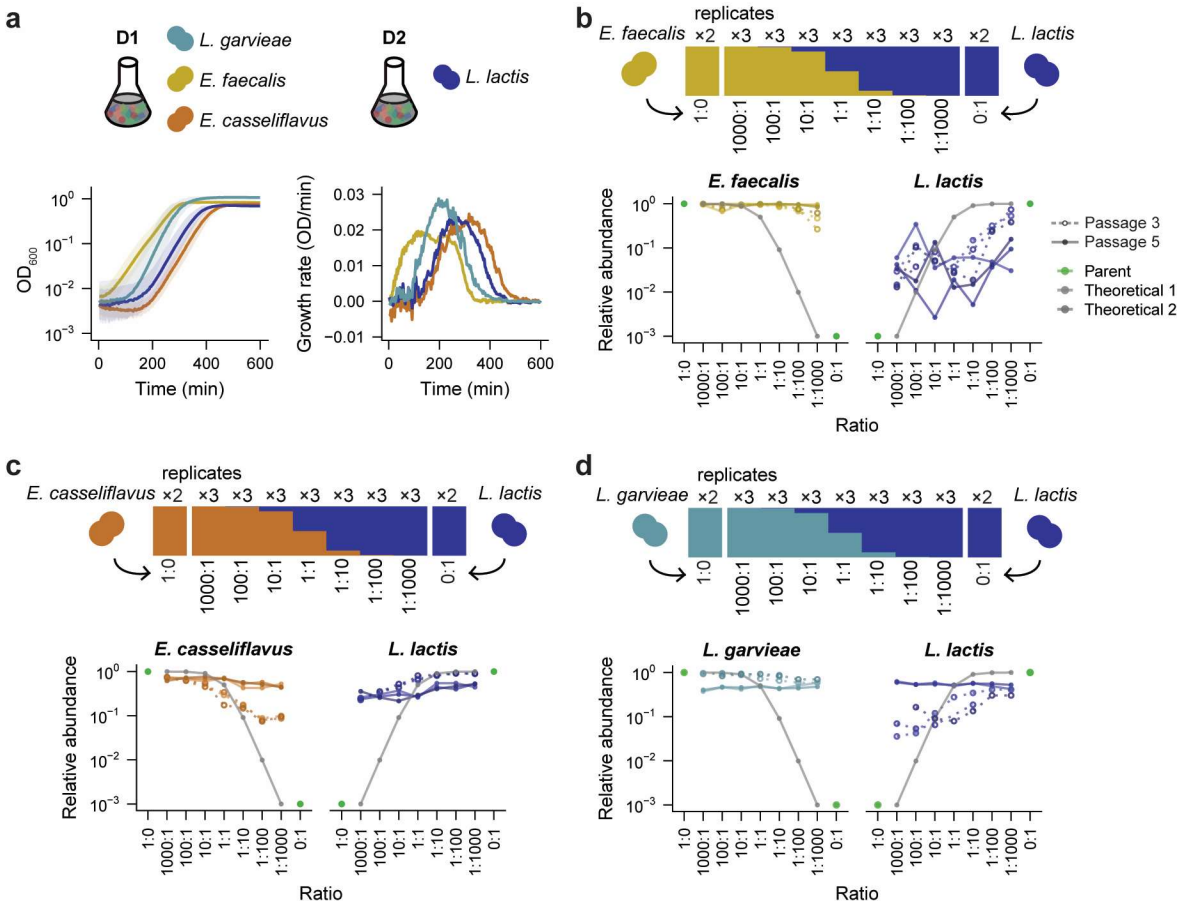


Figure 4: Transient dose-dependent colonization in pairwise strain mixtures.

a) Growth of the *Enterococcus* and *Lactococcus* strains in monoculture. Instataneous growth rates (bottom right) were calculated from blank-subtracted growth curves (bottom left) after smoothing over a 5-min window (**Methods**).

b-d) Species relative abundances in co-cultures of (b) *E. faecalis* and *L. lactis*, (c) *E. casseliflavus* and *L. lactis*, and (d) *L. garvieae* and *L. lactis*. Lines are colored as in **Fig. 3a**. Dose dependence decreased substantially from passage three (dotted lines) to passage five (solid lines).

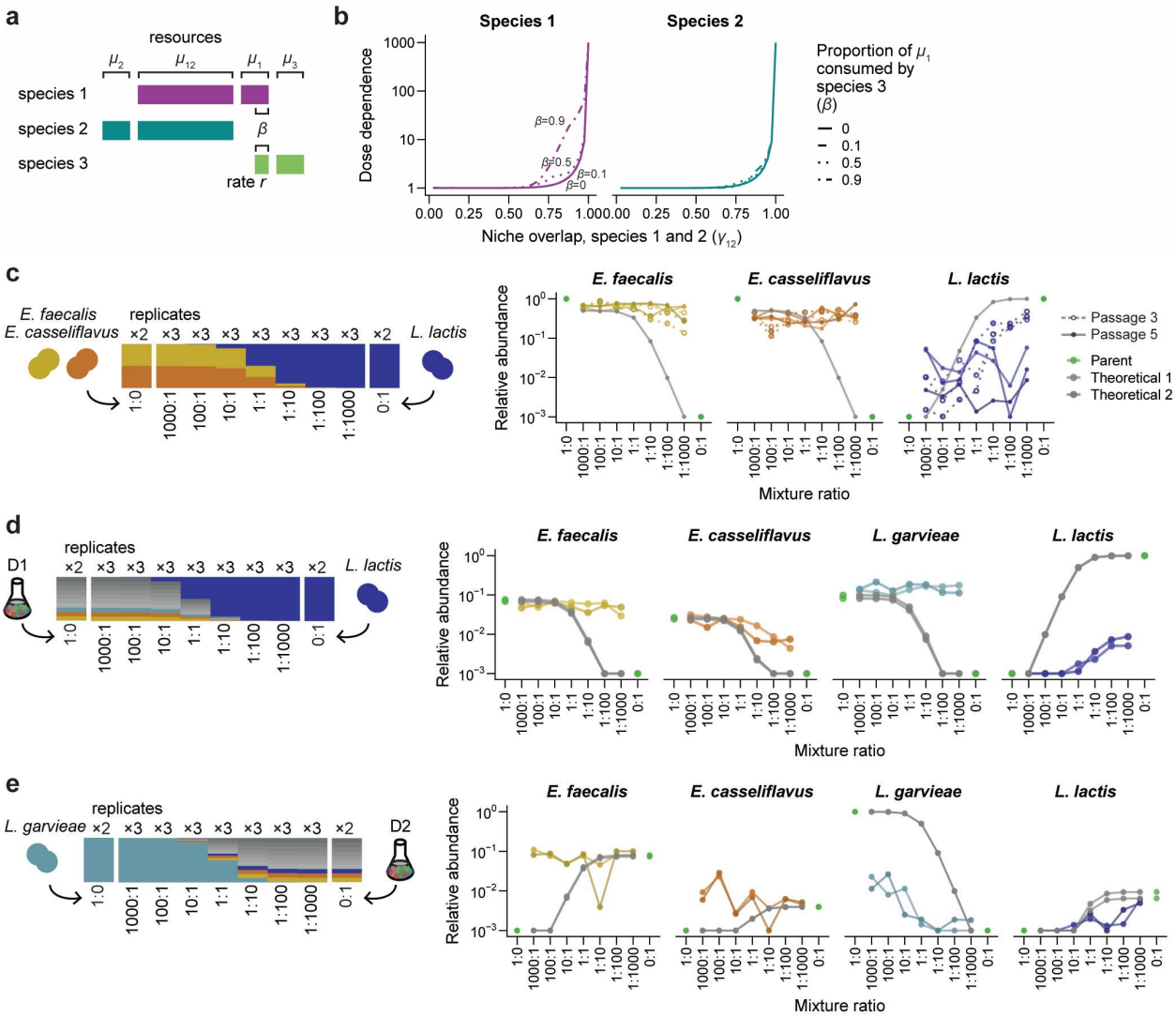


Figure 5: Dose dependence increases in mixtures with more diverse communities.

a) Modification to the consumer-resource model in **Fig. 3a** in which a third species competes with species 1 for a fraction β of resource μ_1 with relative consumption rate r compared to species 1.

b) The dose dependence of species 1 after five passages increases with increasing niche overlap between species 1 and 3 (β) at constant $r = 10$, while the dose

739 dependence of species 2 is largely unaffected. Dose dependence is defined as in

740 **Fig. 3c**.

