Competition for shared resources increases the impact of propagule 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,‡], 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 outcomes of species introductions depend on both propagule size and resource competition, but the effects of these factors are inconsistent across community contexts. Here, we investigate how propagule size and resource competition interact to shape colonization during community coalescence by mixing pairs of *in vitro* gut microbial communities at ratios that vary over seven orders of magnitude. Each resulting co-culture contained species whose relative abundance depended on inoculation dose, causing the composition of some co-cultures to vary substantially across mixture ratios, even after >30 generations of growth. Using a consumer-resource model, we show that dose-dependent colonization can arise when niche overlap is high and species compete near-neutrally for shared resources. This model successfully predicts the outcomes of mixtures of strain isolates, in which propagule size has larger, longer-lasting effects in more diverse communities. Our study establishes a framework for predicting the context-dependent impacts of propagule size, providing principles for the design of microbiome therapeutics.

Introduction

Colonization of an ecosystem by new species plays a major role in shaping community composition and function, but the outcome of species introductions is notoriously difficult to predict[1–7]. To colonize, a species must disperse from its original population and overcome several ecological barriers, including the stochastic effects of ecological drift at small population sizes[8,9] and competition with resident species for a niche in the community[10–14]. Increases in the size and frequency at which propagules of dispersing species are introduced can help to overcome these barriers[15–19]. However, the effects of propagule size and frequency, known together as propagule pressure, are inconsistent across different contexts[5,20–22]. Large propagules can fail to colonize, and small propagules can successfully colonize, when other ecological features such as competition affect species introductions more strongly than propagule pressure[23–28]. Disentangling the influence of these factors is crucial for predicting the outcomes of species introductions.

Understanding the effect of propagule pressure on colonization is especially important in microbial communities, from which species can disperse in propagules varying in size across many orders of magnitude[29–31]. Dispersal between microbial communities can involve dozens or even hundreds of species[32,33], resulting in community coalescence events such as river confluences, leaves falling onto soil, and fecal-oral transmission of gut microbiota[13,30,34–37]. Determining the impacts of propagule size is crucial to predicting the effects of these large-scale introductions.

Understanding the ecological determinants of colonization in the human gut is especially important because of their relevance to the development of successful microbiome therapeutics[38,39]. Studies assessing colonization following fecal microbiota transplants (FMTs) or other probiotic introductions into the gut have shown that both inoculation dose and competition can independently affect colonization outcomes[32,40–44]. Systematic quantification of the relative importance of propagule size and competition is more tractable using laboratory model microbial communities, which can be stable, diverse, and reproducible[45–48]. Stool-derived *in vitro* communities can recapitulate microbiota composition and *in vivo* responses to antibiotic perturbation[49] or iron supplementation[50], making these communities a powerful model system for investigating microbial ecology.

Here, we investigate the effects of propagule size on colonization using controlled mixtures of human stool-derived *in vitro* communities at ratios varying across seven orders of magnitude. While some species reached the same relative abundance regardless of mixture ratio, the relative abundance of other species varied with inoculation dose, even after >30 generations of growth. We utilize a consumer-resource model to show that dose dependence can arise and persist when species have high niche overlap and compete near-neutrally for shared resources, an effect that can be enhanced by the presence of additional species. Using mixtures of strain isolates, we experimentally confirmed the model prediction that the magnitude and duration of dose dependence increases in more diverse communities. This work shows how the effects

of propagule size can be amplified by resource competition in complex communities and have lasting effects on the outcomes of community coalescence.

Results

Diversity and composition of co-culture communities vary by inoculation dose 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[49], 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). 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 had made up >50% of the relative abundance of the stool inoculum (**Extended Data Fig. 1c**). Similarly, family-level relative abundances quickly stabilized (**Fig. 1c**, Pearson's tests, left: n=25, ρ=0.25, P=0.23; right: n=22, ρ=0.93, $P<10^{-9}$, **Extended Data Fig. 1c,d**), with the remaining families after 15 passages collectively representing >75% of the relative abundance of the corresponding stool inoculum in most cases (**Extended Data Fig. 1c**). Thus, these stool-derived *in vitro* communities are diverse and stable models of stool inocula that can be used to study ecological interactions among gut microbes.

110 To systematically assess the effects of propagule size on colonization, we mixed eight
111 pairs of *in vitro* “parent” communities at seven mixture ratios ranging from 1000:1 to
112 1:1000 (**Fig. 1d, Extended Data Fig. 2, Table S1**). We co-cultured mixtures in triplicate
113 for five passages to allow composition to re-equilibrate, and we passaged two replicates
114 of each parent community alongside to allow for comparison. After passaging, we used
115 16S sequencing to assess the effects of inoculation dose on community composition
116 and the colonization outcomes of individual ASVs.
117
118 Ecological neutral theory predicts that if all species are competitively neutral, then
119 propagule size is the only determinant of colonization outcomes[12,28,51]. This theory
120 provides a useful null model for measuring the impact of propagule size on community
121 composition and the colonization of individual species during community coalescence.
122 We compared this model to an alternative hypothesis that non-neutral resource
123 competition determines colonization success, meaning that co-culture compositions
124 should be largely independent of inoculation dose[5,12,52]. To assess the similarity of our
125 data to the neutral model, we computationally generated a set of theoretical mixture
126 communities with compositions based on the inoculation dose of each species
127 (**Extended Data Fig. 3a**), which also represented the initial state of each experimental
128 co-culture before passaging. A large proportion of ASVs in each mixture were shared
129 between parent communities at similar levels (**Extended Data Fig. 3b,c**), and were
130 therefore predicted to maintain their abundance across mixture ratios; all remaining
131 ASVs were predicted to vary in relative abundance with inoculation dose (**Extended
132 Data Fig. 3a**).

The neutral model predicts that diversity (defined here as the effective number of species inferred from the Shannon diversity index) should be maximal at the 1:1 mixture ratio. Even though many ASVs were shared between parent communities, the predicted maximum diversity of theoretical co-cultures at the 1:1 mixture ratio was higher than the diversity of either parent community, accounting for the detection limit of 16S sequencing (**Fig. 1e,f**, Wilcoxon rank-sum one-sided test, $P=6.1\times10^{-5}$, **Extended Data Fig. 4, Methods**). Contrary to the neutral prediction, we found that co-cultures consistently maintained diversities similar to or lower than one of the parent communities, despite some variation over mixture ratios (**Fig. 1e,f, Extended Data Fig. 4**). This finding suggests that the diversity of the parent communities is already nearly maximized such that most niches are occupied by resident species, and that non-neutral competition between residents and new species makes it difficult for all species to coexist during coalescence.

Although propagule size was not the sole determinant of community diversity, diversity sometimes varied over mixture ratios, suggesting that it was impacted by inoculation dose. To further investigate the effects of propagule size on community composition, we compared each experimental co-culture to its corresponding theoretical co-culture using the Jensen-Shannon divergence (JSD, **Methods**). Experimental and theoretical co-cultures were less similar than replicate communities were to each other (**Fig. 1g**, Wilcoxon rank-sum two-sided test, $P<10^{-15}$), further supporting the importance of non-neutral competition during colonization. To assess the relative importance of propagule

size and non-neutral competition across mixtures, we used JSD to quantify the difference between each co-culture and its corresponding parent communities.

Co-cultures consistently differed from the corresponding theoretical expectation at every mixture ratio (**Fig. 1h**). In some mixtures, co-culture composition was similar across mixture ratios (B1/C1, **Fig. 1h**), indicating that propagule size had a weak effect. However, in other mixtures, composition varied substantially with inoculation dose, despite deviating from neutral predictions (D1/D2, **Fig. 1h**). These findings highlight the varied importance of propagule size and competition across mixtures.

Dose-dependent colonizers are present in every community mixture

To investigate why co-culture composition exhibits a stronger dependence on propagule size in some mixtures than in others, we examined colonization outcomes at the ASV level. We compared the relative abundance of each ASV after five passages to the neutral prediction determined solely by its inoculation dose in the initial mixture (**Extended Data Fig. 5**). Of the ASVs that had distinct relative abundances in the two parent communities, some maintained a consistent relative abundance in co-cultures across mixture ratios, suggesting a minimal effect of propagule size on their colonization (**Fig. 2a, Extended Data Fig. 6a,b**). We classified these dose-independent ASVs as strong or weak colonizers depending on whether their relative abundance was greater or less than the neutral prediction, respectively.

178 In contrast, many of the remaining ASVs exhibited dose-dependent colonization, with
179 relative abundances that varied systematically across mixture ratios (**Fig. 2b,c,**
180 **Extended Data Fig. 6c**). Dose-dependent colonizers were present in every community
181 mixture, often in higher numbers than strong or weak colonizers, demonstrating that
182 propagule size can influence the colonization of a sizeable fraction of species in each
183 community during coalescence. Intriguingly, most strong, weak, and dose-dependent
184 colonizers exhibited distinct colonization patterns in different co-cultures and between
185 closely related taxa (**Fig. 2d, Methods**), suggesting that these behaviors were
186 influenced by community context rather than species identity or phylogeny.
187
188 Dose-dependent colonizers accounted for ~10-20% of the relative abundance of each
189 co-culture (**Fig. 2e**) except the one whose composition showed the least dose
190 dependence (B1/C1, **Fig. 1h**). A large fraction of the total relative abundance of each
191 co-culture was accounted for by resident ASVs that were at similar relative abundances
192 in the two parent communities and remained consistently at this relative abundance
193 across mixture ratios (**Fig. 2f, Extended Data Fig. 7a**). Another large fraction of each
194 community consisted of noisy ASVs whose relative abundance varied inconsistently
195 across inoculation doses (**Fig. 2g, Extended Data Fig. 7b**). Notably, many noisy ASVs
196 exhibited large changes in relative abundance across mixture ratios that were
197 consistent among the three replicates of each mixture (**Extended Data Fig. 7b**),
198 suggesting that changes in the surrounding community may have driven these
199 fluctuations. Overall, our results show that propagule size influences the colonization of

200 a distinct set of species in every co-culture, even though the colonization of other
201 species is primarily determined by non-neutral competition.
202
203 **Transient dose dependence arises from high niche overlap in a consumer-**
204 **resource model**
205 We next sought to identify the ecological forces that give rise to dose dependence
206 during community mixtures. In the D1/D2 co-cultures, we observed three dose-
207 dependent *Lactococcus* and *Enterococcus* colonizers from the D1 parent community.
208 The relative abundances of these ASVs were roughly anticorrelated with those of
209 another dose-dependent *Lactococcus* from the D2 community (**Fig. 3a**). This
210 observation led us to hypothesize that dose dependence could arise from competition
211 between phylogenetically related taxa that occupy overlapping niches.
212
213 To understand whether competition for shared resources could explain this dose-
214 dependent colonization, we used a standard consumer-resource (CR) model to
215 investigate the conditions under which dose dependence could arise[52] (**Methods**). We
216 first studied a system in which two species compete for a set of three common
217 resources: μ_1, μ_2, and μ_{12} (**Fig. 3b**). Two of these resources, μ_1 and μ_2, are consumed
218 exclusively by species 1 and 2, respectively, and the third resource, μ_{12}, is consumed
219 by both species at equal rates ($R_{1,12} = R_{2,12}$). We investigated how the niche overlap of
220 species 1 with species 2, defined as $\gamma_{12} = \frac{\mu_{12}}{\mu_1+\mu_{12}}$, would affect the colonization behavior
221 of these two species. We varied γ_{12} from 0.025 to 1 while maintaining the same amount
222 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 CR model to test how species abundances would change over five passages.

In this simple 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, the two species rapidly reached consistent relative abundances across mixture ratios ($\gamma_{12} = 0.75$, **Figure 3c**). By contrast, when niche overlap was high, the relative abundance of both species showed dose dependence, varying by ~20-fold between the 1000:1 and 1:1000 mixture ratios after three passages ($\gamma_{12} = 0.975$, **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 was varied by ~10% (**Extended Data Fig. 8a**). These results show that dose dependence can arise when the growth of two species is dominated by their consumption of shared resources.

