RESEARCH ARTICLE SUMMARY

MICROBIOTA

Imbalance in gut microbial interactions as a marker of health and disease

Roberto Corral López, Juan A. Bonachela*†, Maria Gloria Dominguez-Bello, Michael Manhart, Simon A. Levin, Martin J. Blaser, Miguel A. Muñoz*†

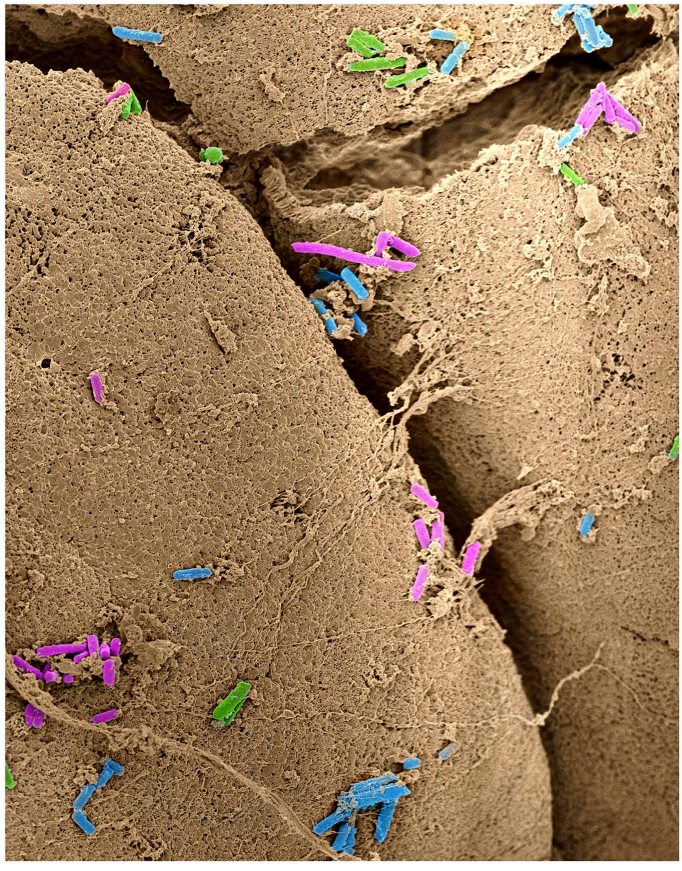
Full article and list of author affiliations: https://doi.org/10.1126/science.ady1729

INTRODUCTION: The human gut microbiome is a complex ecological system crucial for host health. Dysbiosis—i.e., the imbalance of gut microbial communities—is associated with a wide range of diseases, including obesity, diabetes, inflammatory bowel disease (IBD), *Clostridioides difficile* infection (CDI), irritable bowel syndrome (IBS), and colorectal cancer (CRC). Therapeutic approaches, such as fecal microbiota transplantation, dietary interventions, and probiotics, aim to restore balance by reshaping community composition. However, their outcomes remain inconsistent and unpredictable, in part owing to our limited understanding of the metabolic and ecological interactions that govern microbiome dynamics. The latter has also prevented the development of robust biomarkers to distinguish health from disease.

RATIONALE: Previous studies have suggested that health and dysbiosis may represent alternative community states but have not provided a mechanistic justification. Most efforts to define dysbiosis aim to identify bacterial taxa or functions that may differ between healthy and diseased communities or assume that reduced diversity is a universal hallmark of disease. However, such signatures vary across conditions and cohorts and fail to capture the ecological principles that shape disease states. To provide a more mechanistic understanding of gut microbial dynamics in health and disease, we developed a metabolically explicit model in which bacterial interactions arise naturally from competition for shared resources and cross-feeding.

RESULTS: The model reproduces key macroecological patterns and captures the functional redundancy characteristic of real gut microbiomes. Moreover, our model revealed the emergence of two distinct ecological states (healthy and dysbiotic states) whose α and β diversities, dominance indices, and numbers of functions and excreted metabolites closely resembled those observed in real microbiomes. The healthy state was dominated by competitive interactions, whereas the dysbiotic state was shaped by tightly connected cross-feeding consortia. We also developed the ecological network balance index (ENBI), a metric that measures the relative contribution of positive versus negative interactions and reliably separates healthy from dysbiotic states. Calculating the ENBI for the model and metagenomic data for IBD, IBS, CDI, and CRC showed that, in all cases, diseased microbiomes exhibited higher ENBI values. Our metric also correlated with disease stage. These results proved robust across subsampling, geography, taxonomic levels, and profiling methods.

CONCLUSION: Unlike diversity-based metrics, which vary across diseases and cohorts, the ENBI consistently distinguishes healthy from diseased states and even tracks disease progression, offering a path toward robust, noninvasive early warning indicators of disease. The ENBI also provides mechanistic insights: Our results show that dysbiosis is associated with a shift in the community interaction network, with positive interactions increasingly dominating over negative ones. Our framework is both general and extensible, and thus it can be adapted to incorporate additional biological features for the study of specific gut phenomena and be readily applied to other microbiomes (from the vaginal and oral to plant and soil ecosystems), or it can be used for the study and prediction of potential outcomes in therapeutic interventions. By linking microbial ecology with clinical research, our framework advances precision medicine and supports the development of more personalized strategies to maintain or restore gut health. □



Consortia of gut bacterial cells on villus tips (projections from intestinal mucosa). Gut bacterial cells are shown in assorted colors, and the villus tips are brown. The color highlights the contrasting size of the bacteria and the barren-appearing cliffs of intestinal cells and chasms. Our model and metric reveal that positive interactions dominate in diseased gut microbiomes, whereas negative interactions dominate in healthy ones. We track disease emergence and progression monitoring this ecological balance. [Image credit: X. Zhang; edits: A. J. Roldán]

*Corresponding author. Email: mamunoz@onsager.ugr.es (M.A.M.); juan.bonachela@rutgers.edu (J.A.B.) †These authors contributed equally to this work. Cite this article as R. Corral López *et al.*, *Science* **391**, 890 (2026). DOI: 10.1126/science.ady1729

MICROBIOTA

Imbalance in gut microbial interactions as a marker of health and disease

Roberto Corral López[1,2], Juan A. Bonachela[3]*†, Maria Gloria Dominguez-Bello[4], Michael Manhart[5,6], Simon A. Levin[7], Martin J. Blaser[5], Miguel A. Muñoz[1,2]*†

Imbalances in the human gut microbiome, or dysbioses, are associated with multiple diseases but remain poorly understood. Existing biomarkers of dysbiosis fail to capture the ecological mechanisms that differentiate healthy from diseased microbiomes. We have developed a metric, the ecological network balance index (ENBI), that quantifies the balance between positive and negative microbial interactions. This metric was inspired by the phenomenology observed in a model for gut microbiome dynamics that we introduce in this work, which revealed alternative stable states with distinct emergent microbial communities: a healthy state dominated by negative interactions and a dysbiotic state dominated by positive interactions. The ENBI robustly differentiates these states in both simulated and empirical datasets spanning multiple diseases and correlates with disease progression in conditions such as colorectal cancer, which underscores its potential as a diagnostic tool.

The human gut microbiome is a dynamic and complex system, comprising assemblages of diverse microorganisms, that plays a crucial role in maintaining health through its interactions with the host and environment (*1*–*4*). Innovations in DNA sequencing techniques over the past two decades have improved our ability to quantify microbiome composition and functional capabilities (*5*–*7*). These studies have shown that imbalances in the gut microbiome, or dysbioses, associate with several diseases, including obesity, diabetes, inflammatory bowel disease (IBD), *Clostridioides difficile* infection (CDI), irritable bowel syndrome (IBS), colorectal cancer (CRC), liver disease, pancreatic disorders, psoriatic arthritis, and celiac disease (*1*, *2*, *5*, *8*–*13*). Treatments, such as fecal microbiota transplantation (FMT), diet-based interventions, and probiotics (*14*–*16*), aim to restore microbial balance and reverse dysbiosis by altering community composition; however, these treatments show highly variable and unpredictable success rates. One reason for this unpredictability is the limited understanding of the mechanisms underlying gut microbiome dynamics and structure. Specifically, the metabolic and ecological interactions among microorganisms are poorly understood (*1*, *17*, *18*). This knowledge gap has hindered the identification of microbiome biomarkers that correlate with—and accurately differentiate—health and disease. Our work seeks to address both of these fundamental challenges.

[1]Departamento de Electromagnetismo y Física de la Materia, Universidad de Granada, Granada, Spain. [2]Instituto Carlos I de Física Teórica, Universidad de Granada, Granada, Spain. [3]Department of Ecology, Evolution, and Natural Resources, Rutgers University, New Brunswick, NJ, USA. [4]Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ, USA. [5]Center for Advanced Biotechnology and Medicine, Rutgers University, Piscataway, NJ, USA. [6]Department of Biochemistry and Molecular Biology, Rutgers RWJ Medical School, Piscataway, NJ, USA. [7]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. ***Corresponding author. Email: mamunoz@onsager.ugr.es (M.A.M.); juan.bonachela@rutgers.edu (J.A.B.)** †These authors contributed equally to this work.

The search for a universal indicator of dysbiosis

Dysbiosis is a broad concept that encompasses various compositional and functional attributes across diseases, and it is not unambiguously characterized (*10*). Early efforts focused on identifying single bacterial species or functions whose altered levels correlate with a particular disease (*8*, *10*, *19*, *20*). However, dysbiosis typically affects multiple species and functions and can only be captured by systemic, whole-community approaches (*13*, *21*, *22*). A common claim across studies and conditions is that disease correlates with reduced microbial diversity, which is often suggested as a universal marker of dysbiosis (*5*, *11*, *12*, *23*–*25*). We reexamined empirical data from several diseases and found a more nuanced pattern (Fig. 1), in line with recent studies (*10*, *13*, *21*, *26*). Specifically, although IBD and CDI (Fig. 1, A and B) consistently show reduced microbial diversity compared with healthy microbiomes (*13*, *27*, *28*), conditions such as IBS and CRC (Fig. 1, C and D) show a range of microbial diversity and were less clear-cut (*13*, *29*, *30*). Therefore, the question remains as to whether there are conserved mechanisms driving dysbiosis, which would enable the identification of robust disease indicators.

The dichotomy between health and disease has fueled the hypothesis that dysbiosis emerges from shifts between alternative stable states—that is, between distinct taxonomic or functional configurations of the gut microbiome (*31*–*37*). However, this idea remains under debate owing to the varied methodologies used to define and measure such states; the context-dependent nature of the microbiome; and the challenges of controlling multiple external factors, such as diet, medication, and lifestyle (*5*, *8*, *38*). All these factors influence gut composition and function, which complicates the identification of stable states and consistent disease markers.

Theoretical models, nonetheless, allow us to fine-tune definitions and environmental factors (*18*, *38*–*41*). For example, models have identified microbial taxa capable of inhibiting *C. difficile* (CD), a prediction subsequently validated experimentally to confirm that the identified taxa indeed conferred CD colonization resistance (*42*). In another example, machine learning models have systematically tested competing hypotheses regarding the success of FMT (*43*). Despite these advances, few models designed to study gut microbiome dynamics predict or generate alternative stable states, which limits our understanding of their emergence and fundamental properties (*39*, *44*). To address this gap, we introduce a mathematical model specifically designed to explore the ecological dynamics of the gut microbiome and mechanisms underlying its emerging collective states [Fig. 2A and materials and methods (*45*)].

A model for the dynamics of the gut microbiome

Our consumer-resource model describes bacterial and nutrient dynamics and incorporates the possibility of cross-feeding as new metabolites are produced by different bacterial strains (see materials and methods). Nutrients are introduced at a rate h (representing host diet); new bacterial species immigrate at a rate U; and bacteria, nutrients, and metabolites are removed at a dilution rate δ, which determines intestinal transit time.