741 c-e) Species relative abundances in co-cultures of (c) *E. faecalis, E. casseliflavus*, and

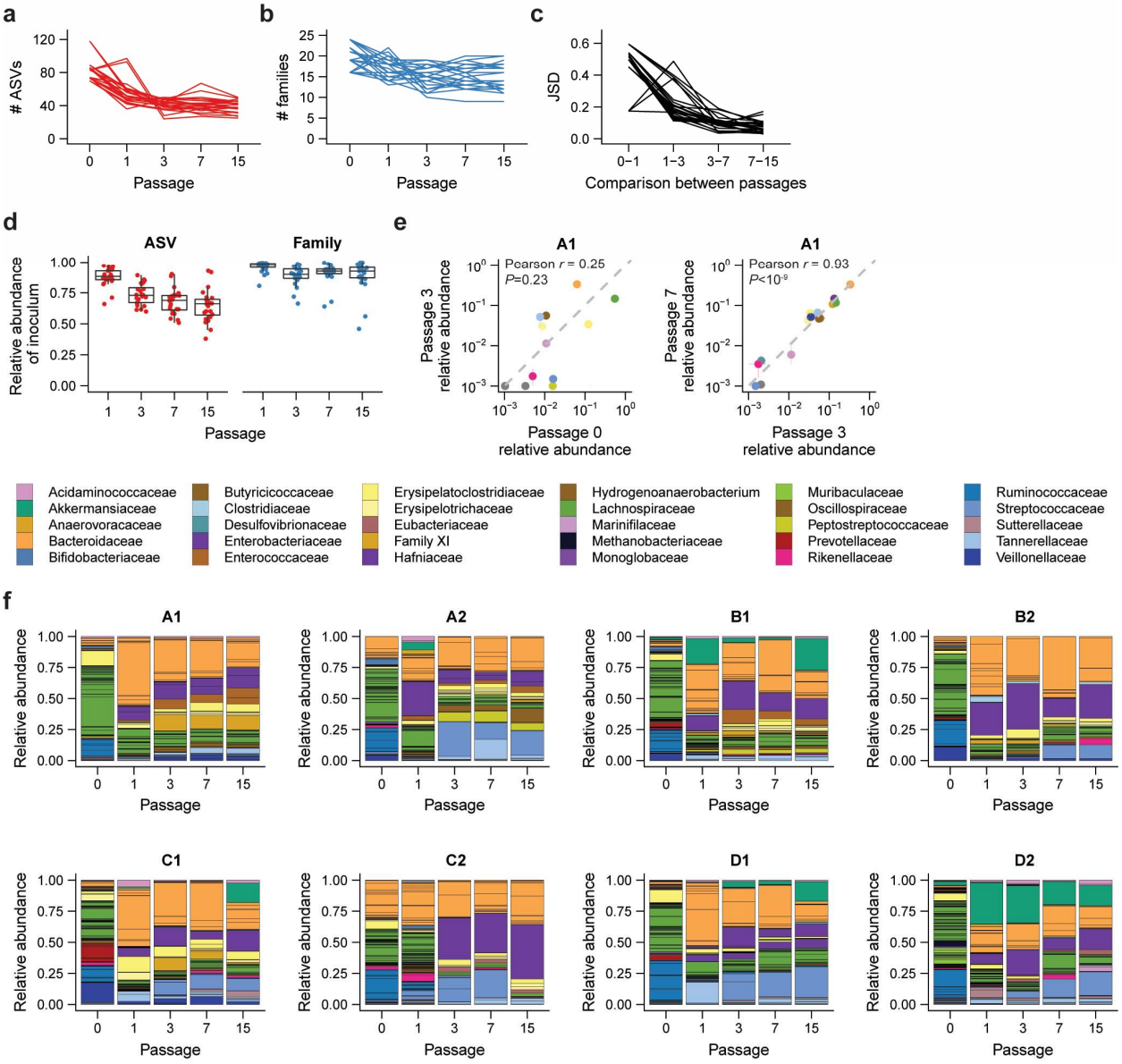
742 *L. lactis*, (d) the D1 community and *L. lactis*, and (e) *L. garvieae* and the D2

743 community. Lines are colored as in **Fig. 3a**. Dotted lines with open points show data

744 after three passages, and solid lines with closed points show passage five.

745 **Extended Data Figures**

746



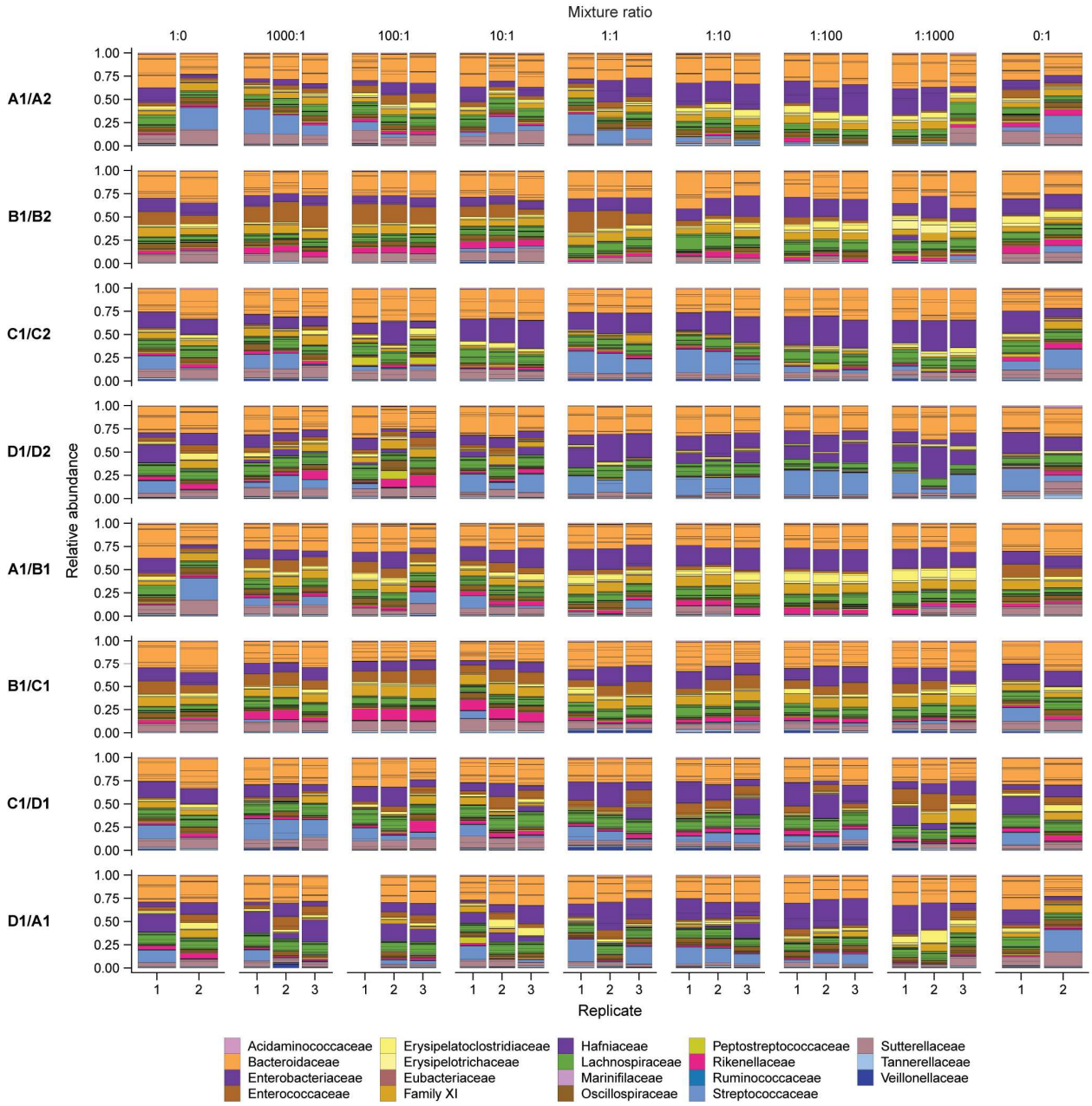
747

748 **Extended Data Figure 1: Diversity and composition of stool-derived communities.**