Although these species showed dose dependence when niche overlap was high, the magnitude of this dose dependence declined over time (**Fig. 3c, Extended Data Fig. 8b**). For example, when $\gamma_{12} = 0.975$, the dose dependence of the two species declined from ~20-fold to ~10-fold from passage three to passage five (**Fig. 3c**). This declining dose dependence was consistent with decreases in dose dependence over passages that we observed among some ASVs during community coalescence (**Extended Data**

245 **Fig. 8c**). This finding demonstrates that high niche overlap can slow a community's
246 convergence to equilibrium, creating transient, neutral-like colonization dynamics.
247
248 **Strain isolates show less dose dependence in pairwise mixtures than during**
249 **community coalescence**
250 Based on the predictions of our CR model, we hypothesized that the *Lactococcus* and
251 *Enterococcus* species that showed dose-dependent colonization in the D1/D2 co-
252 cultures had high levels of niche overlap that would lead them to show dose
253 dependence in pairwise mixtures as well. To test this prediction, we isolated
254 representatives of these ASVs (**Methods**). We performed Sanger sequencing of the
255 16S region and whole-genome sequencing (**Methods**) to identify the strains from the
256 D1 community as *Lactococcus garvieae*, *Enterococcus faecalis*, and *Enterococcus*
257 *casseliflavus*, and the strain from the D2 community as *Lactococcus lactis* (**Fig. 4a**).
258 These strains exhibited similar maximum growth rates in monoculture, although small
259 differences in growth dynamics suggested the potential for distinct profiles of resource
260 utilization (**Fig. 4a**).
261
262 We performed a set of pairwise mixtures of *L. lactis* with the strains isolated from
263 community D1 at ratios from 1000:1 to 1:1000 (**Fig. 4b-d**). Most strains exhibited some
264 degree of dose-dependent colonization after three passages, suggesting that they had
265 some niche overlap (**Fig. 4b-d**). However, after five passages, each strain reached a
266 relative abundance that was consistent across mixture ratios (**Fig. 4b-d**). This lack of
267 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 shows that pairwise strain mixtures are unlikely to fully capture the magnitude and duration of dose dependence that we observe in more complex communities.

Dose dependence is enhanced in mixtures with more diverse communities

We used our CR model to investigate how dose dependence would be affected by the addition of a third species. We introduced a third species into the model that consumes a unique resource that was initialized each passage at the same level as the other resources, such that $\mu_3 = \mu_2$ (**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. To investigate how this new species affects colonization, we varied β from 0 to 0.9 (**Fig. 5b**). As before, we initialized the abundances of species 1 and 2 at a range of ratios from 1000:1 to 1:1000, and we set the abundance of species 3 equal to that of species 2. We simulated the CR 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. 8d**), 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 resource (**Fig. 5b, Extended Data Fig. 8d**). Under these

291 conditions, the dose dependence of species 1 increased substantially due to its
292 increased reliance on the resource shared with species 2 (**Fig. 5b**, **Extended Data Fig.**
293 **8d**). At the same time, the dose dependence of species 2 was minimally affected by the
294 amount of niche overlap between species 1 and 3, showing that the addition of new
295 species does not necessarily alter the colonization behavior of all community members
296 (**Fig. 5b, Extended Data Fig. 8d**). Together, these results predict that initial mixture
297 ratios should have larger, more lasting effects on community composition in more
298 diverse communities, in which competition among multiple species limits the unique
299 resources available for any species.
300
301 To test these predictions, we introduced the *L. lactis* strain into an even mixture of *E.*
302 *faecalis* and *E. casseliflavus* at ratios from 1000:1 to 1:1000. Consistent with our
303 prediction, after three passages, the degree of dose-dependence exhibited by *L. lactis*
304 increased to ~200-fold (**Fig. 5c**) from ~60-fold in pairwise co-cultures (**Fig. 4b,c**), while
305 neither *E. faecalis* nor *E. casselifavus* were classified as dose-dependent colonizers
306 (**Fig. 5c**). However, after five passages, none of the three strains exhibited dose-
307 dependent colonization (**Fig. 5c**). These results support our model's prediction that the
308 addition of a new species can increase the magnitude of dose dependence, even if this
309 dose dependence still declines more rapidly than during community coalescence.
310
311 To test the prediction that introducing strains into a more diverse community would
312 increase the magnitude and duration of dose dependence by increasing niche overlap,
313 we performed a set of strain-community mixtures in which we mixed each of the two

314 *Enterococcus* and *Lactococcus* strains into the D1 or D2 communities at ratios from
315 1000:1 to 1:1000 (**Fig. 5d,e, Extended Data Fig. 9**). In striking contrast to the pairwise
316 and three-strain mixtures, a subset of strains in each strain-community mixture showed
317 dose dependence after five passages (**Fig. 5d,e, Extended Data Fig. 9**). For instance,
318 *E. casseliflavus* and *L. lactis* were no longer dose-dependent by the fifth passage in
319 pairwise co-cultures (**Fig. 4c**). However, when *L. lactis* was introduced into the D1
320 community, both strains exhibited dose dependence after five passages, varying by
321 ~10-fold in relative abundance across all mixture ratios (**Fig. 5d**). Similarly, when *L.*
322 *garvieae* was introduced into the D2 community, it showed complementary patterns of
323 dose dependence with *L. lactis*, even though *E. casseliflavus* remained at consistent
324 relative abundances across mixture ratios (**Fig. 5e, Extended Data Fig. 9c**). Together,
325 these results demonstrate that dose dependence can increase with community
326 diversity. These findings are consistent with the model prediction that the presence of
327 additional species increases niche overlap, amplifying the effect of inoculation dose on
328 colonization.

Discussion

In this study, we quantified the effect of propagule size on hundreds of simultaneous colonization events during community coalescence. We found that propagule size has persistent effects on colonization for a subset of species in complex communities, even after >30 generations of competition—long enough to effect significant ecological changes in a community regardless of ultimate colonization success[53,54]. Using a consumer-resource model, we showed that the magnitude and duration of dose dependence increase when niche overlap is high, as would be expected in diverse communities. In accordance with model predictions, we observed stronger dose dependence during community coalescence and strain-community mixtures than in two- and three-strain mixtures. Our model and results suggest that neutral-like colonization dynamics can arise through resource competition alone, without a need to invoke more complex phenomena like resource fluctuations[55], higher-order species interactions[56,57], or environmental modification[7,11].

The framework we propose suggests several principles for the design of ecological interventions like microbiome therapeutics[38] and agricultural inoculants[58]. First, our model predicts that introduced species will show higher levels of dose dependence in diverse communities. This prediction suggests that experiments with small, model communities may systematically underestimate the effects of propagule size in more diverse, natural communities, emphasizing the importance of conducting large-scale field experiments and clinical studies to test ecological interventions.

352 Our work identifies niche overlap as the main reason propagule size has stronger
353 impacts on colonization in more diverse communities. While niche overlap is
354 challenging to measure, it is often higher among closely related species, which tend to
355 consume more similar resources[59]. We might therefore predict higher levels of dose
356 dependence when introduced species are taxonomically related to existing residents, as
357 our *Lactococcus* and *Enterococcus* example illustrates. A more precise method for
358 predicting niche overlap in microbial communities is metabolomics, which can be a
359 powerful tool for quantifying resource utilization[46,60,61]. Future studies that combine
360 controlled colonization experiments with metabolomics of microbial isolates will shed
361 further light on the ecological factors that determine colonization success.
362

363 **Materials and Methods**

364

365 **Data availability**

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

367 publication.

368

369 **Code availability**

370 The computer code that performs the analysis is available at

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

372

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

374 Stool samples were collected as a part of a study[62] approved by the Stanford University

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

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

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

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

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

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

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

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

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

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

385 years old.

386

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

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

393

394 We generated *in vitro* communities following methods established in our previous
395 studies[49,50]. All steps of community inoculation and passaging were performed in an
396 anaerobic chamber (Coy Instruments), using sterilized pipette tips. First, a 50-mg
397 fragment of each stool core was resuspended in 500 µL of filter-sterilized PBS. Next, 20
398 µL of the stool resuspension were added to 180 µL of Brain Heart Infusion (BHI, BD
399 Biosciences 237200) supplemented with 0.2 mg/mL L-tryptophan, 1 mg/mL L-arginine,
400 0.5 mg/mL L-cysteine, 5 µg/mL vitamin K, 1 µg/mL haemin, and 5 g/mL mucin (referred
401 to here as mBHI+mucin) in a flat-bottom 96-well plate (Greiner Bio-One 655161). Each
402 sample was inoculated as three replicates in separate wells. All media were autoclaved
403 before inoculation to ensure sterility.

404

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

408 stored each plate of saturated cultures at -80 °C after passaging by mixing 100 μL of
409 saturated culture with 100 μL of 50% filter-sterilized glycerol.
410
411 **Whole community mixtures**
412 We thawed glycerol stocks of one replicate of each community from passage 15 and
413 reinoculated these cultures in an anaerobic chamber (Coy Instruments) by adding 3 μL
414 of glycerol stock to 197 μL of mBHI+mucin. These communities were grown at 37 °C for
415 48 h.
416
417 To create community mixtures, we first generated a dilution series of each of the
418 unmixed parent communities. Fresh mBHI+mucin was added to each saturated culture
419 to create communities with concentrations 1/10, 1/100, and 1/1000 of the parent
420 community. We generated each dilution in triplicate from the single replicate of each
421 parent community, resulting in 24 communities for each dilution factor from the eight
422 parent communities, and 72 dilution communities in total across all three dilution factors.
423
424 To generate each pairwise mixture, we added 1 μL of each dilution community to 1 μL
425 of another parent community and 198 μL of fresh mBHI+mucin. We mixed communities
426 in eight combinations: A1/A2, B1/B2, C1/C2, D1/D2, A1/B1, B1/C1, C1/D1, and D1/A1.
427 Four pairs of communities (A1/A2, B1/B2, C1/C2, D1/D2) were derived from subjects
428 from the same household, while the other four mixtures (A1/B1, B1/C1, C1/D1, D1/A1)
429 involved communities from different households. Each pair of communities was mixed
430 at seven ratios: 1000:1, 100:1, 10:1, 1:1, 1:10, 1:100, and 1:1000. Each mixture was

431 passaged in triplicate. This protocol resulted in 168 co-culture communities (8 mixtures
432 × 7 ratios × 3 replicates), which were passaged in two flat-bottom 96 well plates. We
433 also inoculated 2 μL of each undiluted parent community in 198 μL of mBHI+mucin in
434 duplicate, comprising 16 unmixed control parent communities. The four corners of each
435 96-well plate were filled with 200 μL of mBHI+mucin as blank controls to monitor
436 contamination.
437
438 Both 96-well plates containing mixture co-cultures and unmixed control communities
439 were grown at 37 °C and passaged every 48 h five times by adding 1 μL of saturated
440 culture to 199 μL of mBHI+mucin. Saturated cultures from each passage were stored at
441 -80 °C.
442
443 **16S rRNA gene sequencing**
444 DNA was extracted from 50 μL of each community from each mixture using the Qiagen
445 DNeasy Ultraclean 96 Microbial Kit (Qiagen 10196-4) according to the manufacturer's
446 instructions. Extracted DNA was stored at -20 °C.
447
448 We performed 16S rRNA gene amplicon sequencing targeted to the V4 region.
449 Extracted DNA was amplified using primers modified from the Earth Microbiome Project
450 spanning 515F–806R, with gene-specific sequences GTGYCAGCMGCCGCGGTAA for
451 the forward primer and GGACTACNVGGGTWTCTAAT for the reverse primer[63,64]. The
452 16S region was amplified using 27 cycles of PCR with an annealing temperature of 50
453 °C, and successful amplification of the 16S region was confirmed using gel

electrophoresis. Sequencing primers were attached using 10 cycles of PCR with an annealing temperature of 54 °C. Excess primer was removed using a 0.8X Ampure bead cleanup. Libraries were pooled to contain equal volume from each sample and sequenced using an Illumina NovaSeq SP at the DNA Services Lab, Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign.

Raw reads were annotated and demultiplexed using UMI-tools[65], and primer and adapter sequences were trimmed using cutadapt[66]. To remove the blank media controls, samples with <100 reads were excluded. DADA2[67] was used to filter and truncate reads and to assign amplicon sequence variant (ASV) taxonomy based on the SILVA (release 138) reference database[68], and to obtain a phylogeny of all ASVs.

Our final community coalescence dataset contained 91,504,441 reads with a median of 201,425 reads (standard deviation 69,902) per non-blank sample. Two samples in our dataset (A1 replicate 1 from passage 3, and D1/A1 100:1 replicate 1 from passage 5) received zero sequencing coverage, so these were excluded prior to subsequent analyses. 1715 unique ASVs were detected across all samples, and 198 of these were present at $\geq 10^{-3}$ relative abundance in at least one sample, which we estimated as the limit of detection. Although subjects from the same households often shared a higher proportion of gut microbial strains compared to subjects from different households[62], we observed no differences in the proportion of ASVs shared in mixtures of communities from the same versus different households (**Extended Data Fig. 3**). We used the relative abundance of each ASV to calculate Jensen-Shannon divergence.