We also aim to capture the role that metabolic pathways play in shaping gut communities (*46*–*50*). Thus, in our model, each bacterial lineage is characterized by the set of metabolic transformations that it can perform, defined by the network of possible conversions between nutrients and metabolites, along with the metabolic preferences specific to the bacterial strain. Although, for simplicity, only catabolic reactions are explicitly modeled, we use a trade-off function to assign a cost to the set of available pathways; this cost arises from the energy allocated to constructive metabolism and therefore effectively implements anabolism. Thus, our model captures both the energy gain and enzymatic cost for each metabolic reaction.

To model the metabolic conversion of a substance S_α into another substance S_β, we monitor (i) the free energy ($E_{\alpha\beta} > 0$) released during

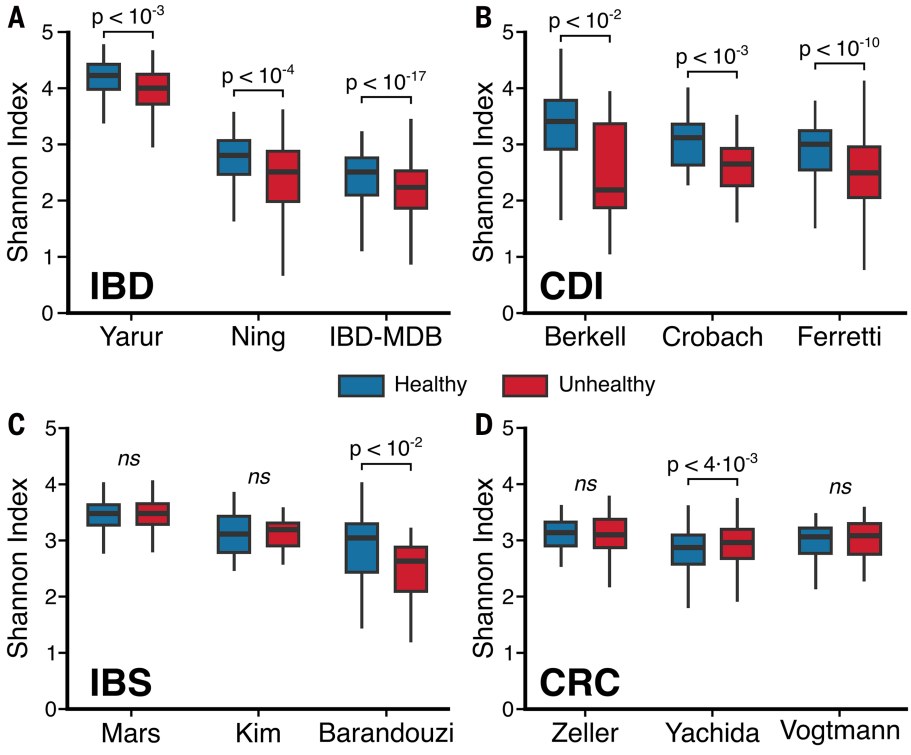


Fig. 1. α diversity in healthy and unhealthy individuals. (**A** to **D**) Comparison of the Shannon index [see materials and methods (*45*)] obtained using existing metagenomic analyses from healthy individuals (blue) versus individuals affected by a specific disease (red): (A) IBD, using data from Yarur *et al*., Ning *et al*., and the Human Microbiome Project (*84*–*86*). (B) CDI, using data from Berkell *et al*., Crobach *et al*., and Ferretti *et al*. (*87*–*89*). (C) IBS, using data from Mars *et al*., Kim *et al*., and Barandouzi *et al*. (*90*–*92*). (D) CRC, using data from Zeller *et al*., Yachida *et al*., and Vogtmann *et al*. (*30*, *93*–*95*). A significant decrease in diversity is evident for IBD and CDI but not for IBS or CRC (statistical significance calculated with two-sided Mann-Whitney *U* test; ns, not significant).

the process and (ii) the enzymatic cost ($C_{\alpha\beta} > 0$), reflecting the number and complexity of enzymes involved. Each bacterial lineage i is defined by a unique pathway matrix $P^{i}_{\alpha\beta}$ (random binary matrix with the metabolic routes available to that bacterial lineage) and a coefficient γ_i that controls the fraction of free energy used for biomass growth. Finally, we used a bottom-up approach to parameterize the model, by which we assigned biologically plausible values to the parameters instead of relying on data fitting exercises. A detailed description of the model is included in the materials and methods, including the definition of species, nutrients, metabolites, the implementation of bacteria-specific metabolic preferences, species immigration, pathways, trade-off function, and chosen parameterization.

Our model successfully reproduced well-known macroecological patterns of the gut microbiome that are observed empirically across multiple systems. In particular, the 95% confidence interval obtained across replicates of the model captures the distributions of species abundance mean and fluctuations, Taylor's law (i.e., the relationship between mean and variance), species prevalence (the distribution of the fraction of samples in which a species is present), and the dependence of prevalence on mean abundance (fig. S1) (*51*, *52*). Additionally, fig. S2 shows that our model replicates the functional redundancy observed in the gut microbiome, where multiple taxonomic compositions encode the same metabolic functions (*53*). In our framework, functions are represented by metabolites grouped into coarse-grained energy levels (*45*). These broader functional categories are akin to Clusters of Orthologous Groups of proteins (COG) or Kyoto Encyclopedia of Genes and Genomes (KEGG) classifications for empirical data (*54*, *55*). Functional redundancy becomes apparent at sufficiently coarse-grained resolutions for both our model and observed data.

Alternative states in the gut microbiome model may reflect health and disease

The dynamics of our simulated gut microbiome community suggest two qualitatively distinct classes of states (Fig. 2B): one characterized by the rapid turnover of many bacterial lineages and another where only a few lineages dominate with quasi-stationary abundances (*45*). We labeled these two classes of states as healthy and dysbiotic, respectively, for reasons that will become apparent in comparison with real data. Notably, these alternative states emerge consistently across a wide range of parameter values, which indicates that they result from the mechanisms encoded in the model rather than specific parameterizations (figs. S3 to S6).

Existing community-level theoretical models, such as the generalized Lotka-Volterra (gLV) (*56*, *57*) and consumer-resource models with random interactions (*58*, *59*), report dynamics reminiscent of either one of our two states (often referred to as "chaotic fluctuations" and "stable fixed points"). However, these models do not capture multistability between the different states. Explicitly incorporating metabolic pathways and letting the structure of the community emerge from the dynamic interactions allowed our model to capture diverse alternative states and the dynamic transitions between them.

Compared with the healthy state, the dysbiotic state in our model exhibited a significant reduction in α diversity (measured by the Shannon index; richness or number of lineages; Simpson's index; and Pileou's evenness) (Fig. 2, C and D, and fig. S7A), the number of enzymes (total enzymatic cost of all active pathways), the number of secreted metabolites (substances with nonzero concentrations), and an increase in dominance (relative abundance of most abundant bacteria) (*45*). These findings indicated that the microbial community in the dysbiotic state operates with greater efficiency compared with the community in the healthy state because fewer bacterial lineages and metabolic pathways are required to form consortia able to convert a larger proportion of the available energy into microbial biomass (fig. S8).

Given the dichotomy between these two distinct states, we also measured these biomarkers in healthy individuals and patients with IBD. The resulting pattern mirrors our model predictions (Fig. 2E and fig. S7B), supporting our labeling of the two alternative states of the model as healthy and dysbiotic. The parallelism of model states with real healthy and dysbiotic states is reinforced by the observation that in the model and in IBD, the shift toward dysbiosis involves the loss of specific metabolic pathways, resulting in a disruption of functional capacity (that is, some functions may be impaired or entirely lost, whereas others become overrepresented; Fig. 2, F and G) (*60*). Furthermore, for our model, as well as for a range of diseases, β diversity across healthy states was significantly lower than for dysbiotic states (fig. S9), which indicates that healthy microbiomes are more similar to each other than the similarity shown by dysbiotic communities.

The patterns mentioned above, however, are subtle or inconsistent across diseases (see, for example, β diversity for CRC in fig. S12D or the number of enzymes, diversity of secreted metabolites, richness, dominance, Simpson's index, and evenness for IBS and CRC in figs. S10

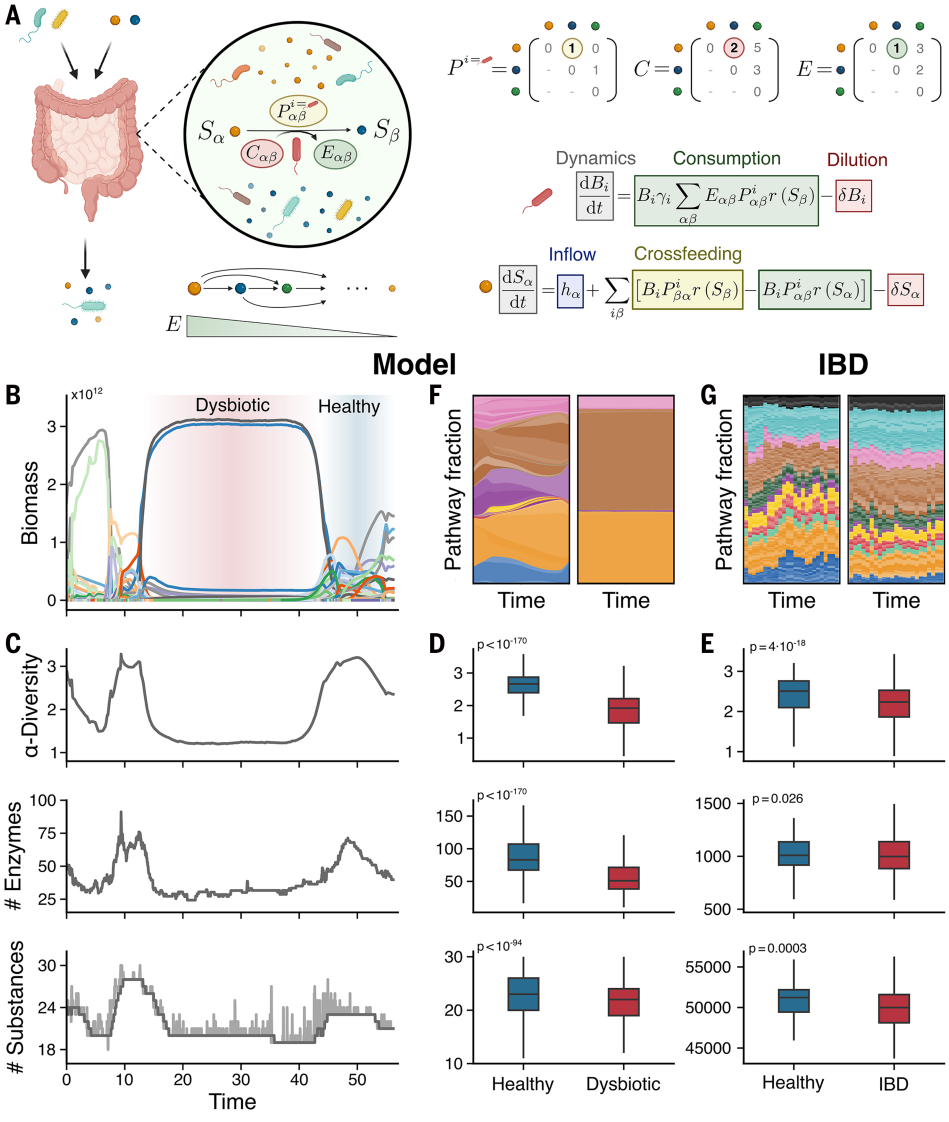


Fig. 2. Gut microbiome model and the spontaneous emergence of alternative stable states. (**A**) Schematic representation of the model. The gut microbiome is conceptualized as living in a chemostat-like environment that includes both bacterial species and substances. Metabolic pathways linking any two substances, S_α and S_β, are characterized by their enzymatic cost, $C_{\alpha\beta}$, and the energy obtained from the conversion, $E_{\alpha\beta}$. Each bacterial species can use specific pathways depending on its pathway matrix, P^i, which determines the metabolic routes available to that species. The model considers only catabolic pathways, where $E_{\alpha\beta} > 0$. (**B**) Biomass (cells) over time for a realization of the model, illustrating the emergence of dysbiotic and healthy states. (**C**) For the model simulation in (B), α diversity (top), number of enzymes (middle), and number of substances (bottom) over time. Dark gray represents rolling average, and light gray represents the actual values. (**D**) Comparison, for the same metrics in (C), of healthy and dysbiotic states in the model. (**E**) Corresponding metrics using existing data for healthy individuals and IBD patients. (**F**) Coarse-grained pathway fraction over time for a healthy (left) and dysbiotic (right) state in the model; colors represent different (coarse-grained) functions. (**G**) Similar plot for healthy versus IBD patients, where colors represent (coarse-grained) functions, and shades represent the various pathways belonging to the given function. All *P* values show two-sided Mann-Whitney *U* tests. Further details are provided in the materials and methods (*45*).

and S7). Beyond this lack of consistency, these metrics merely reflect the shift from one state to the other and do not explain the mechanisms underlying the transitions, nor can they be used as early warning indicators.