749 a) ASV richness of each community over passages. Only ASVs present in the

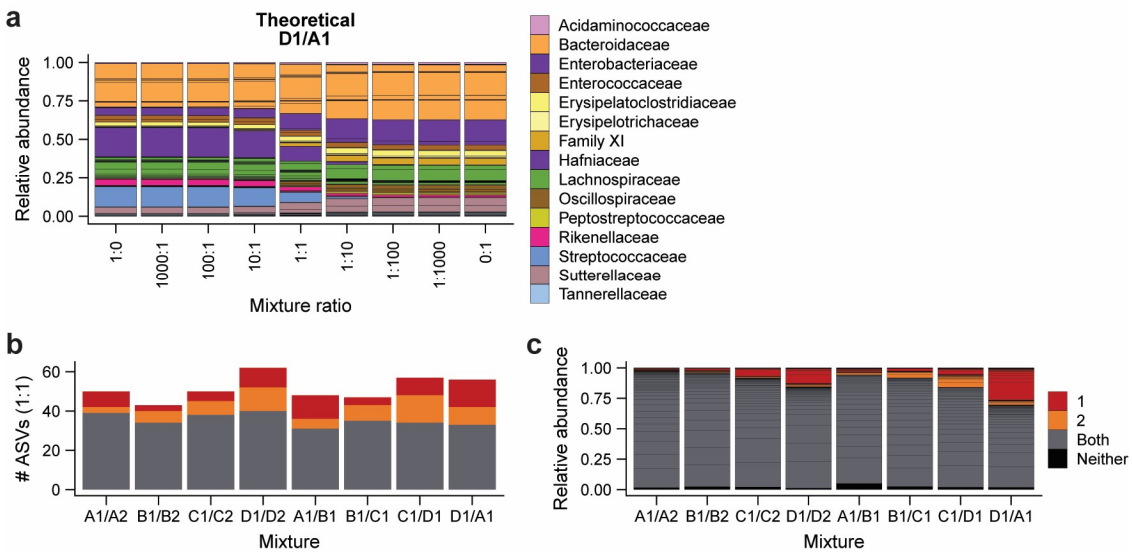
750 community at relative abundance >10^{-3} were included in the richness calculation.

751 b) Family-level richness of each community over passages. Families were counted as
752 present if an ASV in the family was present in the community at relative abundance
753 $>10^{-3}$.
754 c) Stability of each community over passages, as quantified by the Jensen-Shannon
755 divergence (JSD, **Methods**) between consecutive passages.
756 d) Proportion of the stool inoculum (passage 0) retained at each subsequent passage.
757 Each point shows the total ASV or family relative abundance of the stool inoculum
758 represented by taxa retained in communities at a given passage. Taxa were counted
759 as retained if they were detected at a relative abundance $>10^{-3}$.
760 e) Correlation of family-level relative abundance across passages 0, 3, and 7
761 (Pearson's tests, left: n=25, ρ=0.25, P=0.23; right: n=22, ρ=0.93, $P<10^{-9}$). Points are
762 colored by family. The dashed line indicates a perfect correlation in relative
763 abundance between passages.
764 f) Composition of one replicate of each community over passages. ASV relative
765 abundances are separated with solid lines and colored by family. Only the most
766 abundant families are shown in the key.



Extended Data Figure 2: Composition of co-culture communities.

ASV relative abundances are separated with solid lines and colored by bacterial family. Only the most abundant families are shown in the key. The first replicate of the 100:1 mixture ratio of the D1/A1 mixture had low sequencing depth and was removed from subsequent analyses.

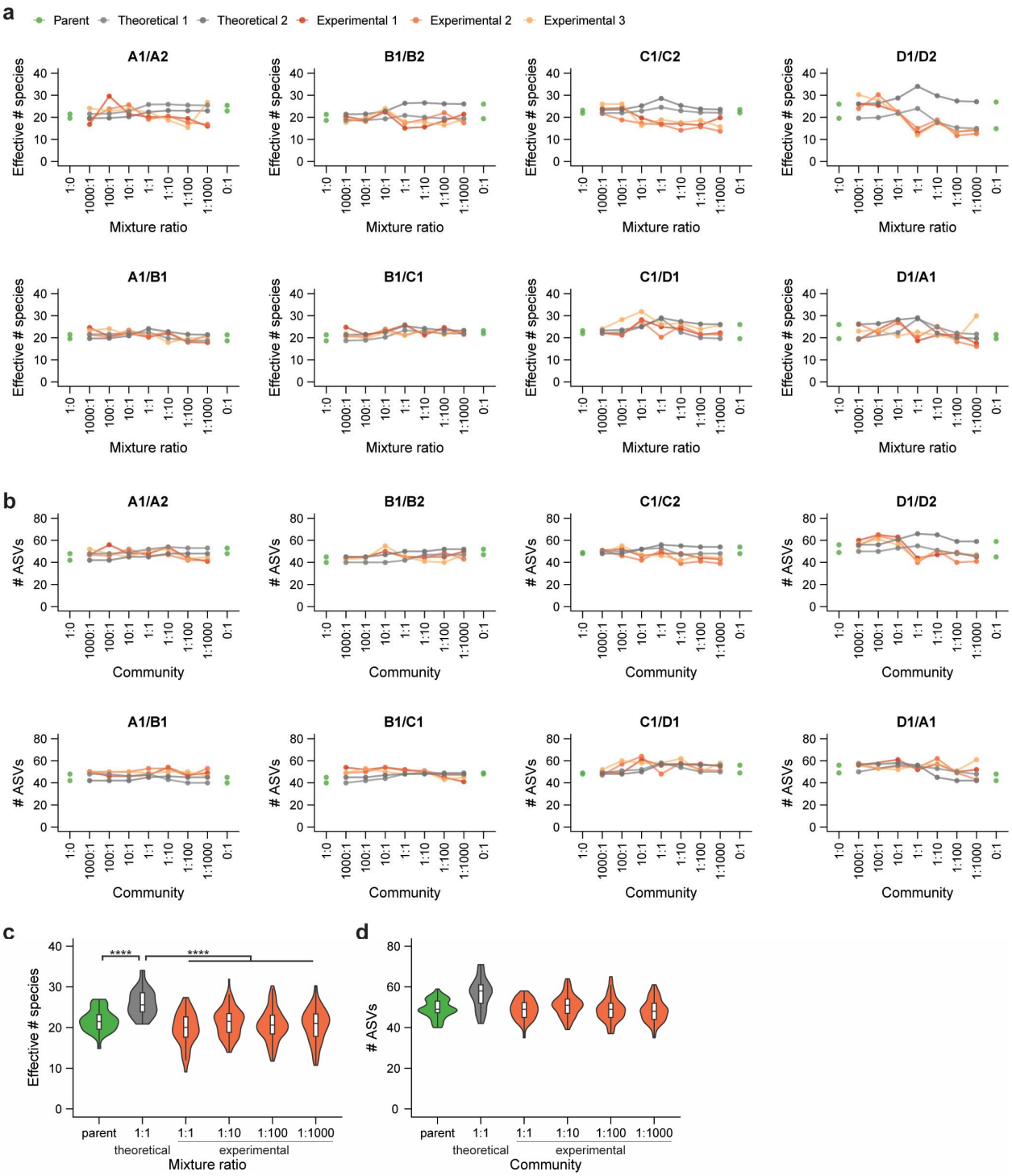


Extended Data Figure 3: Theoretical community composition is dose dependent, even though many ASVs are shared between parent communities.

a) The theoretical neutral community composition of an example mixture after five passages. The 1:0 and 0:1 ratios represent unmixed parent communities. ASV relative abundances are separated with solid lines and colored by family.

b) Most ASVs in each co-culture were present in both parent communities. ASVs were counted as originating from one or both parent communities if they were present at relative abundance $>10^3$.