477

478 In addition to our community coalescence dataset, we obtained a second dataset
479 representing the *L. garvieae*/*L. lactis* co-cultures after five passages and the D2/*L. lactis*
480 co-cultures after five passages. This dataset contained 48,347,562 reads with a median
481 of 683,158 reads (standard deviation 370,504) per non-blank sample. We also obtained
482 a third dataset representing the *E. faecalis*/*L. lactis*, *E. casseliflavus*/*L. lactis*, three-
483 strain, and all strain-community co-cultures after three and five passages. This third
484 dataset contained 9,429,310 reads with a median of 12,927 reads (standard deviation
485 12,375) per non-blank sample. We excluded one replicate each from the *L. garvieae*/*L.*
486 *lactis*, D1/*L. lactis, L. garvieae*/D2, and *E. casseliflavus*/D2 (passage 5) mixtures based
487 on poor sequencing coverage at one or more mixture ratios in those replicates.

488

489 **Estimation of theoretical community compositions**

490 To assess the effect of propagule size in our community mixtures, we computationally
491 generated a set of theoretical community mixtures according to the predictions of
492 ecological neutral theory[51]. The expected abundance of each ASV across mixture ratios
493 was calculated based on a weighted average of the relative abundance in the parent
494 communities: $\frac{x_1}{A} + \frac{x_2}{B}$, where x_1 and x_2 represent the relative abundance of the ASV in
495 parent community 1 and 2, respectively, and A and B represent the dilution factor of
496 parent community 1 and 2, respectively, at the given mixture ratio. For example, the
497 predicted abundance of an ASV in a 1:100 mixture of A1 and A2 in which the
498 inoculation dose of A2 is 100 times greater than that of A1 is $\frac{x_{A1}}{100} + x_{A2}$. We normalized
499 these weighted averages to calculate a theoretical relative abundance for each ASV in a

500 mixture, and we ignored ASVs with predicted relative abundance $<10^{-3}$, our sequencing
501 limit of detection. We generated two theoretical communities for each pair of parent
502 communities based on the two replicates of each parent community in our experiments.
503 We generated separate sets of theoretical communities from passages 3 and 5 of the
504 parent communities to account for changes to parent community composition that may
505 have occurred during passaging.
506
507 **Categorization of ASVs by colonization behavior**
508 We developed a set of summary statistics for each ASV based on its average change in
509 relative abundance across mixture ratios and the deviation from predicted neutral
510 relative abundances, which enabled us to systematically categorize each ASV into one
511 of six distinct patterns: low abundance, noisy, resident, dose-dependent, strong
512 colonizer, or weak colonizer (**Extended Data Fig. 5**).
513
514 For each ASV in each mixture, we calculated four measures. The average $\log_{10}$ fold
515 change in relative abundance across mixture ratios (net dose difference) and average
516 absolute value of $\log_{10}$ fold change in relative abundance across mixture ratios
517 (magnitude dose difference) were used to quantify each ASV's level of dose
518 dependence. The average $\log_{10}$ fold difference between experimental and theoretical
519 relative abundances across mixture ratios (net theoretical difference) and average
520 absolute value of $\log_{10}$ fold difference between experimental and theoretical relative
521 abundances across mixture ratios (magnitude theoretical difference) were used to

522 quantify each ASV's deviation from neutral predictions and distinguish between
523 resident, strong colonizing, and weak colonizing ASVs.
524
525 Net dose difference was calculated as $\frac{\sum(x_i - x_{i-1})}{n}$, where x_i represents the $\log_{10}$(relative
526 abundance) of an ASV at a single mixture ratio (e.g., 100:1), x_{i-1} represents the
527 $\log_{10}$(relative abundance) at the preceding mixture ratio (e.g., 1000:1), and $n = 7$
528 represents the total number of mixture ratios. Magnitude dose difference was calculated
529 as $\frac{\sum|x_i - x_{i-1}|}{n}$. Both net dose difference and magnitude dose difference were first
530 calculated individually for each replicate of the ASV and then averaged across
531 replicates.
532
533 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
534 average experimental and theoretical, respectively, $\log_{10}$(relative abundance) across
535 replicates at a single mixture ratio. Magnitude theoretical difference was calculated as
536 $\frac{\sum|\bar{x}_a - \bar{x}_t|}{n_{\text{ratio}}}$. For ASVs that were not detected in the co-cultures at some mixture ratios
537 and/or parent communities, we set the relative abundance to 10^{-3}.
538
539 To categorize ASVs, we used five binary decisions based on our summary statistics and
540 other measures (**Extended Data Fig. 5**). First, ASVs were classified as "low
541 abundance" if their relative abundance was $<10^{-2.58}$, a threshold slightly above our limit
542 of detection, in both parent communities. Next, ASVs were classified as "noisy" if the
543 ratio of magnitude dose difference to the absolute value of net dose difference was

>2.25, indicating large, non-monotonic changes in relative abundance over mixture ratios. To prevent ASVs with extremely low change over mixture ratios (absolute value net dose difference <0.1) from artificially inflating the number of noisy ASVs, we replaced these low values with 0.1 as a pseudocount.

If ASVs were not classified as low abundance or noisy, they were classified as “dose-dependent colonizers” if the absolute value of the net dose difference was >0.09, meaning that they exhibited large monotonic changes in relative abundance over mixture ratios. Of the remaining ASVs, if the magnitude theoretical difference was >0.3 and the absolute value of the net theoretical difference was ≥0.2, indicating a strong deviation from neutral predictions, they were classified as “strong colonizers” or “weak colonizers” based on whether the sign of net theoretical difference was positive or negative, respectively. All other ASVs, which showed little dose dependence or deviation from neutral predictions, were classified as “resident”. All numeric cutoffs between categories were chosen to minimize the number of ASVs classified as low abundance, noisy, resident, or low abundance, while ensuring that dose-dependent, strong, and weak colonizers could be readily distinguished by inspection.

Consumer-resource model

We implemented a standard consumer-resource model[52]:

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

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

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**).

586 We investigated how dose dependence would be affected by the addition of a third
587 species. We introduced a third species into the model that consumes a unique
588 resource, $\mu_3 = \mu_2 = 1$. Species 3 also competes with species 1 for a fraction β of
589 resource μ_1, which it consumes at relative rate $r = \frac{R_{3,13}}{R_{1,13}}$ compared to species 1. To
590 investigate how the presence of this new species affects colonization, we varied β from
591 0 to 0.9 while maintaining a constant $r = 10$ (**Fig. 5b**). As before, we initialized the
592 abundances of species 1 and 2 at a range of ratios from 1000:1 to 1:1000, and we set
593 the abundance of species 3 equal to that of species 2. To investigate how the
594 consumption rate of this shared resource affects colonization behavior, we also tested
595 the effect of varying r from 10^{-2} to 10^{2} while maintaining a constant $\beta = 0.5$ (**Extended**
596 **Data Fig. 8d**).
597

598 **Strain isolation**

599 *Lactococcus garvieae*, *Lactococcus lactis*, *Enterococcus faecalis*, and *Enterococcus*
600 *casseliflavus* strains representing dose-dependent ASVs were isolated by sorting single
601 cells from communities D1 and D2 into mBHI-filled wells of a 96-well plate using a
602 fluorescence-activated cell sorter (FACS)[69]. To do so, communities were transported
603 from the anaerobic chamber to the FACS facility in an anaerobic container (BD
604 Biosciences) to ensure minimal oxygen exposure. We confirmed the purity of each
605 strain isolate by streaking multiple times on BHI agar plates supplemented with 5%
606 horse blood and submitted individual colonies for Sanger and whole-genome
607 sequencing. Colonies verified as pure were grown for 48 h in 200 µL of mBHI+mucin in

608 a 96-well plate, and 75 µL of each saturated culture were mixed with 75 µL of 50%
609 glycerol and stocked in plastic crimp vials (ThermoFisher C4011-11) for future use.
610
611 To confirm isolate purity, we performed Sanger sequencing targeted to the V4 region of
612 the 16S rRNA gene. We amplified the 16S gene with 35 cycles of PCR with an
613 annealing temperature of 55 °C using the previously described primers. The resulting
614 PCR products were sent to Elim Biopharmaceuticals for sequencing.
615
616 We determined the species identity of strain isolates using whole-genome sequencing.
617 DNA was extracted from isolate cultures using the DNeasy Ultraclean 96 Microbial Kit
618 (Qiagen 10196-4) according to the manufacturer's instructions. Libraries were prepared
619 for whole-genome sequencing using the Nextera DNA Flex Library Prep Kit (Illumina
620 20018705) according to the manufacturer's instructions. The DNA concentration of each
621 library was quantified using the Quant-iT High-Sensitivity dsDNA Assay Kit (Invitrogen
622 Q33130) and a Synergy HT Microplate Reader (Biotek Instruments), and libraries were
623 pooled at equal DNA input. Pooled libraries were sequenced at the DNA Services Lab,
624 Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign.
625
626 **Co-cultures of strain isolates and strain-community mixtures**
627 We performed a set of four mixtures of two or three strains (*L. garvieae*/*L. lactis*, *E.*
628 *faecalis*/*L. lactis*, *E. casseliflavus*/*L. lactis*, and *E. faecalis*/*E. casseliflavus*/*L. lactis*) and
629 five strain-community mixtures (D1/*L. lactis*, *L. garvieae*/D2, *E. faecalis*/D2, *E.*
630 *casseliflavus*/D2, *E. faecalis*/*E. casseliflavus*/D2) using the glycerol stocks of our four

631 strain isolates and of the D1 and D2 parent communities. We performed mixtures at
632 ratios from 1000:1 to 1:1000 in triplicate as described above for whole community
633 mixtures. For mixtures that included both *E. faecalis* and *E. casseliflavus*, 0.5 μL of each
634 strain culture were used. Mixtures were passaged in mBHI+mucin for 48 h at 37 °C five
635 times.
636
637 **Analysis of pairwise mixture composition using Sanger sequencing data**
638 To analyze the composition of co-cultures from the *L. garvieae*/*L. lactis* mixture after
639 three passages, we performed Sanger sequencing of each co-culture as described
640 above and used the CASEU R package[70] to quantify the relative abundance of both
641 strains at each mixture ratio.
642
643 **Quantification of strain isolate growth**
644 Growth curves were obtained using an Epoch 2 Microplate Spectrophotometer (Biotek
645 Instruments). Strains were grown overnight and then diluted 1:10, 1:100, or 1:1000 into
646 fresh BHI+mucin in 96-well plates. We measured the optical density at 600 nm (OD_{600})
647 of 10 replicates of each strain over 48 h of anaerobic growth at 37 °C with continuous
648 orbital shaking. Instantaneous growth rate was quantified as $d\ \ln(OD_{600})/dt$ using the
649 gcplyr R package[71].

Author contributions

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., and R.R.J. performed the research. D.A.G., K.S.X., J.G.L., J.C.C.V., D.A.P., 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 the paper; and all authors reviewed it before submission.