Thus, in the search for a more mechanistic characterization of these shifts in the model, we analyzed the pathway structure for healthy and dysbiotic states (fig. S11). We observed that dysbiotic states contained self-sufficient consortia with minimal pathway overlap—i.e., consisting of bacterial lineages that metabolized the most energetic resources through fewer, more direct metabolic routes, thus enabling efficient exploitation of available resources (fig. S11, A and B) and potentially facilitating host exploitation (*61*). Bacteria in healthy states, by contrast, showed greater functional redundancy, with overlapping metabolic pathways. Bacterial lineages in dysbiotic states were typically closer to the theoretical optimum of the enzymatic trade-off function compared with lineages in healthy states (fig. S11C), which explains the increased persistence (fig. S12A) and dominance (fig. S7A) observed in dysbiosis. Examples of bacteria for which an increased presence has been reported to correlate with real diseases include *Bacteroides fragilis* or *Escherichia coli* in IBD and *Fusobacterium nucleatum* and *Solobacterium moorei* in CRC (*30*, *62*).

Moreover, the model offers mechanistic insight into how transitions between states occur. Transitions to dysbiosis are typically triggered by metabolically compatible bacterial consortia that form a tightly connected metabolic loop or short chain, efficiently metabolizing resources with minimal pathway overlap (fig. S13) and allowing them to rapidly dominate the community. By contrast, recovery from a dysbiotic state requires the establishment of bacterial taxa that can disrupt the metabolic loop or chain of the dominant consortium and establish new pathway interdependence among themselves, thus leading to a collective reorganization of the community structure. These dynamics suggest a possible mechanism for the existence of disease-associated taxa observed across experiments (*10*, *21*) and may help explain why FMT or multistrain probiotic interventions can be effective (*14*, *16*).

Community interactions are disrupted in dysbiotic states and disease

Microbial communities are structured not only by environmental conditions but also by ecological interactions, including competition, cooperation, and cross-feeding (*63*, *64*). Experiments with synthetic communities show that competition often dominates, yet cooperative behaviors such as cross-feeding can arise under certain constraints and promote coexistence (*65*–*67*). Natural communities typically show both interaction types, and their dominance in the community dynamics may shift depending on environmental context (*68*, *69*). Beyond pairwise interactions, recent studies have underscored the role of higher-order effects and ecological feedbacks in community dynamics (*70*, *71*). Together, these findings underscore the importance of explicitly modeling ecological interactions to understand the emergent behaviors of microbial ecosystems.

In our model, bacterial interactions emerge from competition for shared resources and cross-feeding. Community dynamics ultimately determine which substances (and, indirectly, which interactions) persist

in the system. Therefore, our model enables a shift of focus from broad measures of species and functional diversity to the emergent, dynamic ecological network of effective interactions that collectively shape the resulting community (*39*, *72*).

There is currently no consensus on whether competition or cooperation dominates gut microbiota dynamics. Although some empirical studies suggest that competitive interactions dominate owing to their role in promoting stability (*18*, *63*, *73*), others emphasize the critical role of cross-feeding interactions in maintaining a healthy microbial community (*74*–*77*). The relative contributions of these interactions to gut microbiota function in health and disease remain unclear.

To quantify contributions of competition and cross-feeding, we used a net interaction metric ρ, defined as the difference between the total cross-feeding and competitive interactions within the community, normalized by the sum of all interaction strengths [see eq. S10 (*45*)]. Because this metric is normalized, the information that it provides remains independent of community size and diversity.

Our analysis reveals that the healthy state in the model is characterized by a negative net interaction, whereas the dysbiotic state is characterized by a positive net interaction (Fig. 3A), which we use to identify the two states in the model (*45*). This behavior is robust across different values of the parameters of the model (fig. S4). Moreover, fig. S12B shows that the duration of a state correlates with the strength of the interactions: More positive interactions are associated with longer-lasting dysbiotic states (similarly with negative interactions and healthy states). A more detailed view of these interactions reveals that dysbiotic states are characterized by a decrease in competition and an increase in cross-feeding (more specifically, commensalism and mutualism) compared with the healthy state (fig. S14).

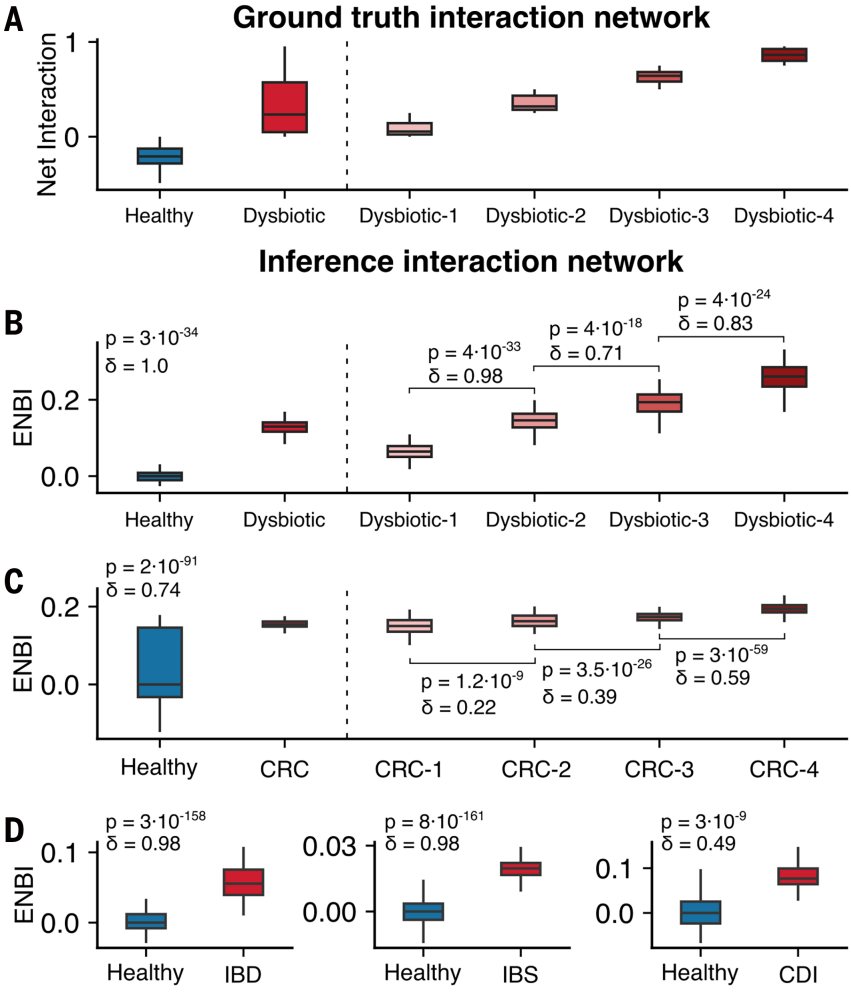


Fig. 3. The ENBI discerns healthy from pathological communities and correlates with disease progression. (**A**) Net interaction, ρ, calculated from the actual interaction network in the model for healthy and dysbiotic states separately (left side of the dashed line) and different consecutive configurations of unhealthy states (right side). For the latter, configurations were grouped and ordered in four sets (i.e., quartiles) from smaller to larger ρ. (**B**) ENBI calculated for model data, showing that this indicator captures the same qualitative behavior as the net interaction. (**C**) ENBI for healthy and CRC patients (left side of dashed line) showing that the latter have a positive ENBI, which suggests a more positively interacting microbial community compared with healthy patients; the ENBI increases with disease progression (right side), supporting our hypothesis of a progressive shift toward network structures with more positive interactions. (**D**) ENBI for IBD, IBS, and CDI also show that diseased states tend to show more positive interactions than their respective controls. See ENBI definition in eq. S12. Effect sizes (Cliff's delta) and *P* values (two-sided Mann-Whitney *U* test) were calculated as described in the materials and methods (*45*).

On the basis of these observations, we hypothesize that in IBD and potentially in other diseases, the gut microbiome shifts toward greater positive interactions compared with the dynamics observed in healthy controls. The net interaction, which focuses on the properties of the underlying interaction network, captures this transition. On the other hand, coarser metrics such as α diversity indices (e.g., the Shannon index) are broadly distributed for any given value of ρ (fig. S15), which may explain why diversity fails to consistently distinguish healthy from diseased states across conditions.

To test our hypothesis, we used an algorithm for empirical coabundance networks to infer effective interaction networks (*45*, *78*) and calculated the associated net interaction metric, $\rho_{inferred}$, for available real-world cross-sectional data and our model-generated data. The latter allowed us to validate the inferred interaction network against the actual one. The inferred net interaction $\rho_{inferred}$ for different simulated communities captures the qualitative behavior of the actual underlying interaction network (Fig. 3, A and B): First, simulated communities from dysbiotic states show a higher $\rho_{inferred}$ compared with those from healthy states; second, communities in dysbiotic states with higher true ρ values show higher values of $\rho_{inferred}$. Figure S16A shows that the inference method remains robust across different subsampling strategies. Building on this validation, we further applied this methodology across different diseases (IBD, CDI, IBS, and CRC). In all cases, there was a significant increase in $\rho_{inferred}$ compared with their respective healthy control groups (fig. S16, B to E). The qualitative results remained consistent across additional network inference methods, including a method designed for cross-sectional data (*79*) (fig. S17) and another specifically developed for longitudinal data (*80*) (fig. S18).

We can use the healthy or control case as a baseline for each dataset and compute the difference between the $\rho_{inferred}$ values of the dysbiotic and control inferred networks, which we called the ecological network balance index (ENBI) (eq. S12). This index is negative when the community exhibits dynamics with more negative interactions than the control (ENBI < 0) and positive when the community shows more positive interactions than the control (ENBI > 0).

Across all diseases analyzed as well as in our model, the dysbiotic state consistently exhibited a significantly positive ENBI value, reflecting a shift toward

more positively interacting communities relative to the controls (Fig. 3, B to D). In the case of CRC, for which long-term data on disease progression are available, we further observed a progressive increase in ENBI as the disease advances, indicating a strengthening of positive interactions in the community over time (fig. S16B and Fig. 3D). Notably, this shift in ENBI is robust across a wide range of methodological and biological contexts. We observed consistent ENBI patterns across different geographic regions and, by extension, dietary backgrounds (fig. S19); across distinct taxonomic profiling pipelines (fig. S20); and across taxonomic resolutions, from the species to family level (fig. S21). Moreover, ENBI values remained similar for two independent healthy cohorts (fig. S22) as well as in a mild dietary intervention study involving a barley-rich diet versus a controlled diet (fig. S23). Together, these results support the robustness of the ENBI and reinforce its potential as a general indicator of disease-associated dysbiosis.

Discussion

Whether substantial changes in the gut microbiome cause or result from the onset of disease remains debated. In this work, we examined the ecological dynamics and community-level effects shaping the gut microbiome in health and disease. We developed the ENBI to quantify the dominant interaction dynamics, which offered insights into the structure of microbial communities and underlying mechanisms involved. Unlike traditional metrics (e.g., the Shannon index), the ENBI reveals a reproducible pattern across diseases and methodologies, highlighting a microbial shift toward more positively interacting communities in dysbiotic states compared with healthy controls.