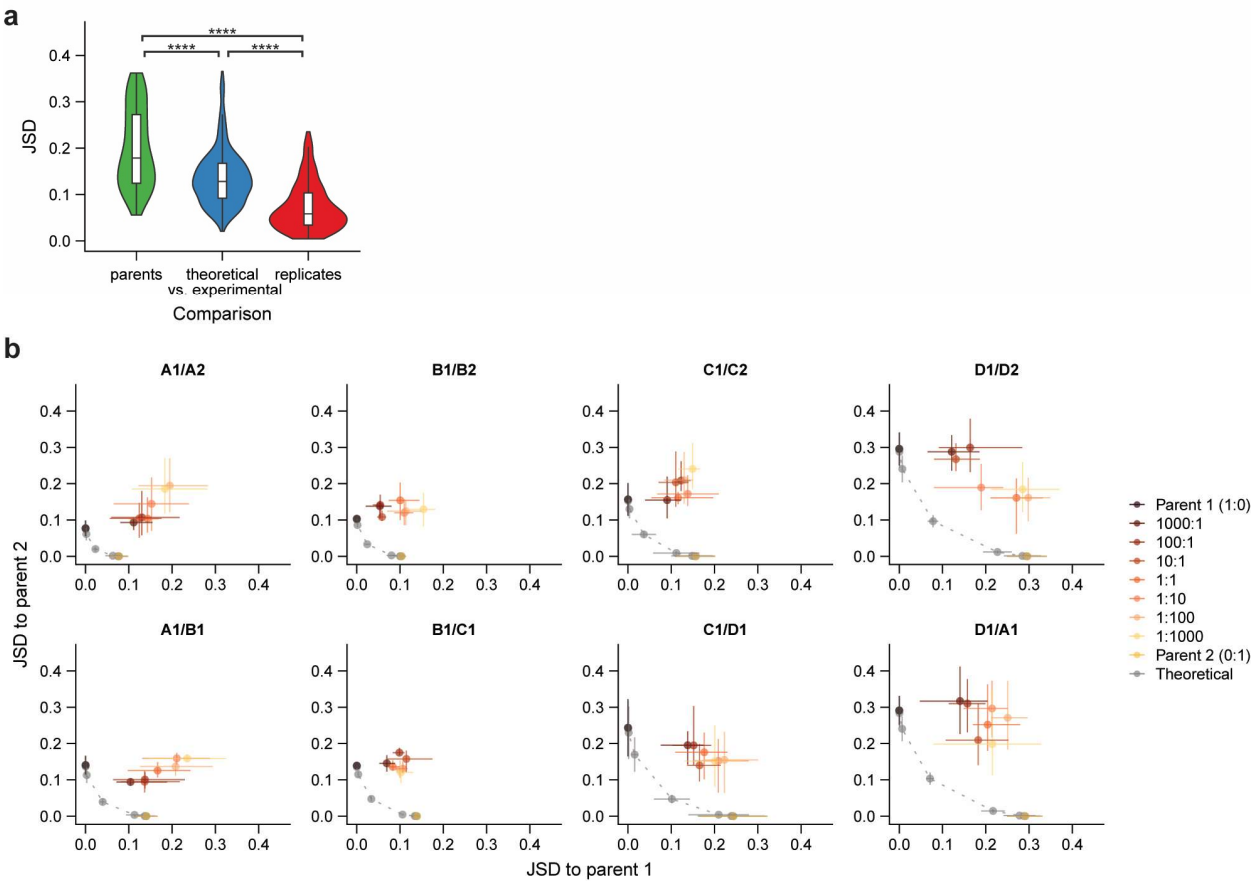
c) Relative abundances of ASVs at the 1:1 mixture ratio, colored by parent community of origin. Black lines indicate separate ASVs. Relative abundances were averaged across the three inoculation replicates.



Extended Data Figure 4: Diversity is limited in co-culture communities.

a,b) Diversity of each co-culture community, quantified by (**a**) the effective number of species ($e^{H'}$) calculated from the Shannon diversity index (H'), or (**b**) ASV richness,

789 assuming a detection limit of 10^{-3} relative abundance. Lines are colored as in **Fig.**
790 **1d**.
791 c) Distribution of diversity across parent and co-culture communities. The effective
792 number of species at each mixture ratio more closely resembled the diversity of the
793 parent communities than the higher-diversity theoretical prediction (Wilcoxon rank-
794 sum one-sided tests, parent (n=32) vs. theoretical (n=16) co-cultures: $P=6.1\times10^{-5}$;
795 theoretical vs. experimental 1:1 (n=24) co-cultures: $P=6.6\times10^{-7}$; theoretical vs. other
796 (n=48 at each mixture ratio) co-cultures: $P<10^{-4}$).
797 d) Distribution of ASV richness across parent and co-culture communities.

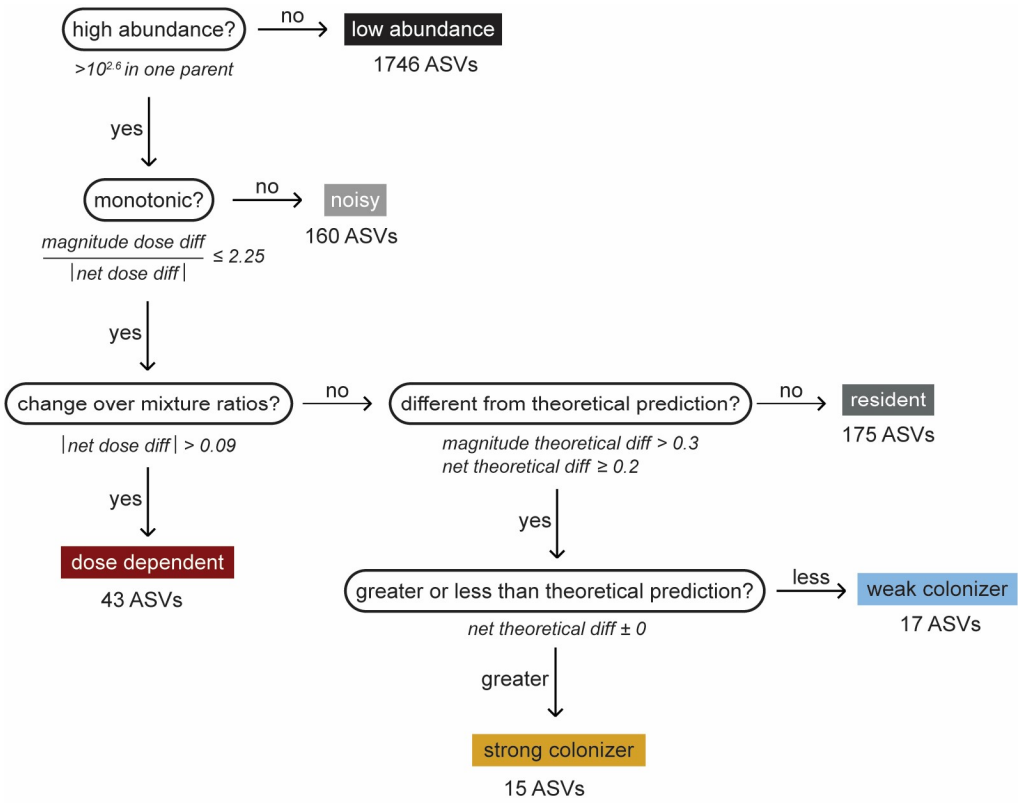


Extended Data Figure 5: Composition of co-culture communities varies by mixture ratio but differs from neutral predictions.

a) Comparisons of co-culture communities to theoretical communities. Violin plots represent the distribution of Jensen-Shannon divergence (JSD) between all pairs of parent communities (including those that were not mixed), theoretical and experimental mixtures at the 1:1 mixture ratio, and inoculation replicates of the 1:1 mixture ratio (significance calculated using Wilcoxon rank-sum one-sided tests, parents (n=56) vs. replicates (n=174): $P<10^{-15}$; parents vs. theoretical-experimental (n=334): $P=1.8\times10^{-7}$; theoretical-experimental vs. replicates: $P<10^{-15}$).