655 **Acknowledgements**

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

Figures

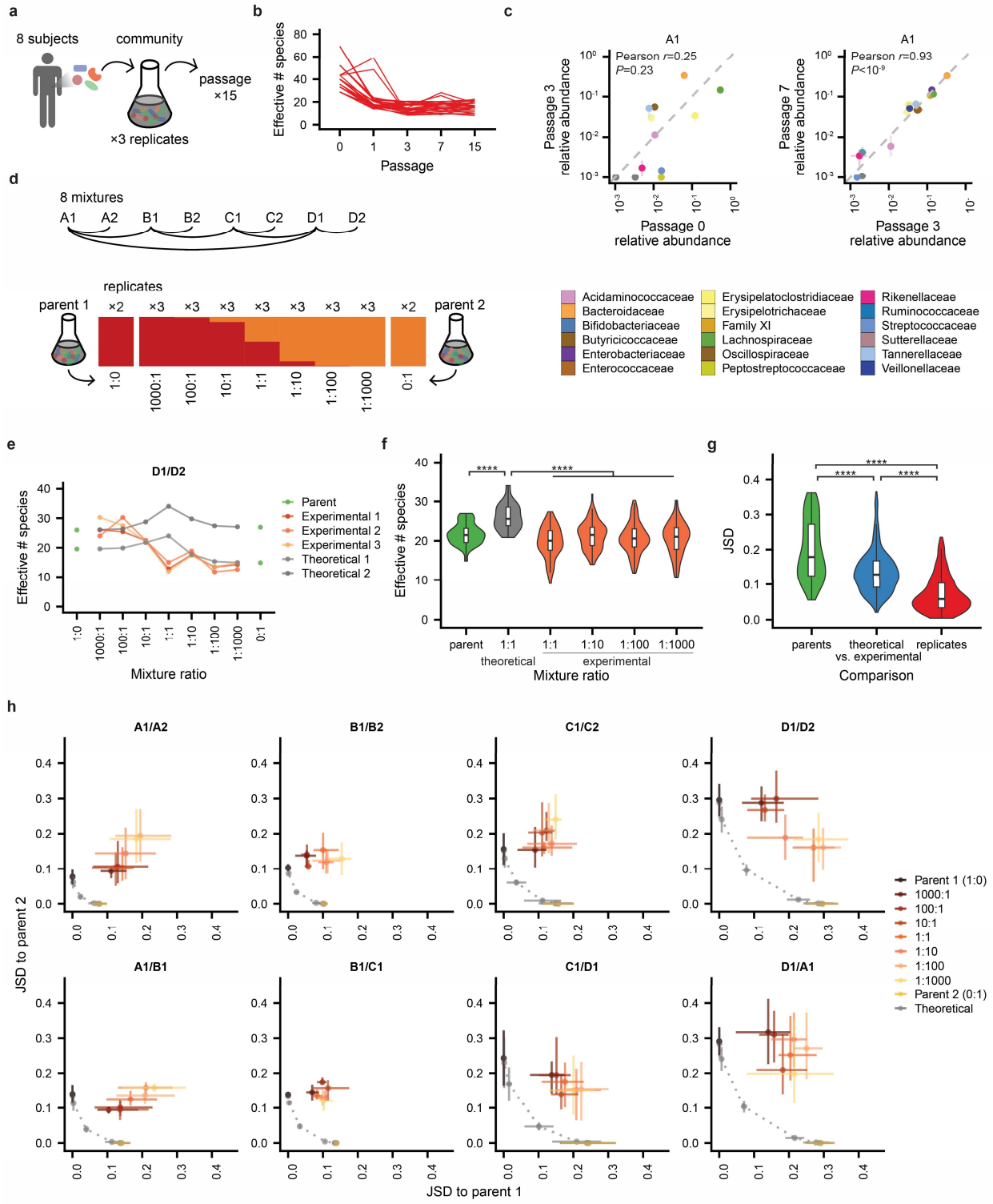

a
8 subjects
community
passage
×15
×3 replicates
b
Effective # species
Passage
c
A1
Pearson r=0.25
P=0.23
Passage 3
relative abundance
Passage 0
relative abundance
A1
Pearson r=0.93
P<10^-9
Passage 7
relative abundance
Passage 3
relative abundance
d
8 mixtures
A1 A2 B1 B2 C1 C2 D1 D2
replicates
parent 1
parent 2
1:0
1000:1
100:1
10:1
1:1
1:10
1:100
1:1000
0:1
Acidaminococcaceae
Bacteroidaceae
Bifidobacteriaceae
Butyricicoccaceae
Enterobacteriaceae
Enterococcaceae
Erysipelatoclostridiaceae
Erysipelotrichaceae
Family XI
Lachnospiraceae
Oscillospiraceae
Peptostreptococcaceae
Rikenellaceae
Ruminococcaceae
Streptococcaceae
Sutterellaceae
Tannerellaceae
Veillonellaceae
e
D1/D2
Effective # species
Mixture ratio
Parent
Experimental 1
Experimental 2
Experimental 3
Theoretical 1
Theoretical 2
f
Effective # species
parent
1:1
theoretical
1:1
1:10
1:100
1:1000
experimental
Mixture ratio
g
JSD
parents
theoretical
vs. experimental
replicates
Comparison
h
A1/A2
B1/B2
C1/C2
D1/D2
A1/B1
B1/C1
C1/D1
D1/A1
JSD to parent 2
JSD to parent 1
Parent 1 (1:0)
1000:1
100:1
10:1
1:1
1:10
1:100
1:1000
Parent 2 (0:1)
Theoretical

Figure 1: Diversity and composition of co-cultures vary by inoculation dose 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) Correlation of family-level relative abundance across passages 0, 3, and 7 (Pearson's tests, left: n=25, ρ=0.25, P=0.23; right: n=22, ρ=0.93, $P<10^{-9}$). Points are colored by family. The dashed line indicates a perfect correlation in relative abundance between passages.

d) 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.

e) 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').

f) Distribution of diversity across parent and co-culture communities. The effective number of species at each mixture ratio more closely resembled the diversity of the parent communities than the higher-diversity theoretical prediction (Wilcoxon rank-

694 sum one-sided tests, parent (n=32) vs. theoretical (n=16) co-cultures: $P=6.1\times10^{-5}$;
695 theoretical vs. experimental 1:1 (n=24) co-cultures: $P=6.6\times10^{-7}$; theoretical vs. other
696 (n=48 at each mixture ratio) co-cultures: $P<10^{-4}$).
697 g) Comparisons of co-culture communities to theoretical communities. Violin plots
698 represent the distribution of Jensen-Shannon divergence (JSD) between all pairs of
699 parent communities (including those that were not mixed), theoretical and
700 experimental mixtures at the 1:1 mixture ratio, and inoculation replicates of the 1:1
701 mixture ratio (significance calculated using Wilcoxon rank-sum one-sided tests,
702 parents (n=56) vs. replicates (n=174): $P<10^{-15}$; parents vs. theoretical-experimental
703 (n=334): $P=1.8\times10^{-7}$; theoretical-experimental vs. replicates: $P<10^{-15}$).
704 h) Variation of community composition across mixtures and mixture ratios. Each plot
705 shows the Jensen-Shannon divergence (JSD) of all co-culture and theoretical
706 communities relative to both parent communities (brown and gold) for a single
707 mixture. Grey points along the dotted line represent theoretical mixtures, and colored
708 points represent data from experimental co-culture communities. Points show
709 averages and error bars show full range of values across inoculation replicates.

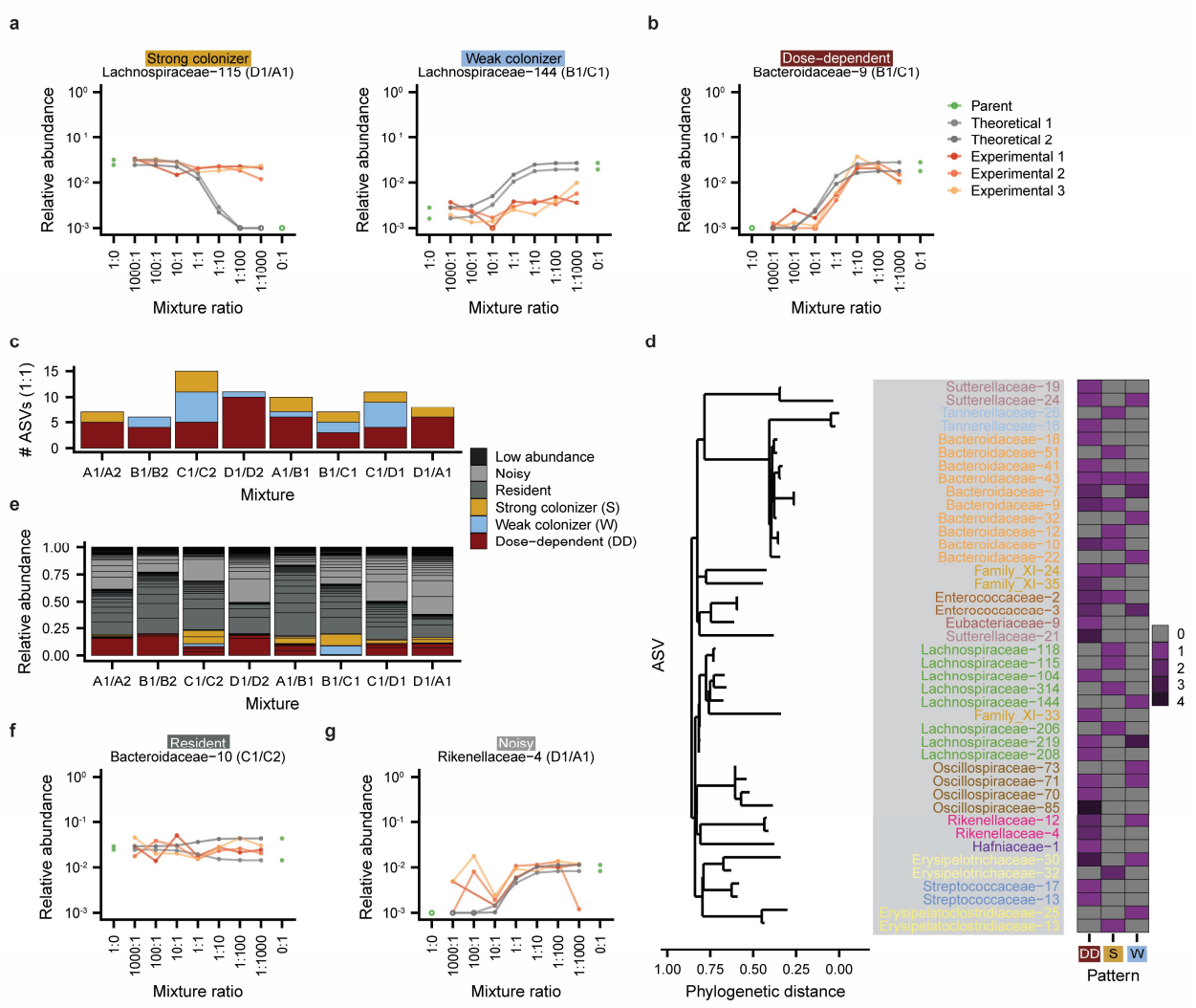


Figure 2: Colonization behaviors of individual ASVs influence community composition.

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**).

c) The number of ASVs that display each colonization pattern in each set of mixtures. Only strong, weak, and dose-dependent colonizers are shown. ASVs were counted as present if they were detected in the 1:1 mixture at the fifth passage.

d) Colonization behavior is not associated with ASV phylogeny. The heatmap shows the number of times each ASV exhibited each colonization pattern in any mixture. Only ASVs exhibiting dose-dependent (DD), strong (S), or weak (W) colonization in any mixture are shown.

e) Total relative abundance of each type of colonizer in each set of mixtures. Only the 1:1 mixture is shown. Black lines indicate separate ASVs. Relative abundances are averaged across the three inoculation replicates.

f) A representative example of a resident ASV that was present at similar relative abundances in both parent communities. Lines are colored as in (**a**).

g) A representative example of a noisy ASV whose relative abundance showed large fluctuations between adjacent mixture ratios. Lines are colored as in (**a**).

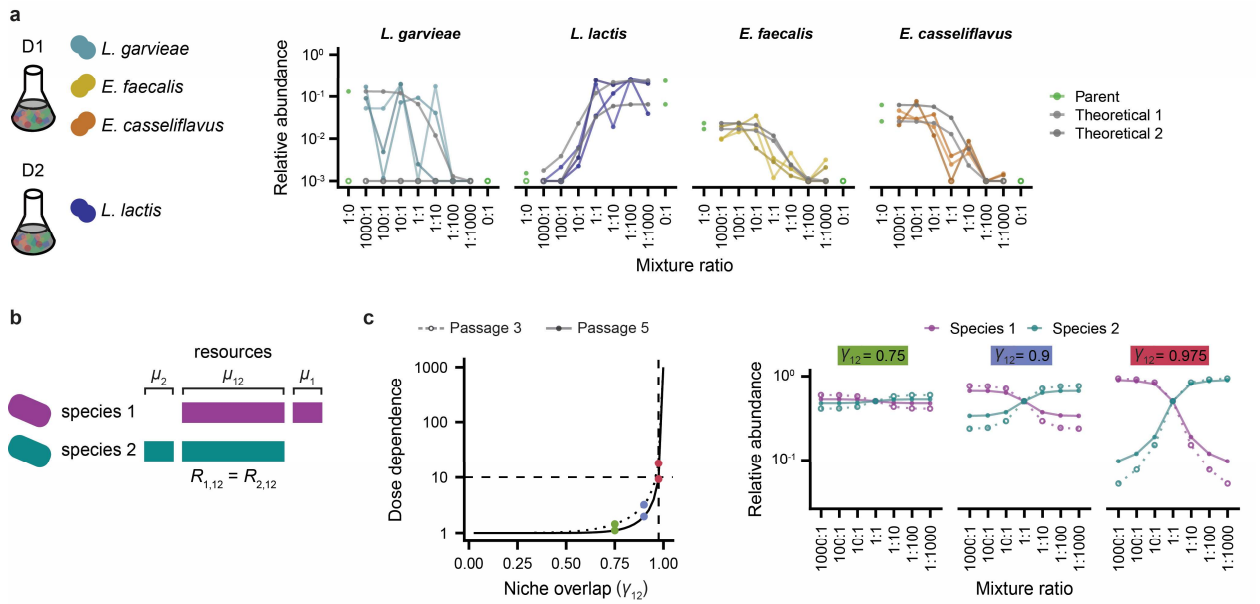


Figure 3: A consumer-resource model predicts transient dose dependence.

a) ASVs from the *Enterococcus* and *Lactococcus* genera showed dose dependence during community coalescence. Origin of strain isolates from the D1/D2 co-cultures. *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 near 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

749 each species are shown at right. Dose dependence is defined as the ratio of relative

750 abundances after passaging starting from a 1000:1 versus 1:1000 mixture ratio.