Specifically, our model indicates that diseased states are associated with a shift in microbial interaction networks, by which overall positive interactions (commensalism and mutualism) increasingly dominate over negative interactions (competition). This relative increase reflects a global shift in the balance of microbial interactions within the community, which does not contradict previous studies describing reductions in specific cross-feeding interactions (*76*, *77*). Furthermore, in line with these previous studies, we also observe an increased metabolic self-sufficiency in dysbiosis: In the model, microbial communities in dysbiotic states are dominated by self-sufficient consortia composed of bacterial lineages that efficiently exploit available resources. These consortia can outcompete other species, potentially depleting key metabolites and altering or even eliminating functions important for host health. This exclusion often, but not always, leads to reduced α diversity, which explains why some diseases exhibit a loss of diversity whereas others do not.

Moreover, for colorectal cancer, for which data on disease progression are available, the ENBI correlates with disease advancement. If this trend holds across other diseases, the finding carries important clinical implications, particularly because our metric can be quantified noninvasively from stool metagenomic data. Therefore, monitoring the ENBI of a patient can help devise early warning indicators for disease onset and monitor progression, which is especially important for conditions such as CRC for which early detection is key to the success of the treatment and improvement of survival rates (*81*). Nonetheless, despite the broad applicability of the ENBI, the index by itself does not identify the specific disease causing dysbiosis. Thus, to maximize its diagnostic potential, the ENBI must be combined with other diagnostic tools for precise disease identification (*81*).

Our model follows a metabolically explicit approach, but, unlike genome-scale metabolic models and kinetic models (*39*), which require detailed genome annotations and well-defined objective functions with extensive parameter fitting, our simpler model focuses on essential ecological and metabolic principles and a biologically plausible parameterization. Our bottom-up approach thus has the potential to unveil the mechanistic underpinnings of gut microbiome emergent properties, such as diversity, multistability, and metabolic organization (*40*). For example, our model provides an explanation for the Anna Karenina principle (AKP)—the observation that the microbiomes of healthy individuals are seemingly similar to each other, but dysbiotic communities are all distinct (fig. S9). Healthy states in our model show similar functional composition, which constrains the communities to exhibit a certain degree of similarity, even though the taxa performing those functions may differ. By contrast, changes in the interaction network structure occurring in dysbiotic states partially or totally disrupt some of those functions in a nondeterministic manner that leads to more variable outcomes, with disease data pointing to potential disease-specific disruptions (*10*). Finally, the subtle behavior for β diversity observed in the model, and the lack of consistency observed in our analysis of β diversity across diseases, agree with the lack of generality of the AKP in empirical data, as proposed in the past (*82*).

Because the model captures ecological interactions in a broad and general way, it can be readily adaptable to other microbiomes, from other tissues (e.g., the vaginal or oral microbiomes) to plant or soil microbiomes. It can also easily expand to incorporate therapeutic interventions, such as probiotics or fecal transplantation. Nonetheless, there are limitations to be considered when using our framework. In the effort to isolate the core ecological mechanisms underlying the emergence of alternative community states, we deliberately excluded several processes that are known to influence gut microbiota composition. These include spatial structure, immune regulation, oxygen gradients, and explicit interactions with the host, currently represented in this work solely through resource input and dilution. Additionally, the model simplifies key aspects of microbial physiology; for example, it represents anabolic reactions indirectly (as opposed to explicitly) and uses a reduced parameterization for energy use and enzymatic costs. Other ecological factors, such as bacteriophage dynamics and chemical antagonism between bacteria (e.g., bacteriocin production), are also omitted. Expanding the model to incorporate a more realistic representation of these aspects is an important avenue for future work. Furthermore, our approach to analyzing empirical data relies on inferred ecological interactions, which intrinsically depend on data quality and quantity and come with inherent assumptions and challenges (e.g., a positive interaction may reflect either mutualism or coblooming owing to environmental change) (*38*, *83*). Other limitations are inherent to the heterogeneity of the methods used across studies. For example, cross-disease comparisons could provide valuable insights into microbiome dynamics for different conditions but are currently hindered by variability in sampling and different taxonomic profiling methods used across studies. Such methodological differences directly affect the inference of the ecological interaction network, which makes quantitative comparisons across diseases difficult. For that reason, we focused on identifying qualitative trends rather than absolute numerical differences. Addressing these challenges will require a multidisciplinary effort to integrate empirical studies and clinical data with theoretical approaches such as ours.

REFERENCES AND NOTES

1. W. M. de Vos, E. A. de Vos, *Nutr. Rev.* **70**, S45–S56 (2012).
2. W. M. de Vos, H. Tilg, M. Van Hul, P. D. Cani, *Gut* **71**, 1020–1032 (2022).
3. J. A. Gilbert *et al.*, *Nature* **535**, 94–103 (2016).
4. K. Hou *et al.*, *Signal Transduct. Target. Ther.* **7**, 135 (2022).
5. Human Microbiome Project Consortium, *Nature* **486**, 207–214 (2012).
6. A. Zhernakova *et al.*, *Science* **352**, 565–569 (2016).
7. G. Vrancken, A. C. Gregory, G. R. B. Huys, K. Faust, J. Raes, *Nat. Rev. Microbiol.* **17**, 754–763 (2019).
8. I. Cho, M. J. Blaser, *Nat. Rev. Genet.* **13**, 260–270 (2012).
9. J. Jovel, L. A. Dieleman, D. Kao, A. L. Mason, E. Wine, in *Metagenomics*, M. Nagarajan, Ed. (Academic Press, 2018), pp. 197–213.
10. Y. Fan, O. Pedersen, *Nat. Rev. Microbiol.* **19**, 55–71 (2021).
11. F. Sommer, J. M. Anderson, R. Bharti, J. Raes, P. Rosenstiel, *Nat. Rev. Microbiol.* **15**, 630–638 (2017).
12. A. M. Valdes, J. Walter, E. Segal, T. D. Spector, *BMJ* **361**, k2179 (2018).
13. C. Duvallet, S. M. Gibbons, T. Gurry, R. A. Irizarry, E. J. Alm, *Nat. Commun.* **8**, 1784 (2017).
14. Y. Wang, S. Zhang, T. J. Borody, F. Zhang, *Chin. Med. J.* **135**, 1927–1939 (2022).
15. Q. Yang *et al.*, *Nutrients* **12**, 381 (2020).
16. O.-A. Petrariu *et al.*, *Front. Microbiol.* **14**, 1296447 (2024).

17. T. S. B. Schmidt, J. Raes, P. Bork, *Cell* **172**, 1198–1215 (2018).
18. M. Kumar, B. Ji, K. Zengler, J. Nielsen, *Nat. Microbiol.* **4**, 1253–1267 (2019).
19. I. Sekirov, S. L. Russell, L. C. M. Antunes, B. B. Finlay, *Physiol. Rev.* **90**, 859–904 (2010).
20. A. B. Shreiner, J. Y. Kao, V. B. Young, *Curr. Opin. Gastroenterol.* **31**, 69–75 (2015).
21. V. K. Gupta *et al.*, *Nat. Commun.* **11**, 4635 (2020).
22. M. Seppi, J. Pasqualini, S. Facchin, E. V. Savarino, S. Suweis, *Biomolecules* **14**, 5 (2023).
23. A. Mosca, M. Leclerc, J. P. Hugot, *Front. Microbiol.* **7**, 455 (2016).
24. M. Kriss, K. Z. Hazleton, N. M. Nusbacher, C. G. Martin, C. A. Lozupone, *Curr. Opin. Microbiol.* **44**, 34–40 (2018).
25. C. Petersen, J. L. Round, *Cell. Microbiol.* **16**, 1024–1033 (2014).
26. Z. S. Ma, L. Li, N. J. Gotelli, *ISME J.* **13**, 1911–1919 (2019).
27. L. Aldars-García, M. Chaparro, J. P. Gisbert, *Microorganisms* **9**, 977 (2021).
28. E. Martinez, B. Taminiau, C. Rodriguez, G. Daube, *Pathogens* **11**, 781 (2022).
29. R. Pittayanon *et al.*, *Gastroenterology* **157**, 97–108 (2019).
30. A. M. Thomas *et al.*, *Nat. Med.* **25**, 667–678 (2019).
31. S. A. Shetty, F. Hugenholtz, L. Lahti, H. Smidt, W. M. de Vos, *FEMS Microbiol. Rev.* **41**, 182–199 (2017).
32. M. Van de Guchte *et al.*, *Microbiome* **8**, 153 (2020).
33. M. Arumugam *et al.*, *Nature* **473**, 174–180 (2011).
34. L. Lahti, J. Salojärvi, A. Salonen, M. Scheffer, W. M. de Vos, *Nat. Commun.* **5**, 4344 (2014).
35. S. Zaoli, J. Grilli, *Sci. Adv.* **7**, eabj2882 (2021).
36. H. Fujita *et al.*, *Microbiome* **11**, 63 (2023).
37. D. Garza *et al.*, bioRxiv 2024.11.28.625814 (2024).
38. K. Faust, J. Raes, *Nat. Rev. Microbiol.* **10**, 538–550 (2012).
39. D. Rios Garza, D. Gonze, H. Zafeiropoulos, B. Liu, K. Faust, *Cell Syst.* **14**, 109–121 (2023).
40. N. I. van den Berg *et al.*, *Nat. Ecol. Evol.* **6**, 855–865 (2022).
41. Y.-Y. Liu, *Cell Syst.* **14**, 135–159 (2023).
42. C. G. Buffie *et al.*, *Nature* **517**, 205–208 (2015).
43. T. S. B. Schmidt *et al.*, *Nat. Med.* **28**, 1902–1912 (2022).
44. G. Aguadé-Gorgorió, J. F. Arnoldi, M. Barbier, S. Kéfi, *Ecol. Lett.* **27**, e14413 (2024).
45. Materials and methods are available as supplementary materials.
46. S. Krishnan, N. Alden, K. Lee, *Curr. Opin. Biotechnol.* **36**, 137–145 (2015).
47. K. Oliphant, E. Allen-Vercoe, *Microbiome* **7**, 91 (2019).
48. K. A. Krautkramer, J. Fan, F. Bäckhed, *Nat. Rev. Microbiol.* **19**, 77–94 (2021).
49. I. Rowland *et al.*, *Eur. J. Nutr.* **57**, 1–24 (2018).
50. Z.-N. Ling *et al.*, *Signal Transduct. Target. Ther.* **8**, 345 (2023).
51. J. Grilli, *Nat. Commun.* **11**, 4743 (2020).
52. P.-Y. Ho, B. H. Good, K. C. Huang, *eLife* **11**, e75168 (2022).
53. S. Louca *et al.*, *Nat. Ecol. Evol.* **2**, 936–943 (2018).
54. R. L. Tatusov, M. Y. Galperin, D. A. Natale, E. V. Koonin, *Nucleic Acids Res.* **28**, 33–36 (2000).
55. H. Ogata *et al.*, *Nucleic Acids Res.* **27**, 29–34 (1999).
56. G. Bunin, *Phys. Rev. E* **95**, 042414 (2017).
57. E. Mallmin, A. Traulsen, S. De Monte, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2312822121 (2024).
58. I. Dalmedigos, G. Bunin, *PLOS Comput. Biol.* **16**, e1008189 (2020).
59. E. Blumenthal, J. W. Rocks, P. Mehta, *Phys. Rev. Lett.* **132**, 127401 (2024).
60. M. Schirmer, A. Garner, H. Vlamakis, R. J. Xavier, *Nat. Rev. Microbiol.* **17**, 497–511 (2019).
61. E. J. Stevens, K. A. Bates, K. C. King, *PLOS Pathog.* **17**, e1009514 (2021).
62. J. Lloyd-Price *et al.*, *Nature* **569**, 655–662 (2019).
63. K. Z. Coyte, J. Schluter, K. R. Foster, *Science* **350**, 663–666 (2015).
64. J. D. Palmer, K. R. Foster, *Science* **376**, 581–582 (2022).
65. K. R. Foster, T. Bell, *Curr. Biol.* **22**, 1845–1850 (2012).
66. O. S. Venturelli *et al.*, *Mol. Syst. Biol.* **14**, e8157 (2018).
67. W. Lopes, D. R. Amor, J. Gore, *Nat. Commun.* **15**, 4709 (2024).
68. D. Machado *et al.*, *Nat. Ecol. Evol.* **5**, 195–203 (2021).
69. J. Camacho-Mateu, A. Lampo, M. Sireci, M. A. Muñoz, J. A. Cuesta, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2309575121 (2024).
70. A. Goyal, L. S. Bittleston, G. E. Leventhal, L. Lu, O. X. Cordero, *eLife* **11**, e74987 (2022).
71. C.-Y. Chang, D. Bajić, J. C. C. Vila, S. Estrela, A. Sanchez, *Science* **381**, 343–348 (2023).
72. K. Z. Coyte, S. Rakoff-Nahoum, *Curr. Biol.* **29**, R538–R544 (2019).
73. P.-Y. Ho, T. H. Nguyen, J. M. Sanchez, B. C. DeFelice, K. C. Huang, *Nat. Microbiol.* **9**, 1036–1048 (2024).
74. E. J. Culp, A. L. Goodman, *Cell Host Microbe* **31**, 485–499 (2023).
75. A. Goyal, T. Wang, V. Dubinkina, S. Maslov, *Nat. Commun.* **12**, 1335 (2021).
76. V. R. Marcelino *et al.*, *Nat. Commun.* **14**, 6546 (2023).
77. A. R. Watson *et al.*, *Genome Biol.* **24**, 78 (2023).
78. J. Tackmann, J. F. Matias Rodrigues, C. von Mering, *Cell Syst.* **9**, 286–296.e8 (2019).
79. C. Li *et al.*, *PLOS Comput. Biol.* **17**, e1009343 (2021).
80. C. K. Fisher, P. Mehta, *PLOS ONE* **9**, e102451 (2014).
81. S. S. Gude *et al.*, *Cureus* **15**, e37622 (2023).
82. S. D. Jurburg, S. A. Blowes, A. Shade, N. Eisenhauer, J. M. Chase, *Microbiome* **12**, 79 (2024).
83. S. Weiss *et al.*, *ISME J.* **10**, 1669–1681 (2016).
84. A. J. Yarur *et al.*, *Gastroenterology* **165**, 963–975.e5 (2023).
85. L. Ning *et al.*, *Nat. Commun.* **14**, 7135 (2023).
86. Human Microbiome Project, The Inflammatory Bowel Disease Multiomics Database; https://ibdmdb.org/.
87. M. Berkell *et al.*, *Nat. Commun.* **12**, 2241 (2021).
88. M. J. T. Crobach *et al.*, *Microorganisms* **8**, 677 (2020).
89. P. Ferretti *et al.*, *eLife* **12**, RP90111 (2023).
90. R. A. T. Mars *et al.*, *Cell* **182**, 1460–1473.e17 (2020).
91. G.-H. Kim, K. Lee, J. O. Shim, *Microbiol. Spectr.* **11**, e0212522 (2023).
92. Z. A. Barandouzi *et al.*, *Sci. Rep.* **12**, 1648 (2022).
93. G. Zeller *et al.*, *Mol. Syst. Biol.* **10**, 766 (2014).
94. S. Yachida *et al.*, *Nat. Med.* **25**, 968–976 (2019).
95. E. Vogtmann *et al.*, *PLOS ONE* **11**, e0155362 (2016).
96. R. Corral Lopez, Code for article: Imbalance in gut microbial interactions as a marker of health and disease, Zenodo (2025).