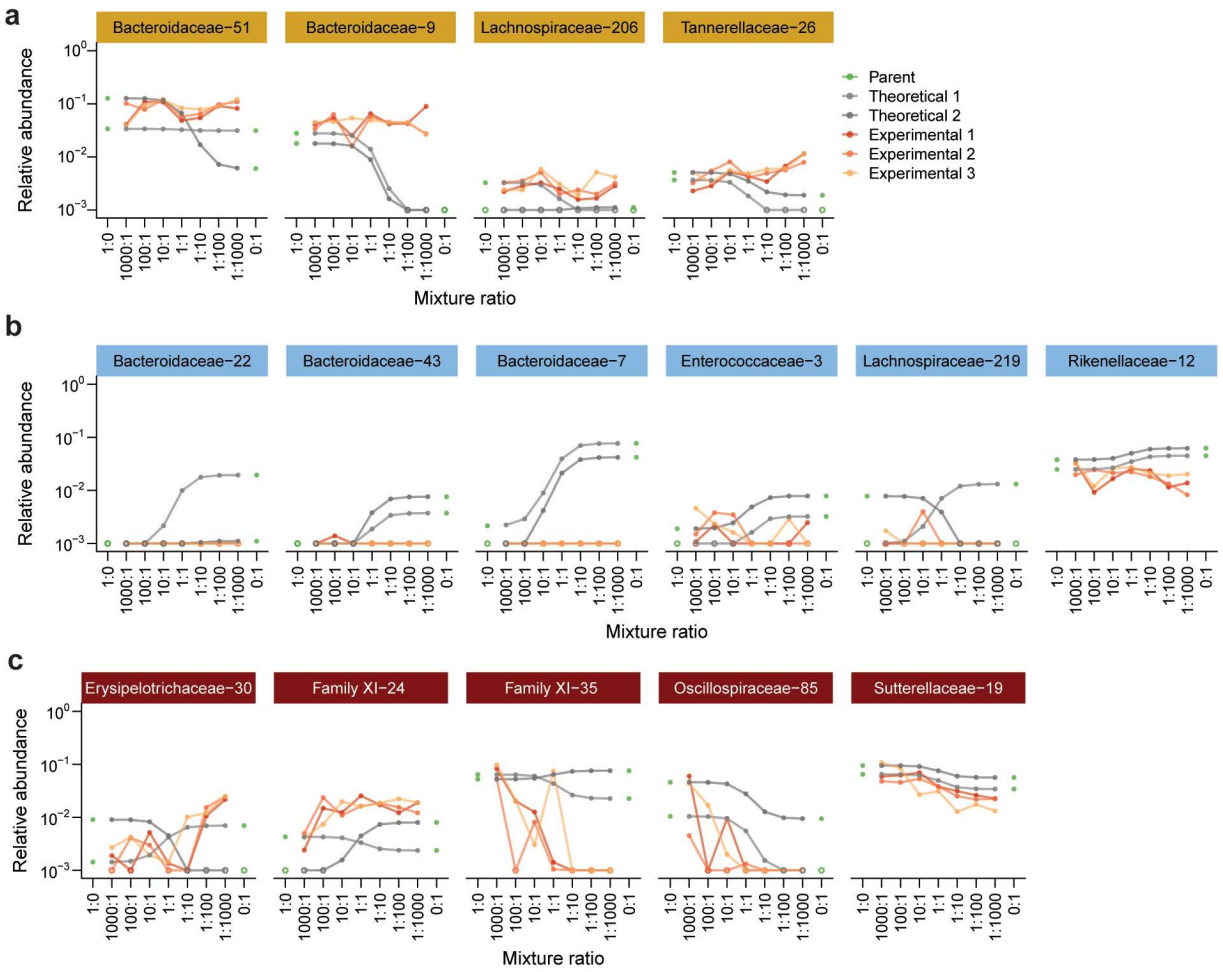
b) Variation of community composition across mixtures and mixture ratios. Each plot shows the Jensen-Shannon divergence (JSD) of all co-culture and theoretical

communities relative to both parent communities (brown and gold) for a single mixture. Points are colored as in **Fig. 1e**.



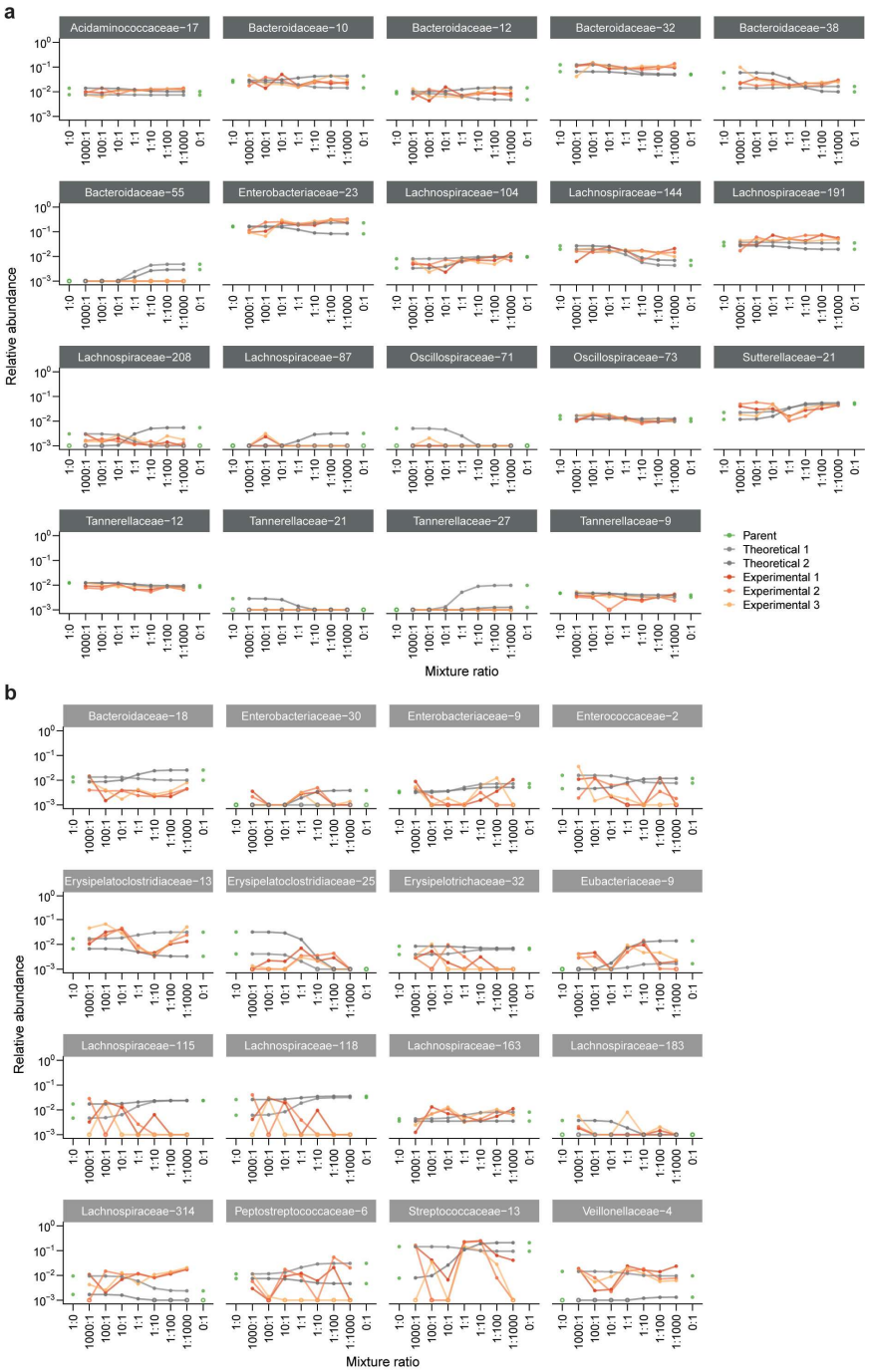
Extended Data Figure 6: Classification scheme for ASV colonization behavior.