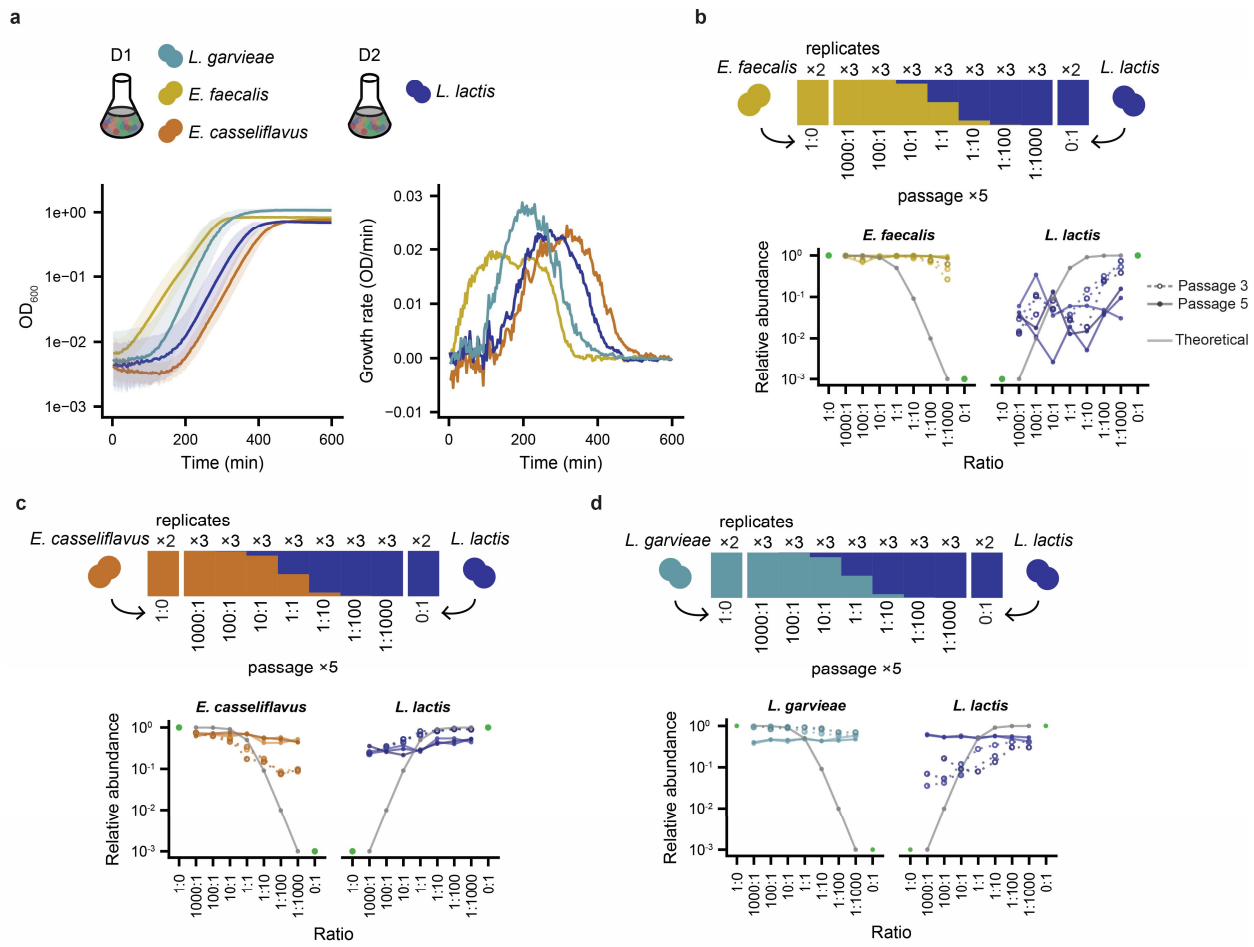


Figure 4: Transient dose-dependent colonization in pairwise strain mixtures.

a) Growth of the *Enterococcus* and *Lactococcus* strains in monoculture. Blank values were subtracted from the OD_{600} measurement of each strain, and growth curves were then $\log_{10}$ transformed. Per-minute growth rates were then calculated from these curves and smoothed over intervals of five minutes (**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).

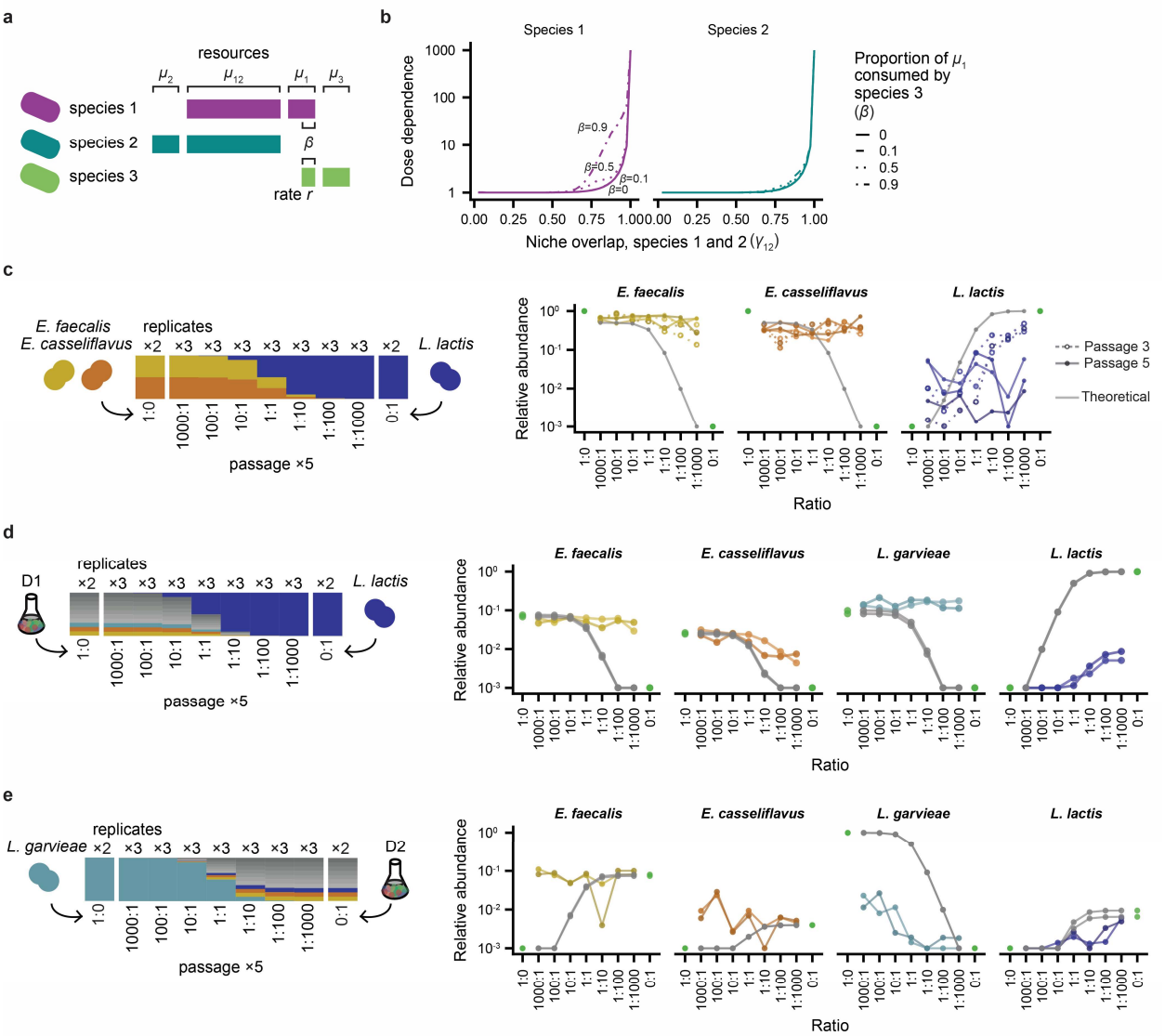


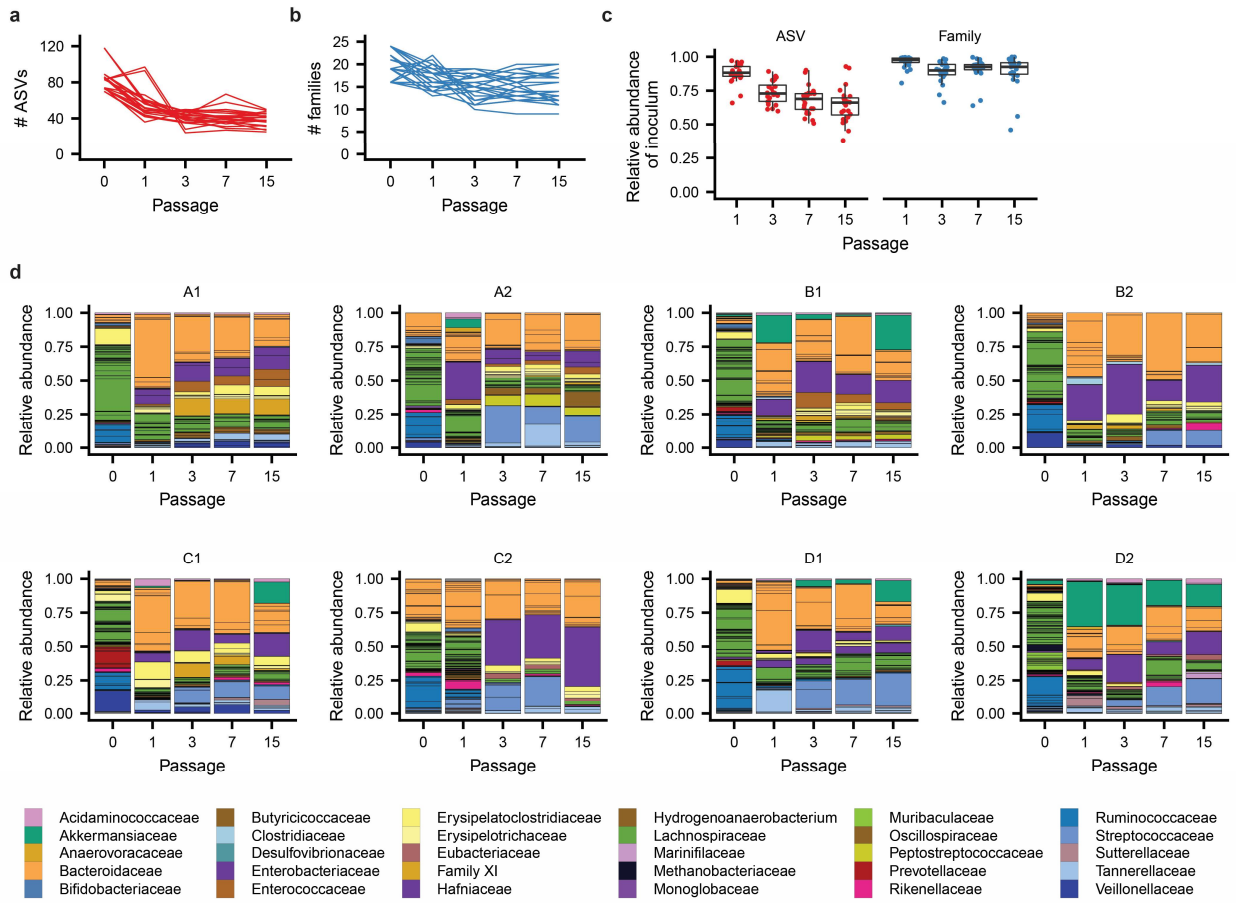
Figure 5: Dose dependence increases in mixtures with more diverse communities.