ACKNOWLEDGMENTS

The authors thank J. Grilli, M. Sireci, and the members of the Bonachela laboratory for helpful discussions. **Funding:** We acknowledge the Spanish Ministry of Research and Innovation and Agencia Estatal de Investigación (AEI), MICIN/AEI/10.13039/501100011033, for financial support through project no. PID2023-149174NB-I00, funded also by the European Regional Development Fund (ERDF) of the European Union, and project no. PID2020-113681GB-I00 funded by MICIN/AEI/10.13039/501100011033. This work was also supported by a Simons Early Career Investigator in Marine Microbial Ecology and Evolution Award (award no. 826106), National Science Foundation grant DMS-2052616 to J.A.B., and National Institutes of Health grant U01 AI22285 to M.J.B. R.C.L. acknowledges funding from the Spanish Ministry and AEI (grant no. FPU19/03887) and Fulbright Spain. **Author contributions:** J.A.B. and M.A.M. designed and supervised the research. R.C.L., J.A.B., and M.A.M. developed the model with input from S.A.L., M.J.B., M.M., and M.G.D.-B. R.C.L. collected the data and performed statistical and modeling analyses. All authors wrote and reviewed the manuscript. **Competing interests:** The authors declare no competing interests. **Data, code, and materials availability:** No empirical data were produced in this study. All the publicly available datasets are cited in the "Datasets" section of the materials and methods. Our code is available on Zenodo (*96*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ady1729
Materials and Methods; Figs. S1 to S24; Table S1; References (*97*–*112*);
MDAR Reproducibility Checklist

Submitted 11 April 2025; resubmitted 30 July 2025; accepted 20 November 2025

10.1126/science.ady1729

Science
AAAS

Supplementary Materials for

Imbalance in gut microbial interactions as a marker of health and disease

Roberto Corral López *et al*.

Corresponding authors: Juan A. Bonachela, juan.bonachela@rutgers.edu; Miguel A. Muñoz, mamunoz@onsager.ugr.es

Science **391**, 890 (2026)
DOI: 10.1126/science.ady1729

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Materials and Methods

Measures

- α**-Diversity:** measured using the Shannon Index, defined as

$$\text{Shannon Index} = -\sum_i b_i \ln b_i \tag{S1}$$

where b_i represents the relative abundance (biomass) of type i in the model, or the relative abundance (proportion of counts) of species i in the metagenomic data.

- **Richness:** measured using the number of distinct bacterial types in the model, or taxonomic entities in the real-world data.

- **Dominance:** measured as the relative abundance of the most abundant bacterial type (model) or taxonomic entity (real-world data).

- **Simpson Index:** defined as

$$\text{Simpson} = 1 - \sum_i b_i^2 \tag{S2}$$

where b_i represents the relative abundance of bacterial type i (model) or species i (metagenomic data).

- **Pielou's evenness:** calculated as the Shannon index divided by the logarithm of richness:

$$\text{Pielou} = \frac{\text{Shannon Index}}{\ln(\text{Richness})} \tag{S3}$$

- **β-diversity:** measured using the Bray-Curtis (BC) dissimilarity index between two samples, a and d, defined as:

$$BC_{a,d} = 1 - \frac{2\sum_i \min\left(b_i^a, b_i^d\right)}{\sum_i \left(b_i^a + b_i^d\right)} \quad \text{(S4)}$$

where b_i^a and b_i^d represent, for samples a and d, the relative abundance of type i (model) or species i (metagenomic data). A value of $BC_{a,d} = 0$ indicates identical community compositions, while $BC_{a,d} = 1$ indicates completely distinct communities.

- **Number of Enzymes:** In real data, we calculated the number of enzymes as the sum of enzymes with non-zero abundance in the sample. In the model, because enzymes are present through the enzymatic cost of a pathway, we calculated a proxy for the number of enzymes, the total enzymatic cost in the sample:

$$\text{Total Enzymatic Cost} = \sum_{i\alpha\beta} P_{\alpha\beta}^i \, C_{\alpha\beta} \quad \text{(S5)}$$

where $P_{\alpha\beta}^i$ represents the pathway matrix of bacterial type i and $C_{\alpha\beta}$ the associated cost stemming from enzyme production.

- **Number of Metabolites:** Calculated as the sum of non-zero metabolites

$$\text{Number of Metabolites} = \sum_{\alpha} \theta\left(S_\alpha\right) \quad \text{(S6)}$$

where $\theta(x)$ is the Heaviside function (i.e. $\Theta(x) = 1$ if $x > 0$, and $\Theta(x) = 0$ otherwise) and S_α is the density of metabolite α, both in the model and metabolomic data.

- **Pathway fraction:** In real data, the pathway fraction is computed as the relative abundance of each metabolic pathway. In the model, we define it as:

$$\text{Pathway Fraction}_{\alpha\beta} = \frac{\sum_i P^i_{\alpha\beta} C_{\alpha\beta} B_i}{\sum_{i,\alpha,\beta} P^i_{\alpha\beta} C_{\alpha\beta} B_i} \tag{S7}$$

This formulation incorporates both the enzymatic cost and the biomass of the species expressing each pathway, providing a proxy for the functional signal that would be obtained from pathway abundance profiles in metagenomic data.

- **Cross-feeding:** calculated as

$$c^+_{ij} = \sum_{\alpha\beta\delta} P^i_{\alpha\beta} P^j_{\beta\delta} r_i(S_\alpha) r_j(S_\beta) B_i B_j \tag{S8}$$

which captures all the cross-feeding interactions from type i to type j, mediated by all possible metabolic pathways. Specifically, c^+_{ij} considers all instances where species i consumes substance S_α to generate substance S_β, which is then consumed by species j to produce any byproduct S_δ. Each interaction is weighted by the corresponding species abundances (B_i and B_j) and consumption factor r_i, the latter given by the Monod function $r(S) = r^{\max} \frac{S}{K+S}$. The matrix composed of all c^+_{ij} is always positive and non-symmetric.

- **Competition:** calculated as

$$c^-_{ij} = \sum_{\alpha\beta\delta} P^i_{\alpha\beta} P^j_{\alpha\delta} r_i(S_\alpha) r_j(S_\alpha) B_i B_j \tag{S9}$$

which captures the competition between types i and j arising from their shared preference for the same substance S_α. Each competitive interaction is weighted by the corresponding species abundances and consumption functions. The matrix composed using all c_{ij}^- is always positive and symmetric.

- **Net interaction:** defined as

$$\rho = \frac{\sum_{i \neq j} \left(c_{ij}^+ - c_{ij}^- \right)}{\sum_{i \neq j} \left(c_{ij}^+ + c_{ij}^- \right)} \tag{S10}$$

where the terms c_{ij}^+ and c_{ij}^- represent, respectively, the elements of the cross-feeding and competition matrices above.

- **Inferred net interaction:** defined as

$$\rho_{\text{inferred}} = \frac{\sum_{i \neq j} \left(\omega_{ij}^+ - \omega_{ij}^- \right)}{\sum_{i \neq j} \left(\omega_{ij}^+ + \omega_{ij}^- \right)} \tag{S11}$$

where ω_{ij} represent the positive (ω_{ij}^+) or negative (ω_{ij}^-) weights between types (in the model) or species (metagenomic data) i and j in the inferred interaction network. See section "Inferring the interaction network" below for details.

- **Ecological Network Balance Index (ENBI)**: caculated as

$$\text{ENBI} = \rho_{\text{inferred}}^{\text{D}} - \text{Median}(\rho_{\text{inferred}}^{\text{H}}) \tag{S12}$$

where ρ_{inferred} are the unhealthy (dysbiotic, $\rho_{\text{inferred}}^{\text{U}}$) and healthy (healthy, $\rho_{\text{inferred}}^{\text{H}}$) inferred net interactions for the different cases analyzed (diseases and model).

- **p-value:** We assessed statistical significance using the Mann-Whitney U test (two-sided).

- **Effect size:** We quantified the effect size of pairwise comparisons using Cliff's delta, a non-parametric measure that captures the degree of overlap between two distributions. It is defined as:

$$\delta = \frac{\text{Number of times}(x_i > y_j) - \text{Number of times}(x_i < y_j)}{mn} \quad \text{(S13)}$$

where $x_i \in X$ and $y_j \in Y$ are observations from two groups of sizes m and n, respectively. The numerator counts the number of times a value from group X is greater than or less than a value from group Y. A value of $\delta = \pm 1$ indicates complete separation between the distributions (i.e., all $x_i > y_j$ or all $x_i < y_j$), while $\delta = 0$ corresponds to complete overlap.