The total number of ASVs across all mixtures is provided for each category of behavior.



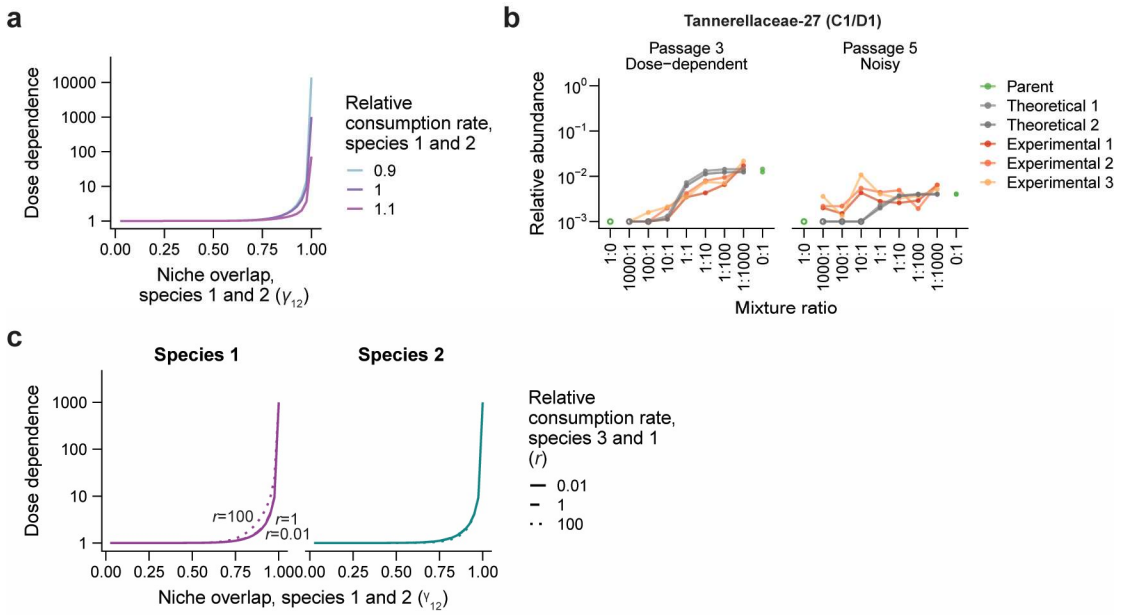
Extended Data Figure 7: Examples of strong, weak, and dose-dependent colonizers from the C1/C2 mixture.

a-c) Relative abundance of (**a**) strong, (**b**) weak, and (**c**) dose-dependent colonizers from the C1/C2 mixture. Each panel shows one ASV. Lines are colored as in **Fig. 2a**.



Extended Data Figure 8: Examples of resident and noisy ASVs from the C1/C2 mixture.

a,b) Relative abundance of (**a**) resident and (**b**) noisy ASVs from the C1/C2 mixture. Each panel shows one ASV. Lines are colored as in **Fig. 2a**.



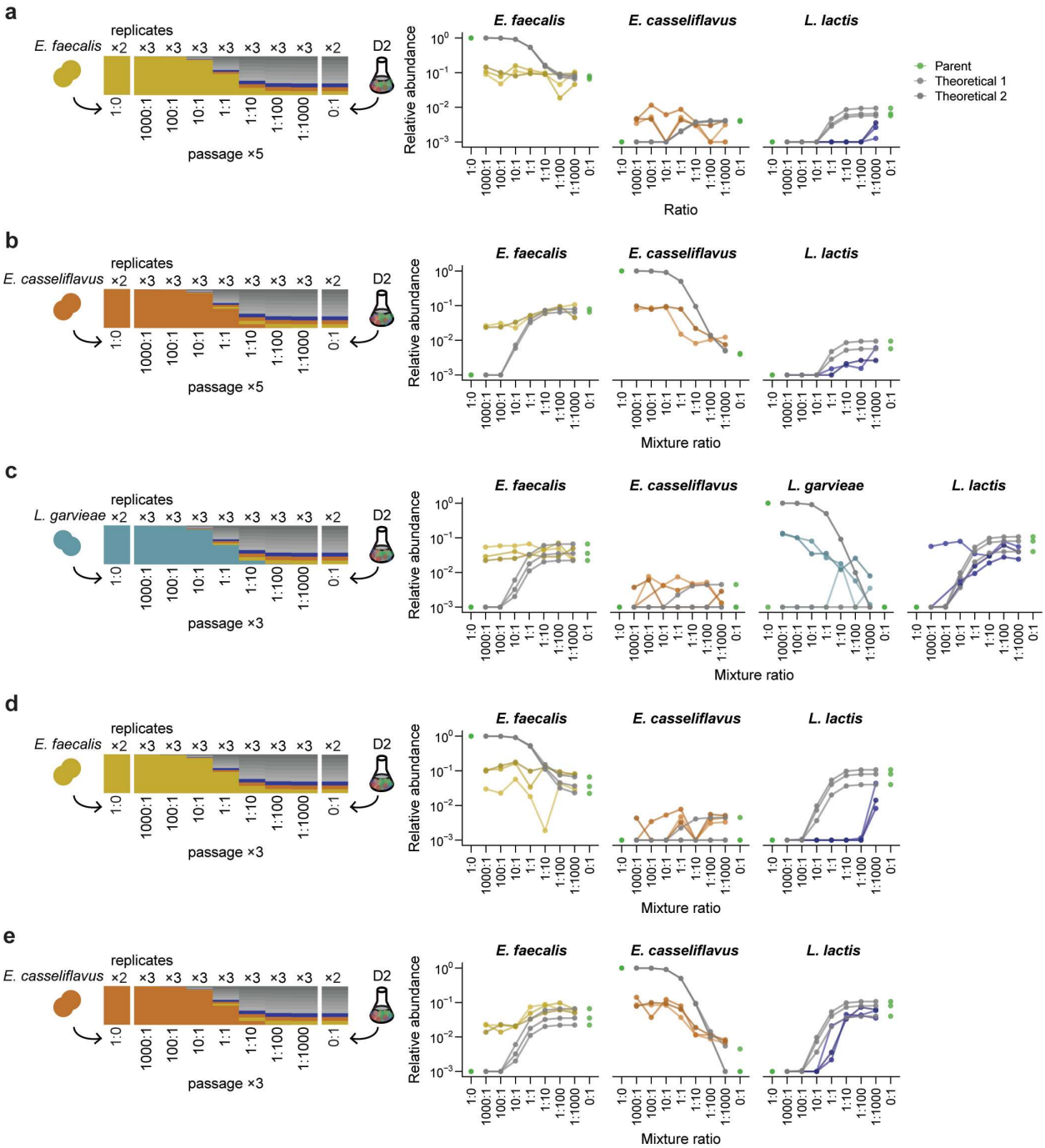
Extended Data Figure 9: Predicted effects of niche overlap and resource consumption rates in a consumer-resource model.

a) Dose dependence after five passages in a two-species consumer-resource model is robust to ~10% variation in relative consumption rates $R_{1,12}$ and $R_{2,12}$ of a shared resource μ_{12}.

b) The dose dependence of an ASV in the Tannerellaceae family declined substantially between passages three and five. Lines are colored as in **Fig. 2a**.

c) In our consumer-resource model, the dose dependence of species 1 after five passages increases with the relative resource consumption rate of species 3 compared with species 1 (r) at constant $\beta = 0.5$, while the dose dependence of species 2 is largely unaffected.

Dose dependence is defined as in **Fig. 3c**.



Extended Data Figure 10: Colonization behavior in strain-community mixtures.

a) Species relative abundances in the *E. faecalis*/D2 co-cultures after five passages. Only *L. lactis* was classified as dose dependent.

843 b) Species relative abundances in the *E. casseliflavus*/D2 co-cultures after five
844 passages. *E. faecalis*, *E. casseliflavus*, and *L. lactis* were classified as dose
845 dependent.
846 c) Species relative abundances in the *L. garvieae*/D2 co-cultures after three passages.
847 *L. garvieae* and *L. lactis* were classified as dose dependent.
848 d) Species relative abundances in the *E. faecalis*/D2 co-cultures after three passages.
849 Only *L. lactis* was classified as dose dependent.
850 e) Species relative abundances in the *E. casseliflavus*/D2 co-cultures after three
851 passages. *E. faecalis*, *E. casseliflavus*, and *L. lactis* were classified as dose
852 dependent.
853 Lines are colored as in **Fig. 2a**.