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 niche overlap between species 1 and 3 (β) at constant $r = 10$, while the dose dependence of species 2 is largely unaffected. Dose dependence is defined as in **Fig. 3c**.

770 c-e) Species relative abundances in co-cultures of (c) *E. faecalis, E. casseliflavus*, and
771 *L. lactis*, (d) the D1 community and *L. lactis*, and (e) *L. garvieae* and the D2
772 community. Lines are colored as in **Fig. 3a**. Dotted lines with open points show data
773 after three passages, and solid lines with closed points show passage five.

Extended Data Figures



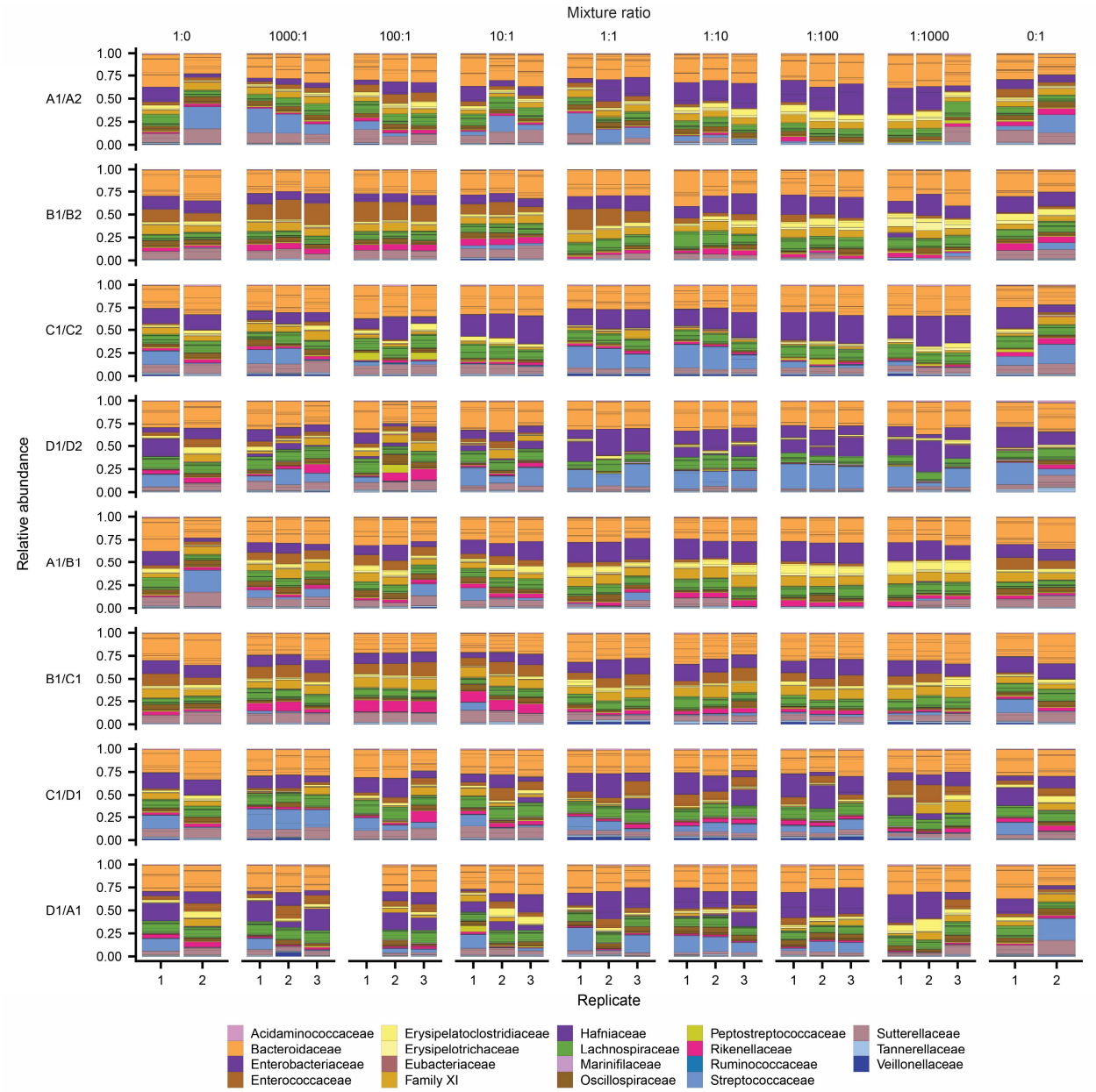
Extended Data Figure 1: Diversity and composition of stool-derived communities.

a) ASV richness of each community over passages. Only ASVs present in the community at relative abundance >10^{-3} were included in the richness calculation.

b) Family-level richness of each community over passages.

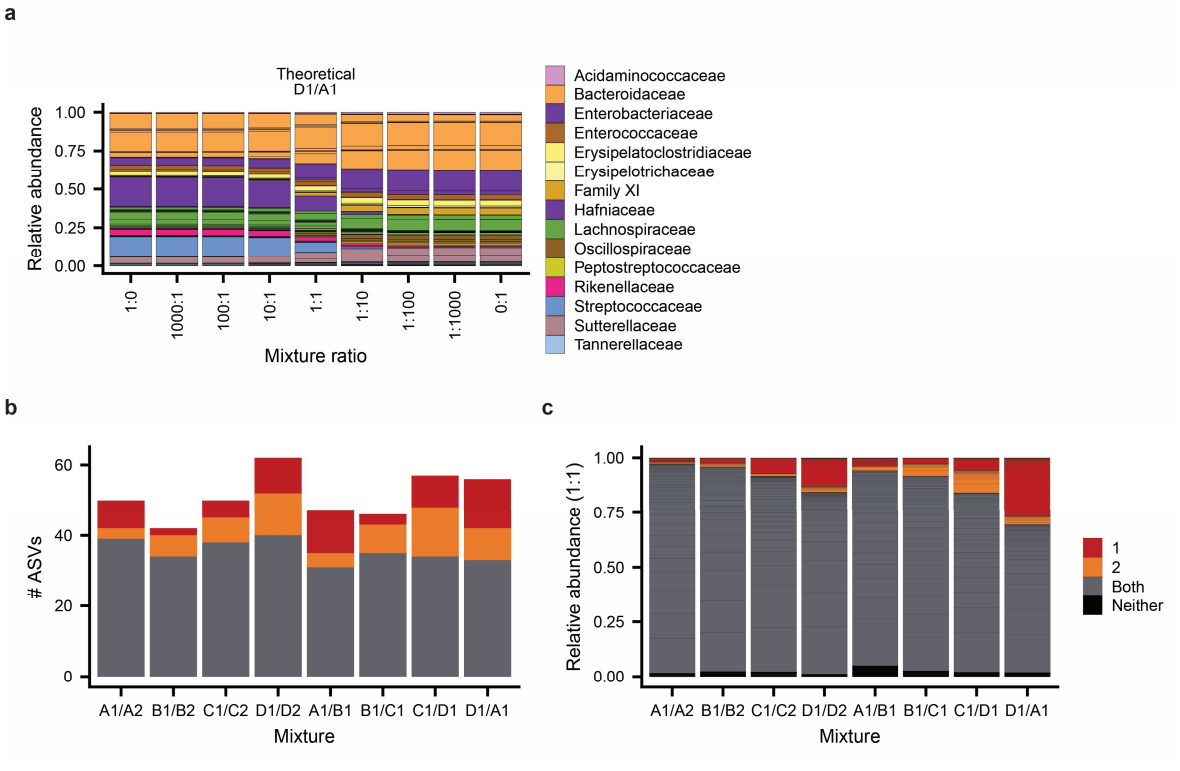
c) Proportion of the stool inoculum (passage 0) retained at each subsequent passage. Each point shows the total ASV or family relative abundance of the stool inoculum represented by taxa retained in communities at a given passage. Taxa were counted as retained if they were detected at a relative abundance >10^{-3}.

785 d) Composition of one replicate of each community over passages. ASV relative
786 abundances are separated with solid lines and colored by family. Only the most
787 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.

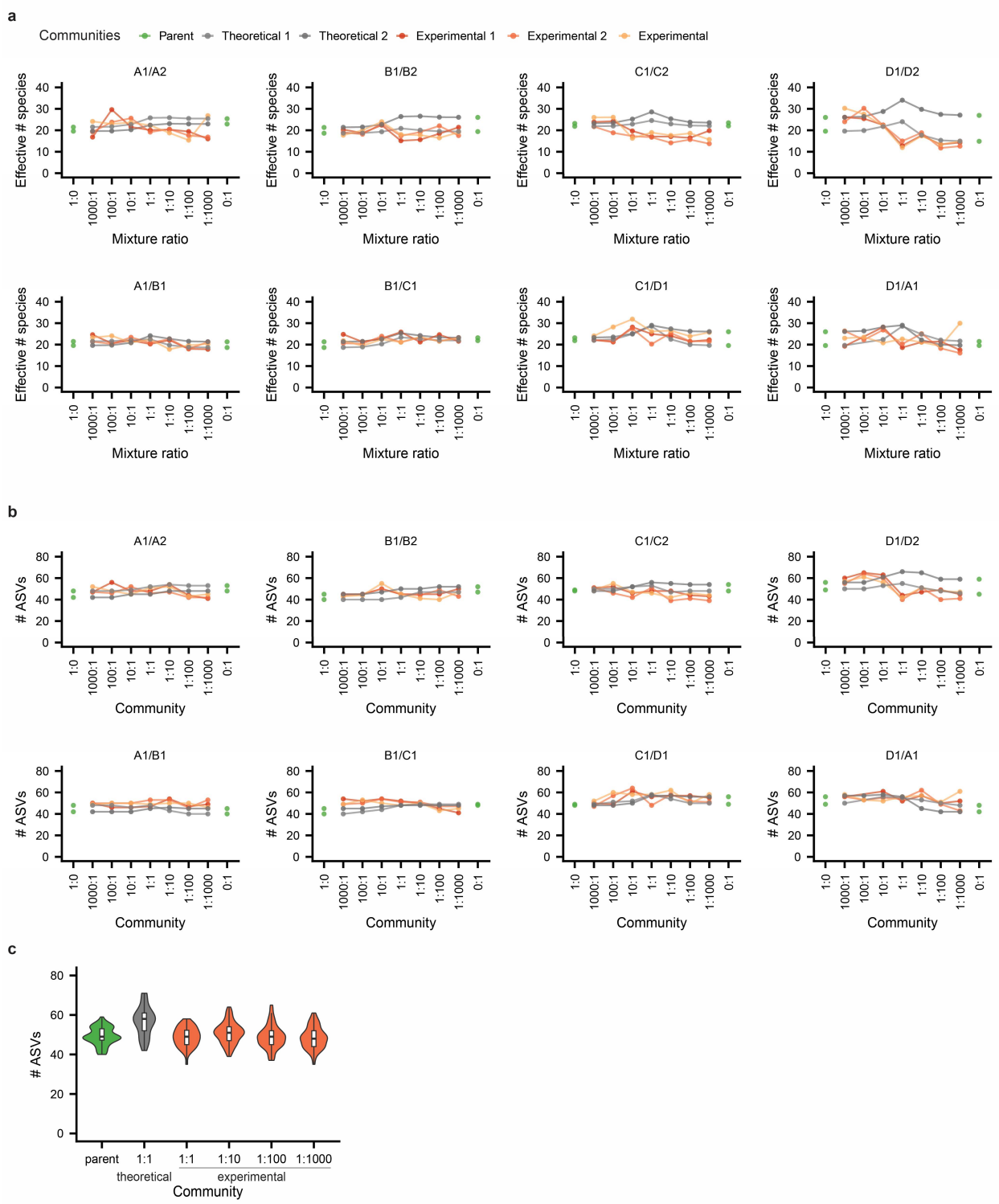


Extended Data Figure 3: Theoretical community composition is dose-dependent, even though many ASVs are shared between parent communities.

a) The theoretical neutral community composition of an example mixture after five passages. 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 are averaged across the three inoculation replicates.


a
Communities
Parent
Theoretical 1
Theoretical 2
Experimental 1
Experimental 2
Experimental
A1/A2
B1/B2
C1/C2
D1/D2
A1/B1
B1/C1
C1/D1
D1/A1
Effective # species
Mixture ratio
1:0
1000:1
100:1
10:1
1:1
1:10
1:100
1:1000
0:1
b
ASVs
Community
c
parent
theoretical
experimental

806

807 **Extended Data Figure 4: Diversity is limited in co-culture communities is robust**

808 **across multiple metrics.**

809 a-b) Diversity of each co-culture community, quantified by (a) the effective number of