Effect sizes can be interpreted as follows: (i) *negligible* for $|\delta| < 0.15$; (ii) *small* for $0.15 \leq |\delta| < 0.33$; (iii) *medium* for $0.33 \leq |\delta| < 0.47$; and (iv) *large* for $|\delta| \geq 0.47$. See (97) for further details.

Definition of healthy and dysbiotic states

To classify states beyond a simple visual inspection of the temporal dynamics (e.g. Fig. 2B), we define the **healthy state** as the configuration of the system in which the net interaction among present bacteria is negative ($\rho < 0$). Conversely, the **dysbiotic state** is characterized by a positive net interaction among bacteria in the system ($\rho > 0$).

We examined multiple alternative definitions based on other biomarkers, including combinations of the Shannon index, richness, and the number of substances or metabolites, among others.

However, the criterion above proved to be the definition most in agreement with the classification of states obtained through visual inspection.

Inferring the interaction network

Inferring interaction networks from abundance correlations is a long-standing challenge, with well-known caveats discussed extensively in the literature (38, 80, 98, 99). To address these limitations and improve robustness, we employed three complementary approaches that move beyond purely correlational metrics: two suitable for cross-sectional datasets (i.e., data from different subjects) and another one specifically devised for longitudinal datasets (i.e., data collected over multiple time points from the same subject). All three algorithms operate on relative abundance data.

- **Method 1.** ***FlashWeave***: This algorithm infers interaction networks from cross-sectional data using co-abundance networks. These networks are built on the principle that bacterial species with positively correlated abundances likely exhibit positive interactions, while negatively correlated abundances suggest negative interactions. Importantly, the algorithm goes beyond simple correlation by incorporating ways to filter out indirect interactions mediated by other species, and distill direct ones. See (78) for further details.

  As output, this algorithm generates a single undirected interaction network for a given set of samples. To obtain statistical insights, we used bootstrapping (or subsampling) on each dataset. The latter involves generating multiple subsamples from the full dataset and inferring the interaction network for each subsample. For instance, Fig. S16 shows the inferred net interaction, ρ_{inferred}, for 500 bootstrapped subsamples across all

datasets, including both model outputs and various disease datasets, using fixed fractions of the total sample set. Additionally, Fig. 3 displays the results for a specific subset obtained with a fraction of 0.8. Across all cases, we consistently observed that the net interaction of the inferred network is higher in diseased (dysbiotic) states compared to control (healthy) states for both real datasets and the model, respectively.

For the datasets corresponding to different diseases (IBD, IBS, CDI, CRC), we used the *sensitive* and *non-heterogeneous* options of FlashWeave in conjunction with its default parameters (78). For the data generated with the model, where the number of samples exceeds one thousand, we applied the *heterogeneous* option, recommended for large datasets in the original publication.

- **Method 2. *BEEM-Static***: This algorithm, which we used as an additional validation of the results obtained with method 1, infers ecological interactions from cross-sectional data by estimating the parameters of a generalized Lotka-Volterra (gLV) model. *BEEM-Static* explicitly models species interactions by assuming that microbiome communities are at or near equilibrium. Specifically, the method employs an expectation-maximization algorithm to iteratively estimate both species' biomasses and interaction strengths, filtering out samples that violate the equilibrium assumptions. Notably, this approach enables the inference of directed interaction networks and has been shown to outperform standard correlation-based methods in capturing both the presence and directionality of microbial interactions.

The algorithm includes various tweaking parameters, which are detailed in the original publication (79). In our implementation, we adhered to the default settings with two exceptions: we increased the maximum number of iterations within the model to 50 and applied preprocessing steps to our datasets. Specifically, we used from the datasets only bacterial taxa that were present in at least 30% of the samples and had a mean relative abundance greater than 10^{-4}, as the algorithm did not seem to work otherwise. As with Method 1, we used bootstrapping to assess statistical significance. Fig.S17 presents the results of our subsampling approach. Consistent with the findings from Method 1, the figure shows that the net interaction of the inferred network is systematically higher in diseased (dysbiotic) states compared to control (healthy) states in both real datasets and model-generated data.

- **Method 3.** ***LIMITS***: This algorithm leverages a discrete-time Lotka-Volterra model to infer undirected interaction networks from longitudinal data. The algorithm includes three key tweaking parameters: the number of species considered for network inference and two parameters specific to the algorithm (the error threshold and the bagging—or bootstrap aggregation—determining the number of internal iterations performed). While the inference process is relatively straightforward, we refer the reader to the original publication for a detailed explanation (80).

  For each subject we selected, among subjects with at least 20 temporal samples, a predefined number of species and infer the interaction network across various error thresholds from 1 to 5 (ranges employed in the original publication). We used the algorithm over a large number of iterations (10^6) (real data) and $5 \cdot 10^4$ (model data) to

ensure robustness, although we noticed that the iteration count did not significantly affect the results. The final network was constructed by first merging the outputs of all iterations at each error threshold and then averaging across all error thresholds. In our implementation, we extended the original algorithm by computing results using both the median and the mean or average. As shown in Fig. S18, the overall patterns remain consistent, with the mean yielding slightly more stable estimates and reduced variability across replicates.

Model building

We model the gut microbiome as living in a chemostat-like environment, a common simplifying assumption (in reality, food intake and substance excretion occurs in a discrete, rather than continuous, manner). We model the ecological dynamics of the system using consumer-resource equations, also considering metabolic preferences. Specifically, we consider a number $N(t)$ of bacterial lineages (e.g., species, strains, or another suitable taxonomical group) with abundances $B_i(t)$, where i identifies the bacterial lineage. Additionally, we include a maximum number of possible resources (or substances, as we use indistinguishably the two terms here) R, with concentrations S_α, where α indicates the type of substance. Substances are introduced into the system at an inflow rate h_α and may also be generated as byproducts of bacterial metabolism. New bacterial lineages are introduced at a certain rate U with a small initial relative abundance of 10^{-4}, as long as their invasion fitness —initial per-capita growth rate of the invader in the environment created by the existing bacterial lineages— is positive. If the invasion fitness is negative and the initial biomass of the invading bacteria is sufficiently small, the new bacterial

lineage is declared extinct in the next temporal step without affecting the system. Therefore, only viable invader lineages contribute to the dynamics of the system.

We explicitly incorporate metabolic pathways into the model by representing all possible routes between any two distinct substances. To describe the metabolic pathway that converts a substance S_α into another substance S_β, we introduce two key quantities:

- The energy $E_{\alpha\beta} = E_\alpha - E_\beta$, or benefit, derived from the route, where E_α and E_β are the energies associated with substances α and β, respectively. This $E_{\alpha\beta}$ reflects the free energy released during the reaction, energy that we assume is fully available for growth. Nonetheless, this assumption could be relaxed by incorporating a conversion constant to account for partial energy utilization.

- The enzymatic cost, $C_{\alpha\beta}$, required to complete the pathway, which accounts for the number and complexity of the enzymes involved.

For simplicity, we explicitly focus on catabolic reactions, i.e. those in which the free energy of the reaction is positive, and account for anabolic processes indirectly only (see below). In particular, we assume that all metabolic routes where $E_{\alpha\beta} > 0$ (i.e. catabolism) are possible, and that the energy obtained and the enzymatic cost needed for the route are the same (see below for details on how this is implemented in the model via the trade-offs). Nonetheless, the framework is versatile and can accommodate more complex and realistic metabolic structures through parameter adjustments.

Each bacterial species is characterized by its own specific metabolic pathways, represented by a binary matrix P^i. In this matrix, $P^i_{\alpha\beta} = 1$ indicates that bacterium i can metabolize S_α into S_β, while $P^i_{\alpha\beta} = 0$ indicates it cannot. These pathways are randomly assigned when the bacterium is introduced into the system and remain fixed, therefore assuming no evolutionary adaptation occurs. This immigration of bacteria with random pathways constitutes the only source of stochasticity in the system.

To assign the pathways, a random number of pathways ranging between 1 and a specified maximum is first drawn for each bacterial lineage. The maximum is chosen to be relatively small to reduce computational overhead, but sufficiently large to not limit the introduction of viable bacteria into the system. We conservatively set this maximum to 10 based on the observation that, across all trade-off parameterizations considered, bacterial lineages with more than 8 pathways were unable to successfully invade and persist in the system. Then, the specific pathways are selected randomly from all possible $S_\alpha \rightarrow S_\beta$ combinations, adhering to the catabolic constraint $E_{\alpha\beta} > 0$. This approach ensures a diverse yet computationally manageable set of pathways for each bacterium, reflecting ecological variability while maintaining system feasibility.

Model equations

The equations representing the model are as follows:

$$\frac{dB_i}{dt} = B_i \left(\gamma_i \sum_{\alpha\beta}^{R} r_{i,\alpha}^{\max} \frac{S_\alpha}{K_{i,\alpha} + S_\alpha} P^i_{\alpha\beta} E_{\alpha\beta} - \delta \right) \tag{S14}$$

$$\frac{dS_\alpha}{dt} = h_\alpha + \sum_i^N \sum_\beta^R B_i \left(P^i_{\beta\alpha} r^{\max}_{i,\beta} \frac{S_\beta}{K_{i,\beta} + S_\beta} - P^i_{\alpha\beta} r^{\max}_{i,\alpha} \frac{S_\alpha}{K_{i,\alpha} + S_\alpha} \right) - \delta S_\alpha \qquad \text{(S15)}$$

where γ_i is the yield or conversion constant from energy to biomass for bacterium i; $r^{\max}_{i,\alpha}$ and $K_{i,\alpha}$ are, respectively, the uptake rate and half-saturation constant for bacterium i feeding on substance α; δ is the outflow rate; and h_α is the inflow rate for substance α. The values for all parameters are fully described in Table S1. We assume growth parameters, $r^{\max}$ and K, to be the same across all bacterial types and substances, and that only the most energetic substance has a non-zero inflow rate ($h_0 \equiv h$; $h_\alpha = 0$ for $\alpha > 0$). For convenience, we express the (unique) energy associated with a substance α with different units, using the conversion $e_\alpha = E_\alpha \frac{M_\alpha}{V_{\text{colon}}}$; we further define such energy as a fraction of a maximum energy, i.e. $e_\alpha = e_{\max}(1 - \alpha/R)$.

Finally, to ensure a balanced and realistic pathway structure within the system, a trade-off is introduced, imposing disadvantages on bacteria with numerous and costly pathways. This trade-off prevents unrealistic scenarios in which a single bacterial type that shows all pathways ultimately dominates the community. The trade-off is implemented as a growth penalty proportional to the number and cost of pathways utilized by the bacteria:

$$\gamma_i \propto \gamma \cdot g\left(\sum_{\alpha,\beta} C_{\alpha\beta} \, P^i_{\alpha\beta} \right) \qquad \text{(S16)}$$

where g is a decreasing function. For the latter, we have explored the following two families of functions:

$$g(x) = \frac{b}{1 + c \cdot x^\nu} \qquad \text{(S17)}$$

$$g(x) = be^{-c \cdot x} \tag{S18}$$

where b, c and ν are three positive constants that help us vary the different trade-offs inside the same family, and verify the robustness of the main results.

Simulation algorithm

We start our simulations with only one substance (the most energetic one, which is introduced in the system at a non-vanishing rate) and only one bacterial lineage able to metabolize this substance (and thus transform it into a random, different substance). From those initial conditions, the simulations proceed according to the following algorithm:

0. Run the ecological dynamics, given by Eqs. S14-S15 until the next invasion step.

1. At each step, eliminate species that have fallen below the extinction threshold (set to be a fraction 10^{-4} of the total biomass, that is, an abundance $10^{-4} \sum_i B_i$).