854 **Table S1. Sample origin of each *in vitro* community.**

855 *In vitro* communities were derived from day 0 pre-antibiotic samples collected from

856 healthy adults[75].

Community	Subject	Household
A1	XBA	XB
A2	XBB	XB
B1	XCA	XC
B2	XCB	XC
C1	XDA	XD
C2	XDB	XD
D1	XFA	XF
D2	XFB	XF

857

858 **References**

859 1. Elton, C. S. *The Ecology of Invasions by Animals and Plants*. (Springer New
860 York, 1958). doi:10.1007/978-1-4899-7214-9.

861 2. Lodge, D. M. Biological invasions: Lessons for ecology. *Trends Ecol Evol* **8**,
862 133–137 (1993).

863 3. Williamson, M. H. & Fitter, A. The characters of successful invaders. *Biol*
864 *Conserv* **78**, 163–170 (1996).

865 4. Mack, R. N. *et al.* Biotic Invasions: Causes, Epidemiology, Global
866 Consequences, and Control. *Ecological Applications* **10**, 689 (2000).

867 5. Simberloff, D. The role of propagule pressure in biological invasions. *Annu*
868 *Rev Ecol Evol Syst* **40**, 81–102 (2009).

869 6. Dassonville, N., Guillaumaud, N., Piola, F., Meerts, P. & Poly, F. Niche
870 construction by the invasive Asian knotweeds (species complex Fallopia):
871 Impact on activity, abundance and community structure of denitrifiers and
872 nitrifiers. *Biol Invasions* **13**, 1115–1133 (2011).

873 7. Callahan, B. J., Fukami, T. & Fisher, D. S. Rapid evolution of adaptive niche
874 construction in experimental microbial populations. *Evolution (N Y)* **68**,
875 3307–3316 (2014).

876 8. Finlay, B. J. & Clarke, K. J. Ubiquitous dispersal of microbial species.
877 *Scientific Correspondence* **400**, 828 (1999).

878 9. Barbour, K. M., Barrón-Sandoval, A., Walters, K. E. & Martiny, J. B. H.
879 Towards quantifying microbial dispersal in the environment. *Environ*
880 *Microbiol* **25**, 137–142 (2023).

881 10. Custer, G. F., Bresciani, L. & Dini-Andreote, F. Ecological and Evolutionary
882 Implications of Microbial Dispersal. *Frontiers in Microbiology* vol. 13 Preprint
883 at https://doi.org/10.3389/fmicb.2022.855859 (2022).

884 11. Shaffer, M. L. Minimum Population Sizes for Species Conservation.
885 *Bioscience* **31**, 131–134 (1981).

886 12. Lande, R. Genetics and Demography in Biological Conservation. *Science*
887 *(1979)* **241**, 1455–1460 (1985).

888 13. Nummi, P. Introduced Semiaquatic Birds and Mammals in Europe. in
889 *Invasive Aquatic Species of Europe. Distribution, Impacts, and*
890 *Management* (eds. Leppäkoski, E., Gollasch, S. & Olenin, S.) 162–172
891 (Springer Dordrecht, 2002).

14. Fukami, T. Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. *Annu Rev Ecol Evol Syst* **46**, 1–23 (2015).

15. Vellend, M. *The Theory of Ecological Communities*. (Princeton University Press, 2016).

16. Custer, G. F., Bresciani, L. & Dini-Andreote, F. Toward an integrative framework for microbial community coalescence. *Trends Microbiol* (2023) doi:10.1016/j.tim.2023.09.001.

17. Blackburn, T. M. *et al.* A proposed unified framework for biological invasions. *Trends Ecol Evol* **26**, 333–339 (2011).

18. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography*. (Princeton University Press, 2001).

19. Chesson, P. MacArthur's Consumer-Resource Model. *Theor Popul Biol* **37**, 26–38 (1990).

20. Lockwood, J. L., Cassey, P. & Blackburn, T. The role of propagule pressure in explaining species invasions. *Trends Ecol Evol* **20**, 223–228 (2005).

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

22. Ho, P.-Y., Nguyen, T. H., Sanchez, J. M., DeFelice, B. C. & Huang, K. C. Resource competition predicts assembly of in vitro gut bacterial communities. *bioRxiv* (2022) doi:10.1101/2022.05.30.494065.

23. Veltman, C. J., Nee, S. & Crawley, M. J. Correlates of Introduction Success in Exotic New Zealand Birds. *Am Nat* **147**, 542–557 (1996).

24. Drake, J. M., Baggenstos, P. & Lodge, D. M. Propagule pressure and persistence in experimental populations. *Biol Lett* **1**, 480–483 (2005).

25. Memmott, J., Craze, P. G., Harman, H. M., Syrett, P. & Fowler, S. V. The Effect of Propagule Size on the Invasion of an Alien Insect. *Source: Journal of Animal Ecology* **74**, 50–62 (2005).

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

27. Pyšek, P. *et al.* Naturalization of central European plants in North America: Species traits, habitats, propagule pressure, residence time. *Ecology* **96**, 762–774 (2015).

926 28. Chadwell, T. B. & Engelhardt, K. A. M. Effects of pre-existing submersed
927 vegetation and propagule pressure on the invasion success of Hydrilla
928 verticillata. *Journal of Applied Ecology* **45**, 515–523 (2008).

929 29. Evans, S., Martiny, J. B. H. & Allison, S. D. Effects of dispersal and
930 selection on stochastic assembly in microbial communities. *ISME Journal*
931 **11**, 176–185 (2017).

932 30. Albright, M. B. N., Sevanto, S., Gallegos-Graves, L. V. & Dunbar, J. Biotic
933 Interactions Are More Important than Propagule Pressure in Microbial
934 Community Invasions. *ASM Journals mBio* **11**, (2020).

935 31. Alzate, A., Onstein, R. E., Etienne, R. S. & Bonte, D. The role of
936 preadaptation, propagule pressure and competition in the colonization of
937 new habitats. *Oikos* **129**, 820–829 (2020).

938 32. Rojas-Botero, S., Kollmann, J. & Teixeira, L. H. Competitive trait hierarchies
939 of native communities and invasive propagule pressure consistently predict
940 invasion success during grassland establishment. *Biol Invasions* **24**, 107–
941 122 (2022).

942 33. Sierocinski, P., Soria Pascual, J., Padfield, D., Salter, M. & Buckling, A. The
943 impact of propagule pressure on whole community invasions in
944 biomethane-producing communities. *iScience* **24**, (2021).

945 34. Jones, M. L., Ramoneda, J., Rivett, D. W. & Bell, T. Biotic resistance
946 shapes the influence of propagule pressure on invasion success in bacterial
947 communities. *Ecology* **98**, 1743–1749 (2017).

948 35. Walter, J., Maldonado-Gómez, M. X. & Martínez, I. To engraft or not to
949 engraft: an ecological framework for gut microbiome modulation with live
950 microbes. *Curr Opin Biotechnol* **49**, 129–139 (2018).

951 36. Suez, J., Zmora, N., Segal, E. & Elinav, E. The pros, cons, and many
952 unknowns of probiotics. *Nature Medicine* vol. 25 716–729 Preprint at
953 https://doi.org/10.1038/s41591-019-0439-x (2019).

954 37. Gianotti, L. *et al.* A randomized double-blind trial on perioperative
955 administration of probiotics in colorectal cancer patients. *World J*
956 *Gastroenterol* **16**, 167–175 (2010).

957 38. Maldonado-Gómez, M. X. *et al.* Stable Engraftment of Bifidobacterium
958 longum AH1206 in the Human Gut Depends on Individualized Features of
959 the Resident Microbiome. *Cell Host Microbe* **20**, 515–526 (2016).

960 39. Ouwehand, A. C. A review of dose-responses of probiotics in human
961 studies. *Benef Microbes* **8**, 143–151 (2017).