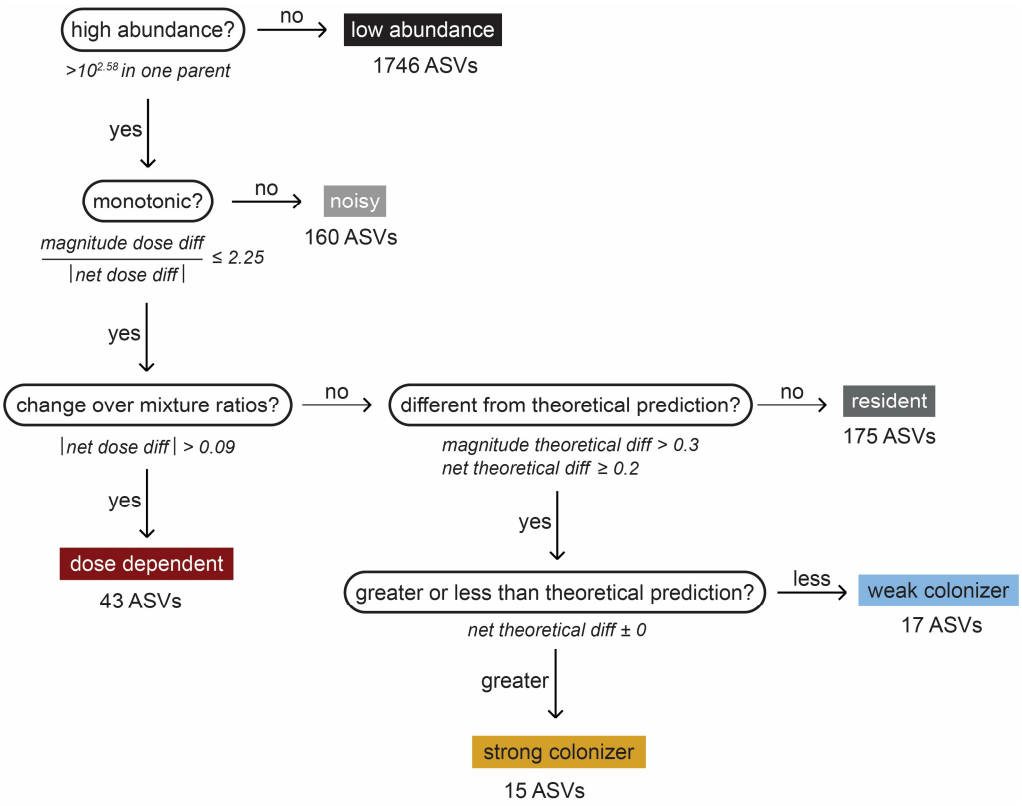
810 species ($e^{H'}$) calculated from the Shannon diversity index (H'), or (b) ASV richness,

811 assuming a limit of detection of 10^{-3} relative abundance. Lines are colored as in **Fig.**

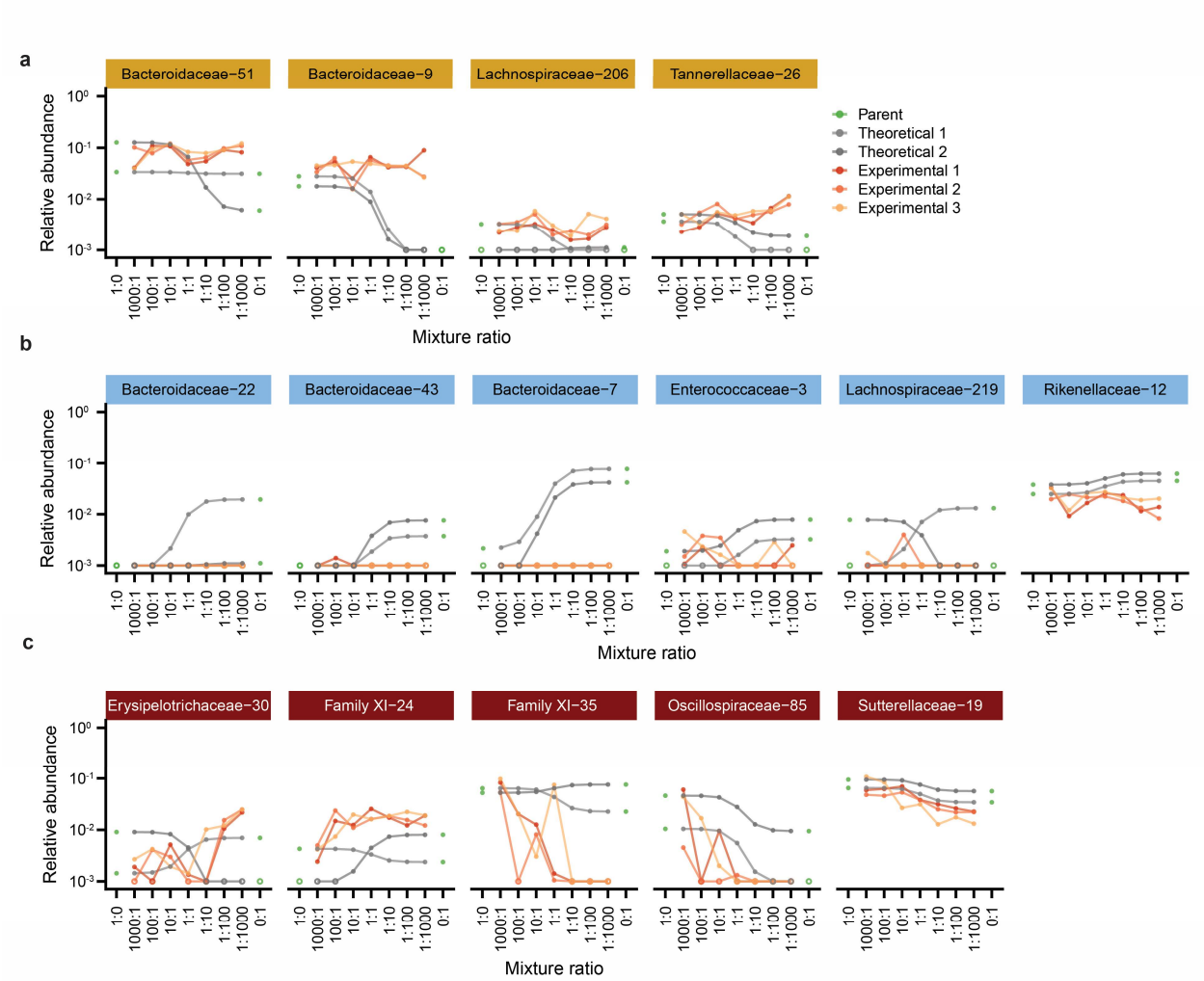
812 **1e**.

813 c) Distribution of ASV richness across parent and co-culture communities.

814

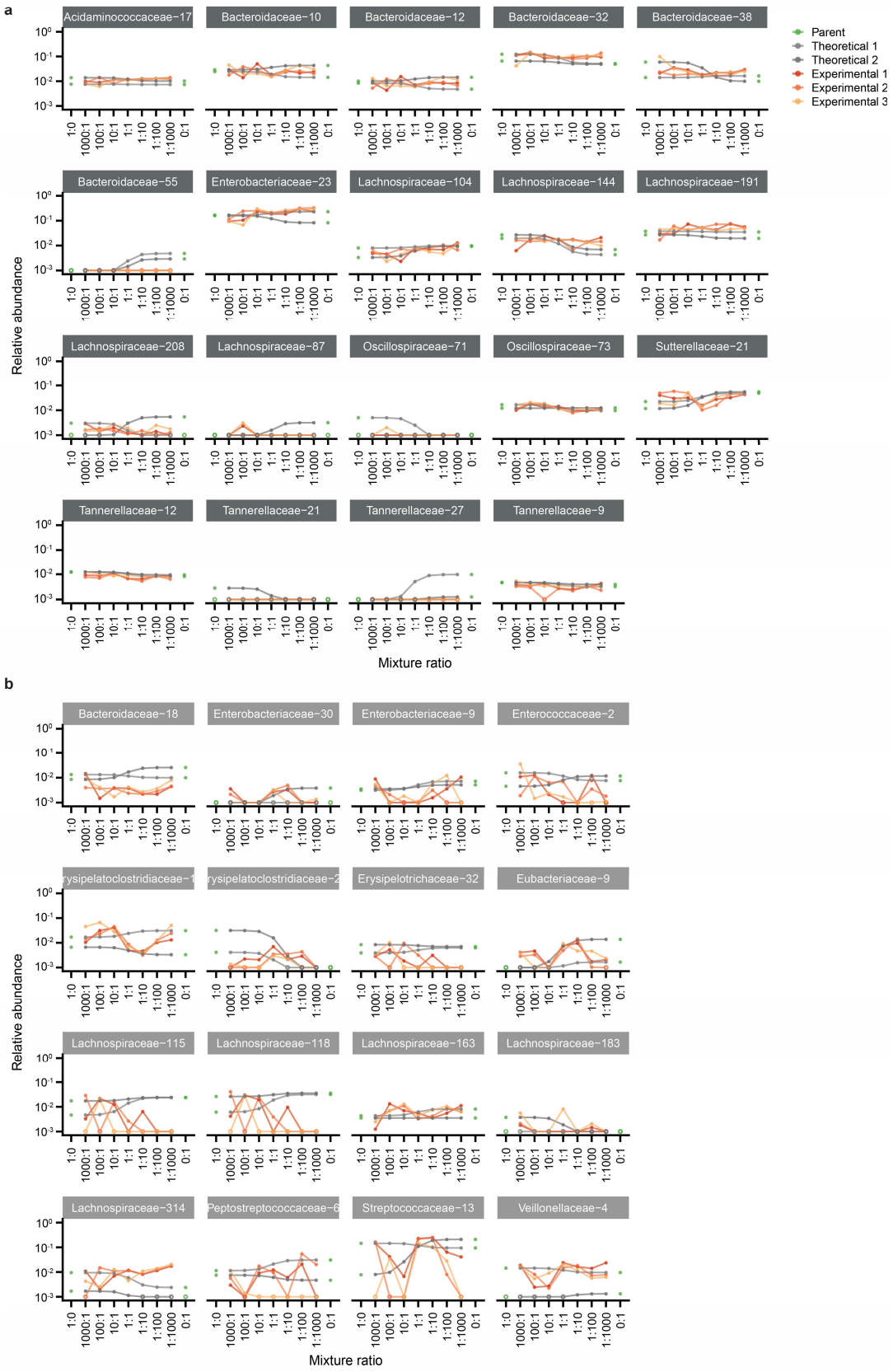


Extended Data Figure 5: Classification scheme for ASV colonization behavior. ASV totals for each category refer to the total number of observations of an ASV with that behavior across all mixtures.



Extended Data Figure 6: Example 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**.