2. Introduce an invader, a bacterial lineage whose pathways are randomly selected (see above).

3. Repeat until a selected ending time.

Influence of the trade-off

As shown in Fig. S3, the qualitative results of our model remain robust across different trade-off functional forms and parameterizations. Nonetheless, as expected given the nature of the trade-

off (see above), certain parameterizations can lead to ill-defined systems in which either no bacteria survive or only a single species dominates.

The trade-off primarily determines the total enzymatic cost allocated by each bacterial lineage and how this cost is distributed—either across multiple low-cost pathways or concentrated in fewer high-cost ones. Fig. S3F shows the distribution of total enzymatic cost per lineage across different trade-offs. We can approximate the peak of this distribution and gain insight into its width by considering the maximization of the per capita growth rate. First, we incorporate the trade-off function γ_i in the definition of the growth rate (Eqs. S14 and S16):

$$\text{Growth rate} \propto g\left(\sum_{\alpha,\beta} C_{\alpha\beta}\, P_{\alpha\beta}^{i}\right)\sum_{\alpha\beta}^{R} \frac{S_\alpha}{K_{i,\alpha}+S_\alpha}\, P_{\alpha\beta}^{i} e_{\alpha\beta} \tag{S19}$$

where $e_{\alpha\beta} = E_{\alpha\beta}\frac{M_\alpha}{V_{\text{colon}}}$ represents the energy of path $S_\alpha \to S_\beta$ in different units (see above). Maximizing this equation exactly is not possible, because it depends on all substances in the system, which are in turn determined by the pathways of the existing bacterial community. However, assuming that each substance reaches approximately similar concentrations, we can simplify the expression to:

$$\text{Growth rate} \propto g\left(C_{\text{T}}^{i}\right) e_{\text{T}}^{i} \tag{S20}$$

where $C_{\text{T}}^{i} = \sum_{\alpha,\beta} C_{\alpha\beta}\, P_{\alpha\beta}^{i}$ represents the total enzymatic cost per bacterial type, and $e_{\text{T}}^{i} = \sum_{\alpha\beta}^{R} P_{\alpha\beta}^{i}\, e_{\alpha\beta}$ denotes the total energy a bacterial type i extracts from its pathways (up to a proportionality constant to maintain units). Given our assumption that the energy and enzymatic

cost matrices are identical (or at least proportional), the optimal total enzymatic cost per type is given by the value that maximizes:

$$\text{Growth rate} \propto g\left(C_{\mathrm{T}}^{i}\right)C_{\mathrm{T}}^{i} \tag{S21}$$

Fig. S3F shows, for different versions of the trade-off, the right-hand side of this equation alongside the distribution of total enzymatic cost per type. As shown in the figure, the peak of this right-hand side coincides with the peak of the distribution, whose width results from the fact that pathway configurations are assigned randomly and suboptimal pathways persist in the system for some time. This width, however, is also influenced by the trade-off as, when the right-hand side of Eq. S21 is flatter around its maximum, the resulting distribution is broader as suboptimal pathway configurations are similarly efficient to the optimal one, thus allowing for greater variability in enzymatic investment strategies.

In other words, if the maximum of Eq. S21 were significantly higher than the maximum energy obtainable per pathway, only bacteria with the most energetically efficient pathway would survive, leading to an ill-defined system. To obtain non-trivial communities, the trade-off function and parameters must ensure that the peak of Eq. S21 remains below the most energetic pathway. Moreover, the closer this maximum is to the mode of the distribution, the richer the microbial community will be, as a greater number of functionally distinct yet competitive enzymatic investment strategies can emerge.

An illustrative example of unrealistic functional form for the trade-off is the following (linear) function:

$$g\left(C_{\mathrm{T}}^{i}\right)=\frac{1}{C_{\mathrm{T}}^{i}} \tag{S22}$$

This function leads to a growth rate (Eq. S21 with no internal maximum, meaning that all types with different pathway configurations are effectively neutral with respect to the cost–growth trade-off. As a consequence, no configuration is favored, and the system lacks a selective pressure toward particular strategies. Fig. S24 shows three realizations for a system with this trade-off. We observe that, due to the lack of selection pressure, once the first bacterial types colonize the system and occupy the available niches, the community composition becomes effectively static. Given that real gut microbiomes exhibit temporal variability (100), this trade-off is biologically implausible.

Model simulation details

We integrated numerically the model equations (S14-S15) by using a multi-step variable-order integration scheme available in Python ($solve\_ivp$), in which we adapted the integration time steps to prevent negative abundances. The parameters used for all simulations (unless otherwise specified) can be found in Table S1. The enzymatic cost matrix, $C_{\alpha\beta}$, is assumed to be identical to the energy matrix, $E_{\alpha\beta}$, as differences are inherently accounted for through the non-linear trade-off function (see Eqs. S17 and S18). Unless otherwise specified, the simulations presented in the main text used a trade-off function defined by Eq. S17 with parameters $b = 8$, $c = 6$, and $\nu = 1.1$.

Functional coarse-graining

In the model, coarse-graining is implemented by clustering substances with similar energy levels into functional groups (see Fig. S2 for illustration). Let us consider the set of all substances S_α which, as explained in the Model Equations section, is ordered by energy level. For fine coarse-graining, substances are grouped in sets of four, meaning that S_1-S_4 are treated as a single functional entity, S_5-S_9 as another, and so on. An exception is made for S_0, which is always considered separately since it is the only substance directly introduced into the system. Once substances are grouped, we consider all possible metabolic routes between these grouped entities, resulting in $n_{\mathrm{fg}} = 43$ functional groups. For broad coarse-graining, a larger grouping size of 10 is used, leading to a more aggregated representation of metabolic pathways. In this case, the number of functional groups is reduced to $n_{\mathrm{fg}} = 9$.

In the IBD dataset, we annotated and grouped metabolic pathways according to the BioCyc ontology classification (101). We assigned a different color to each functional group, with individual pathways represented by different shades within their respective groups. In Fig. 2G, these groups are ordered from bottom to top as follows: Carbohydrate Degradation (blue), Amino Acid Biosynthesis (orange), Nucleoside and Nucleotide Degradation (light green), Carbohydrate Biosynthesis (red), All Other Biosynthetic Pathways (yellow), All Other Pathways (purple), Generation of Precursor Metabolites and Energy (dark green), Cofactor, Carrier, and Vitamin Biosynthesis (brown), Fatty Acid and Lipid Biosynthesis (pink), Nucleoside and Nucleotide Biosynthesis (light blue), Cell Structure Biosynthesis (black).

Datasets

For the analysis of diseases, we used the following datasets:

- **IBD:** Metagenomic, enzymatic and metabolomic data were obtained from the Inflammatory Bowel Disease Multi'omics Database (https://ibdmdb.org/).
- **IBS:** Metagenomic and enzymatic data were obtained from (90). Metabolomic data obtained from (102).
- **CDI:** Metagenomic data obtained from (89).
- **CRC:** Metagenomic, enzymatic, and metabolomic data were obtained from (94). In Fig. 3C CRC-1,2,3,4 correspond to MP (multiple polypoid adenomas with low-grade dysplasia), Stage-0 CRC, Stage-I,II CRC, and Stage-III,IV CRC as obtained from (94).
- **F4 and M3 gut data:** Metagenomic data corresponding to (103) was obtained from https://github.com/twbattaglia/MicrobeDS.
- **Two independent healthy cohorts:** 16S rRNA sequencing data obtained from (104).
- **Healthy data with different diets** 16S rRNA sequencing data obtained from (105).
- **Model data:** Temporal analyses and macroecological pattern assessments were based on weekly sampled data from our simulations. For analyses comparing healthy and dysbiotic states, we used 1000 samples per state (or the maximum number of samples available) for each realization, with approximately 100 independent realizations included in the analysis.

All datasets were obtained in pre-processed form, as curated by their original authors. We refer the reader to the corresponding publications for detailed descriptions of pre-processing steps. For our analyses, we only normalized count data to obtain relative abundances for each sample.

Table S1: Experimental values of the model parameters

Model parameter values

Symbol	Definition	Value	References
γ	Energy to biomass conversion	1.05×10^{13} cell/mol ATP	(106)
δ	Outflow rate	$\frac{1}{24}$ h^{-1}	(107)
h	Inflow rate of most energetic substance	2 g / (L h)	(59, 108)
K	Half-saturation constant	10^{-4} g/L	(109)
r^{max}	Maximum uptake rate	10^{-10} g/(L cell h)	(109)
U	Invasion rate	$\frac{2}{3}$ invasions h^{-1}	-
e_{max}	Maximum energy per resource consumption	3 mol ATP/mol of S	-
M	Molar mass of substances	100 g/mol	-
V_{colon}	Volume of the colon	0.41 L	(110)
R	Total number of substances	30	-

Supplementary Figures

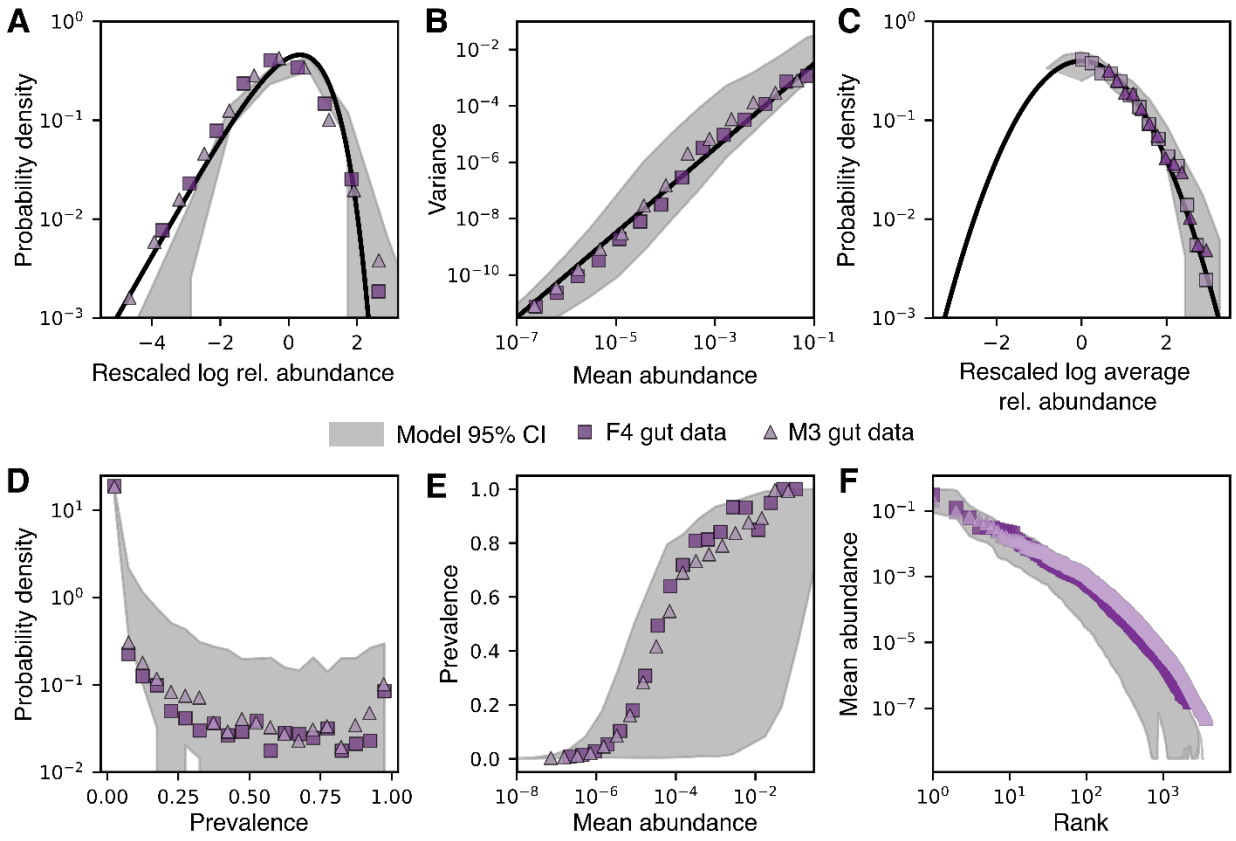


Figure S1: ***Model reproduces well-established macroecological patterns of the gut microbiome without any parameter fitting.*** *(**A-C**) Macroecological patterns described in (51). (**A**) The Abundance Fluctuation Distribution (AFD), which characterizes the distribution of a species' abundances across samples. The solid black line represents the expected distribution obtained by rescaling the log of relative abundance (subtracting the mean and dividing by the standard deviation) when the relative abundance follows a gamma distribution. (**B**) The mean and variance of species abundance distributions follow Taylor's Law, with variance scaling as a power of the mean, specifically with an exponent of 1.5 (black line). (**C**) The Mean Abundance Distribution (MAD), defined as the distribution of species' mean abundance across samples, follows a lognormal distribution (black line). (**D-F**) Macroecological patterns described in (52). (**D**) Distribution of species prevalence, defined as the fraction of samples in which a species is present. (**E**) Relationship between species prevalence and mean abundance. (**F**) Rank distribution of mean abundances.*