40. Zmora, N. *et al.* Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features. *Cell* **174**, 1388-1405.e21 (2018).

41. Li, S. S. *et al.* Durable coexistence of donor and recipient strains after fecal microbiota transplantation. *Science (1979)* **352**, 586–589 (2016).

42. Schmidt, T. S. B. *et al.* Drivers and determinants of strain dynamics following fecal microbiota transplantation. *Nat Med* **28**, 1902–1912 (2022).

43. Jessup, C. M. *et al.* Big questions, small worlds: Microbial model systems in ecology. *Trends Ecol Evol* **19**, 189–197 (2004).

44. Aranda-Díaz, A. *et al.* Assembly of gut-derived bacterial communities follows 'early-bird' resource utilization dynamics. *bioRxiv* (2023) doi:10.1101/2023.01.13.523996.

45. Javdan, B. *et al.* Personalized Mapping of Drug Metabolism by the Human Gut Microbiome. *Cell* **181**, 1661-1679.e22 (2020).

46. Aranda-Díaz, A. *et al.* Establishment and characterization of stable, diverse, fecal-derived in vitro microbial communities that model the intestinal microbiota. *Cell Host Microbe* **30**, 260-272.e5 (2022).

47. Celis, A. I., Relman, D. A. & Huang, K. C. The impact of iron and heme availability on the healthy human gut microbiome in vivo and in vitro. *Cell Chem Biol* **30**, 110-126.e3 (2023).

48. Rillig, M. C. *et al.* Interchange of entire communities: Microbial community coalescence. *Trends Ecol Evol* **30**, 470–476 (2015).

49. Posfai, A., Taillefumier, T. & Wingreen, N. S. Metabolic Trade-Offs Promote Diversity in a Model Ecosystem. *Phys Rev Lett* **118**, (2017).

50. Abreu, C. I., Andersen Woltz, V. L., Friedman, J. & Gore, J. Microbial communities display alternative stable states in a fluctuating environment. *PLoS Comput Biol* **16**, (2020).

51. Pereira, F. C. & Berry, D. Microbial nutrient niches in the gut. *Environ Microbiol* **19**, 1366–1378 (2017).

52. Sprockett, D., Fukami, T. & Relman, D. A. Role of priority effects in the early-life assembly of the gut microbiota. *Nat Rev Gastroenterol Hepatol* **15**, 197–205 (2018).

53. Wootton, J. T. Field parameterization and experimental test of the neutral theory of biodiversity. *Nature* **433**, 309–312 (2005).

54. D'Andrea, R., Gibbs, T. & O'Dwyer, J. P. Emergent neutrality in consumer-resource dynamics. *PLoS Comput Biol* **16**, (2020).

55. Grilli, J., Barabás, G., Michalska-Smith, M. J. & Allesina, S. Higher-order interactions stabilize dynamics in competitive network models. *Nature* **548**, 210–213 (2017).

56. Morin, M. A., Morrison, A. J., Harms, M. J. & Dutton, R. J. Higher-order interactions shape microbial interactions as microbial community complexity increases. *Sci Rep* **12**, (2022).

57. Ludington, W. B. Higher-order microbiome interactions and how to find them. *Trends Microbiol* **30**, 618–621 (2022).

58. Chang, C.-Y., Bajić, D., Vila, J. C. C., Estrela, S. & Sanchez, A. Emergent coexistence in multispecies microbial communities. *Science (1979)* **381**, (2023).

59. Fukami, T. Assembly history interacts with ecosystem size to influence species diversity. *Ecology* **85**, 3234–3242 (2004).

60. Fukami, T. & Nakajima, M. Community assembly: Alternative stable states or alternative transient states? *Ecol Lett* **14**, 973–984 (2011).

61. Song, C., Fukami, T. & Saavedra, S. Untangling the complexity of priority effects in multispecies communities. *Ecol Lett* **24**, 2301–2313 (2021).

62. Mallon, C. A. *et al.* The impact of failure: Unsuccessful bacterial invasions steer the soil microbial community away from the invader's niche. *ISME Journal* **12**, 728–741 (2018).

63. Warren, P. H., Law, R. & Weatherby, A. J. Mapping the Assembly of Protist Communities in Microcosms. *Ecology* **84**, 1001–1011 (2003).

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

65. Chappell, C. R. *et al.* Wide-ranging consequences of priority effects governed by an overarching factor. *Elife* **11**, (2022).

66. Simberloff, D. & Von Holle, B. *Positive interactions of nonindigenous species: invasional meltdown? Biological Invasions* vol. 1 (1999).

67. Miller, T. E., TerHorst, C. P. & Burns, J. H. The ghost of competition present. *American Naturalist* **173**, 347–353 (2009).

68. Olito, C. & Fukami, T. Long-term effects of predator arrival timing on prey community succession. *American Naturalist* **173**, 354–362 (2009).

69. O'Callaghan, M., Ballard, R. A. & Wright, D. Soil microbial inoculants for sustainable agriculture: Limitations and opportunities. *Soil Use Manag* **38**, 1340–1369 (2022).

1034 70. Pudlo, N. A. *et al.* Phenotypic and Genomic Diversification in Complex
1035 Carbohydrate-Degrading Human Gut Bacteria. *mSystems* **7**, (2022).

1036 71. Han, S. *et al.* A metabolomics pipeline for the mechanistic interrogation of
1037 the gut microbiome. *Nature* **595**, 415–420 (2021).

1038 72. Abrams, P. The Theory of Limiting Similarity. *Source: Annual Review of*
1039 *Ecology and Systematics* **14**, 359–376 (1983).

1040 73. Connors, B. M. *et al.* Model-guided design of the diversity of a synthetic
1041 human gut community. *bioRxiv* (2022) doi:10.1101/2022.03.14.484355.

1042 74. Martiny, J. B. H., Jones, S. E., Lennon, J. T. & Martiny, A. C. Microbiomes
1043 in light of traits: A phylogenetic perspective. *Science* vol. 350 Preprint at
1044 https://doi.org/10.1126/science.aac9323 (2015).

1045 75. Xue, K. S. *et al.* Prolonged delays in human microbiota transmission after a
1046 controlled antibiotic perturbation. *bioRxiv* (2023)
1047 doi:10.1101/2023.09.26.559480.

1048 76. Apprill, A., Mcnally, S., Parsons, R. & Weber, L. Minor revision to V4 region
1049 SSU rRNA 806R gene primer greatly increases detection of SAR11
1050 bacterioplankton. *Aquatic Microbial Ecology* **75**, 129–137 (2015).

1051 77. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters:
1052 Assessing small subunit rRNA primers for marine microbiomes with mock
1053 communities, time series and global field samples. *Environ Microbiol* **18**,
1054 1403–1414 (2016).

1055 78. Smith, T., Heger, A. & Sudbery, I. UMI-tools: Modeling sequencing errors in
1056 Unique Molecular Identifiers to improve quantification accuracy. *Genome*
1057 *Res* **27**, 491–499 (2017).

1058 79. Martin, M. Cutadapt removes adapter sequences from high-throughput
1059 sequencing reads. *EMBnet J* **17**, 10–12 (2011).

1060 80. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from
1061 Illumina amplicon data. *Nat Methods* **13**, 581–583 (2016).

1062 81. Quast, C. *et al.* The SILVA ribosomal RNA gene database project:
1063 Improved data processing and web-based tools. *Nucleic Acids Res* **41**,
1064 (2013).

1065 82. Shiver, A. L., Culver, R., Deutschbauer, A. M. & Huang, K. C. Rapid
1066 ordering of barcoded transposon insertion libraries of anaerobic bacteria.
1067 *Nat Protoc* **16**, 3049–3071 (2021).

1068 83. Cermak, N., Datta, M. Sen & Conwill, A. Rapid, Inexpensive Measurement
1069 of Synthetic Bacterial Community Composition by Sanger Sequencing of
1070 Amplicon Mixtures. *iScience* **23**, (2020).

1071 84. Blazanin, M. gcplyr: an R package for microbial growth curve data analysis.
1072 *bioRxiv* (2023) doi:10.1101/2023.04.30.538883.

1073