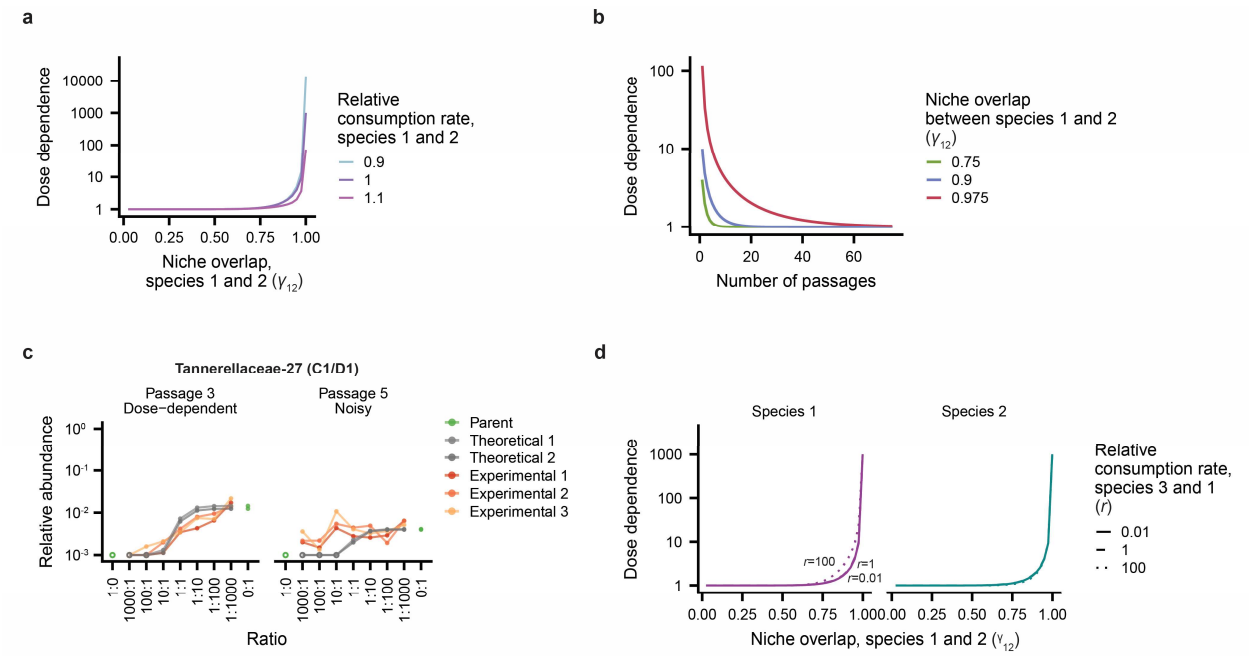

a
Acidaminococcaceae−17
Bacteroidaceae−10
Bacteroidaceae−12
Bacteroidaceae−32
Bacteroidaceae−38
Bacteroidaceae−55
Enterobacteriaceae−23
Lachnospiraceae−104
Lachnospiraceae−144
Lachnospiraceae−191
Lachnospiraceae−208
Lachnospiraceae−87
Oscillospiraceae−71
Oscillospiraceae−73
Sutterellaceae−21
Tannerellaceae−12
Tannerellaceae−21
Tannerellaceae−27
Tannerellaceae−9
Parent
Theoretical 1
Theoretical 2
Experimental 1
Experimental 2
Experimental 3
Relative abundance
Mixture ratio
b
Bacteroidaceae−18
Enterobacteriaceae−30
Enterobacteriaceae−9
Enterococcaceae−2
Erysipelotrichaceae−32
Eubacteriaceae−9
Lachnospiraceae−115
Lachnospiraceae−118
Lachnospiraceae−163
Lachnospiraceae−183
Lachnospiraceae−314
Streptococcaceae−13
Veillonellaceae−4
Relative abundance
Mixture ratio

825

826 **Extended Data Figure 7: Resident and noisy ASVs from the C1/C2 mixture.**

827 a,b) Relative abundance of (a) resident and (b) noisy ASVs from the C1/C2 mixture.

828 Each panel shows one ASV. Lines are colored as in **Fig. 2a**.



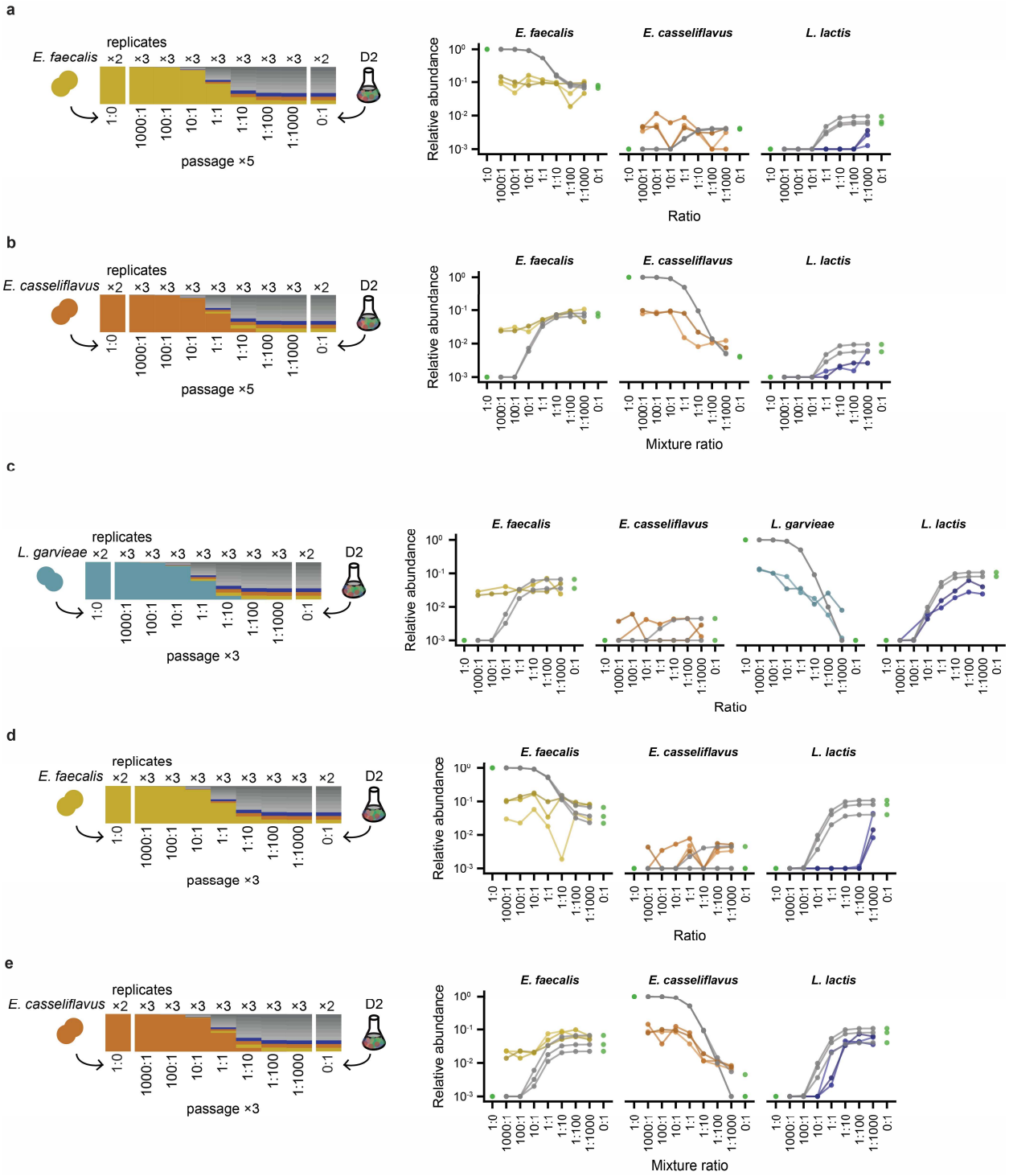
Extended Data Figure 8: Predicted effects of niche overlap and resource consumption rates in a consumer-resource model.

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

b) Dose dependence declines over passages in a two-species consumer resource model (equal consumption rates $R_{1,12} = R_{2,12} = 1$), even when the niche overlap of species 1 and 2 (γ_{12}) is high.

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

d) The dose dependence of species 1 after five passages increases with relative resource consumption rate of species 3 and 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 9: Colonization behavior in strain-community mixtures.

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

848 b) Species relative abundances in the *E. casseliflavus*/D2 co-cultures after five
849 passages. *E. faecalis*, *E. casseliflavus*, and *L. lactis* are classified as dose-
850 dependent.
851 c) Species relative abundances in the *L. garvieae*/D2 co-cultures after three passages.
852 *L. garvieae* and *L. lactis* are classified as dose-dependent.
853 d) Species relative abundances in the *E. faecalis*/D2 co-cultures after three passages.
854 Only *L. lactis* is classified as dose-dependent.
855 e) Species relative abundances in the *E. casseliflavus*/D2 co-cultures after three
856 passages. *E. faecalis*, *E. casseliflavus*, and *L. lactis* are classified as dose-
857 dependent.
858 Lines are colored as in **Fig. 2a**.

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

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

861 healthy adults[62].

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

862

References

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

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

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

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

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

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

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

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

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

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

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

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

13. 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.

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

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

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

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

904 18. Acosta, F., Zamor, R. M., Najar, F. Z., Roe, B. A. & Hambright, K. D. Dynamics of
905 an experimental microbial invasion. *Proc Natl Acad Sci U S A* **112**, 11594–11599
906 (2015).

907 19. Pyšek, P. *et al.* Naturalization of central European plants in North America:
908 Species traits, habitats, propagule pressure, residence time. *Ecology* **96**, 762–774
909 (2015).

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

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

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

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

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

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

926 26. Alzate, A., Onstein, R. E., Etienne, R. S. & Bonte, D. The role of preadaptation,
927 propagule pressure and competition in the colonization of new habitats. *Oikos*
928 **129**, 820–829 (2020).

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

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

934 29. Finlay, B. J. & Clarke, K. J. Ubiquitous dispersal of microbial species. *Scientific*
935 *Correspondence* **400**, 828 (1999).

936 30. Schmidt, T. S. *et al.* Extensive transmission of microbes along the gastrointestinal
937 tract. *Elife* **8**, (2019).

938 31. Barbour, K. M., Barrón-Sandoval, A., Walters, K. E. & Martiny, J. B. H. Towards
939 quantifying microbial dispersal in the environment. *Environ Microbiol* **25**, 137–142
940 (2023).

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

943 33. Mansour, I., Heppell, C. M., Ryo, M. & Rillig, M. C. Application of the microbial
944 community coalescence concept to riverine networks. *Biological Reviews* **93**,
945 1832–1845 (2018).

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

948 35. Castledine, M., Sierocinski, P., Padfield, D. & Buckling, A. Community
949 coalescence: An eco-evolutionary perspective. *Philosophical Transactions of the*
950 *Royal Society B: Biological Sciences* **375**, (2020).

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

954 37. Diaz-Colunga, J. *et al.* Top-down and bottom-up cohesiveness in microbial
955 community coalescence. *PNAS* **119**, (2022).

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

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

962 40. Gianotti, L. *et al.* A randomized double-blind trial on perioperative administration
963 of probiotics in colorectal cancer patients. *World J Gastroenterol* **16**, 167–175
964 (2010).

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

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

43. 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).

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

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

46. 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.

47. 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.

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

49. 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).

50. 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).

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

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

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

54. 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).

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

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

1004 57. Chang, C.-Y., Bajić, D., Vila, J. C. C., Estrela, S. & Sanchez, A. Emergent
1005 coexistence in multispecies microbial communities. *Science (1979)* **381**, (2023).

1006 58. O'Callaghan, M., Ballard, R. A. & Wright, D. Soil microbial inoculants for
1007 sustainable agriculture: Limitations and opportunities. *Soil Use Manag* **38**, 1340–
1008 1369 (2022).

1009 59. Pudlo, N. A. *et al.* Phenotypic and Genomic Diversification in Complex
1010 Carbohydrate-Degrading Human Gut Bacteria. *mSystems* **7**, (2022).

1011 60. Han, S. *et al.* A metabolomics pipeline for the mechanistic interrogation of the gut
1012 microbiome. *Nature* **595**, 415–420 (2021).

1013 61. Connors, B. M. *et al.* Model-guided design of the diversity of a synthetic human
1014 gut community. *bioRxiv* (2022) doi:10.1101/2022.03.14.484355.

1015 62. Xue, K. S. *et al.* Prolonged delays in human microbiota transmission after a
1016 controlled antibiotic perturbation. *bioRxiv* (2023) doi:10.1101/2023.09.26.559480.

1017 63. Apprill, A., Mcnally, S., Parsons, R. & Weber, L. Minor revision to V4 region SSU
1018 rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton.
1019 *Aquatic Microbial Ecology* **75**, 129–137 (2015).

1020 64. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: Assessing
1021 small subunit rRNA primers for marine microbiomes with mock communities, time
1022 series and global field samples. *Environ Microbiol* **18**, 1403–1414 (2016).

1023 65. Smith, T., Heger, A. & Sudbery, I. UMI-tools: Modeling sequencing errors in
1024 Unique Molecular Identifiers to improve quantification accuracy. *Genome Res* **27**,
1025 491–499 (2017).

1026 66. Martin, M. Cutadapt removes adapter sequences from high-throughput
1027 sequencing reads. *EMBnet J* **17**, 10–12 (2011).

1028 67. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina
1029 amplicon data. *Nat Methods* **13**, 581–583 (2016).

1030 68. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data
1031 processing and web-based tools. *Nucleic Acids Res* **41**, (2013).

1032 69. Shiver, A. L., Culver, R., Deutschbauer, A. M. & Huang, K. C. Rapid ordering of
1033 barcoded transposon insertion libraries of anaerobic bacteria. *Nat Protoc* **16**,
1034 3049–3071 (2021).

1035 70. Cermak, N., Datta, M. Sen & Conwill, A. Rapid, Inexpensive Measurement of
1036 Synthetic Bacterial Community Composition by Sanger Sequencing of Amplicon
1037 Mixtures. *iScience* **23**, (2020).

1038 71. Blazanin, M. gcplyr: an R package for microbial growth curve data analysis.
1039 *bioRxiv* (2023) doi:10.1101/2023.04.30.538883.