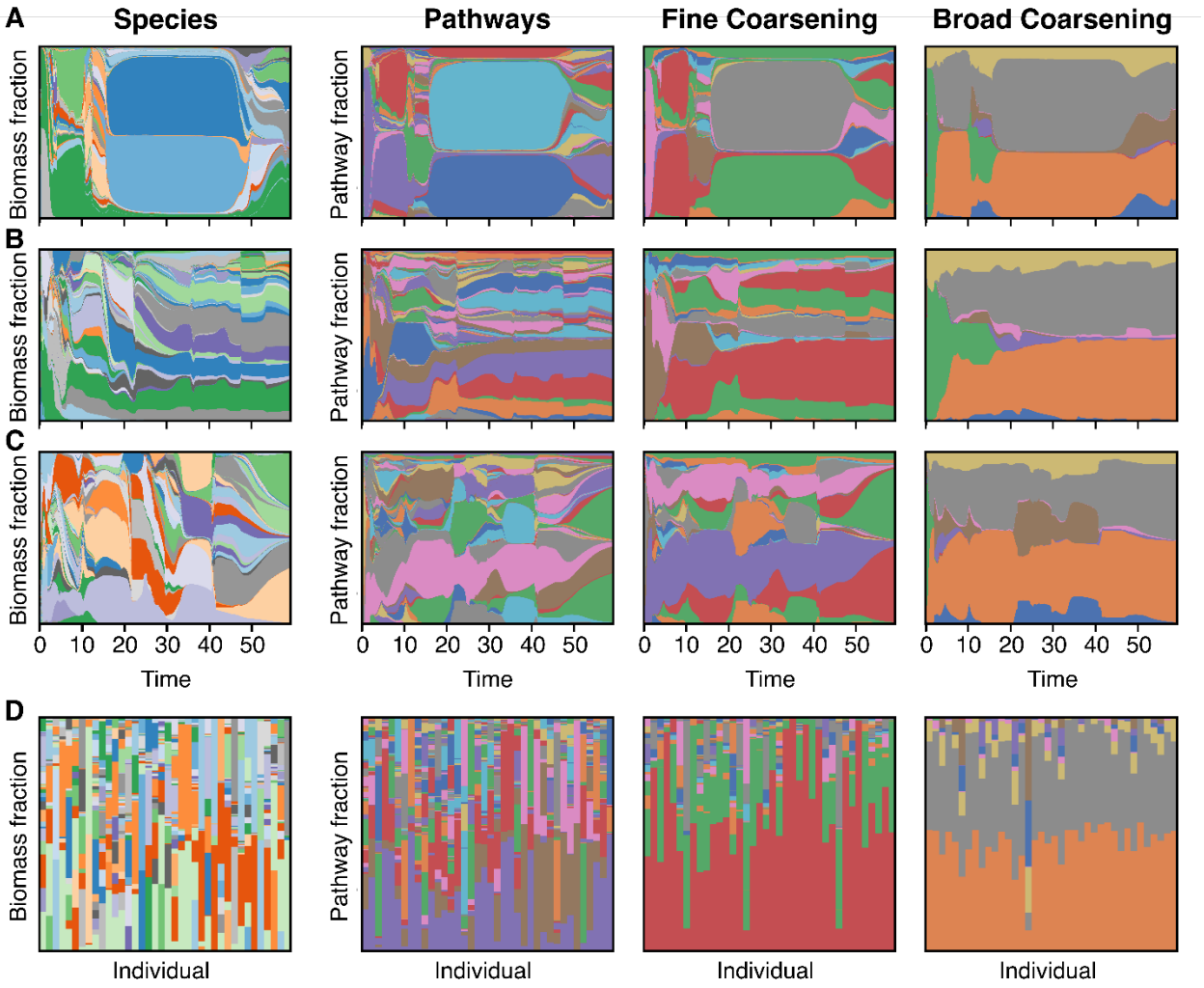


Figure S2: ***Model data reproduces functional redundancy.*** *(**A-C**) Time-series (in years) of relative species abundance (left), enzymatic cost fraction for all pathways (number of resulting functional groups, $n_{fg} = 435$; middle left), for a fine coarsening ($n_{fg} = 43$; middle right), and for a broad coarsening ($n_{fg} = 9$; right) across three different realizations. (**D**) Transversal (across individuals) evaluation of the same metrics: biomass fraction at the species level and enzymatic cost fraction under different coarsening schemes. Coarse-grained pathways exhibit functional redundancy, meaning that taxonomically distinct species share similar functional profiles (53).*

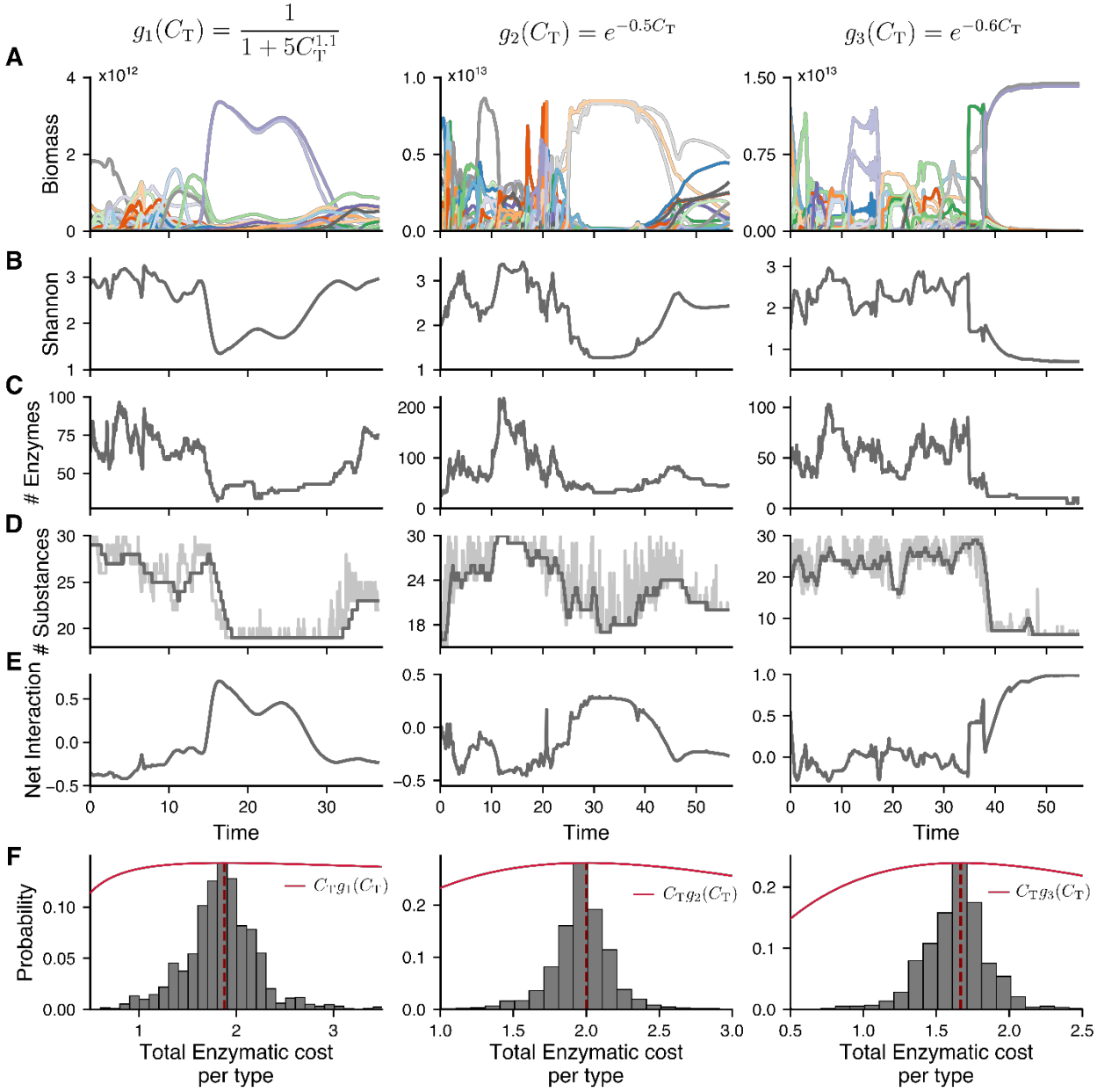


Figure S3: ***Alternative stable states are robust against trade-off changes.*** *For three different trade-off functions, g_1 (left), g_2 (middle), and g_3 (right), the following are shown: (**A-E**) Time-series (in years) of bacterial abundances (**A**), Shannon index (**B**), enzymatic cost (**C**), number of substances (**D**), and net interaction (**E**). (**F**) The total enzymatic cost per lineage (C_T) distribution shows a peak at the maximum of $C_T g_k(C_T)$, suggesting an optimal enzymatic investment strategy. The consistency of these patterns across different trade-off formulations and parameter choices indicates the robustness of the alternative stable states.*

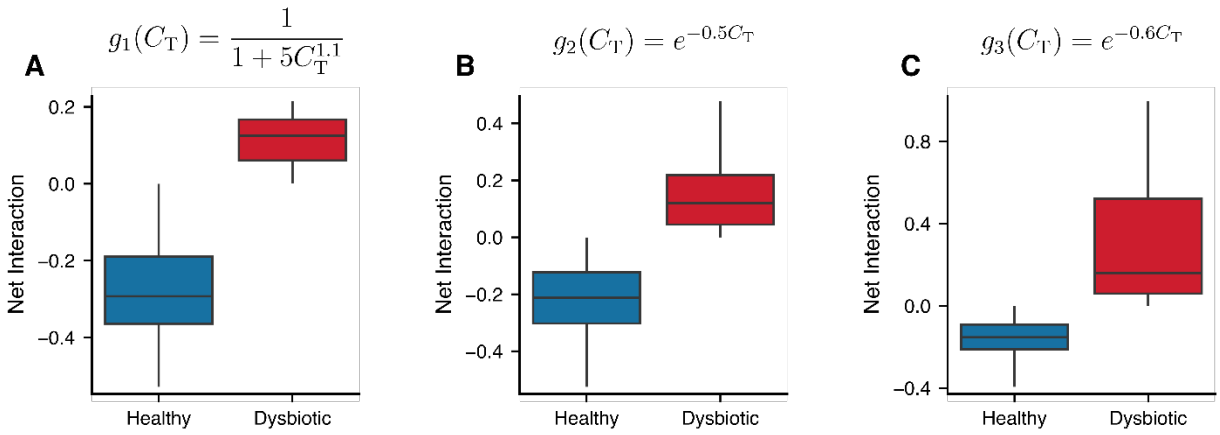


Figure S4: ***Robustness of alternative stable states to different trade-off functions.*** *Net interaction values for healthy and dysbiotic states across 10 independent simulations under three different trade-off functions: g_1 (left), g_2 (middle), and g_3 (right). Despite the different forms and parameterizations, the model consistently exhibits two distinct ecological regimes, with the dysbiotic state showing higher net interaction than the healthy one.*

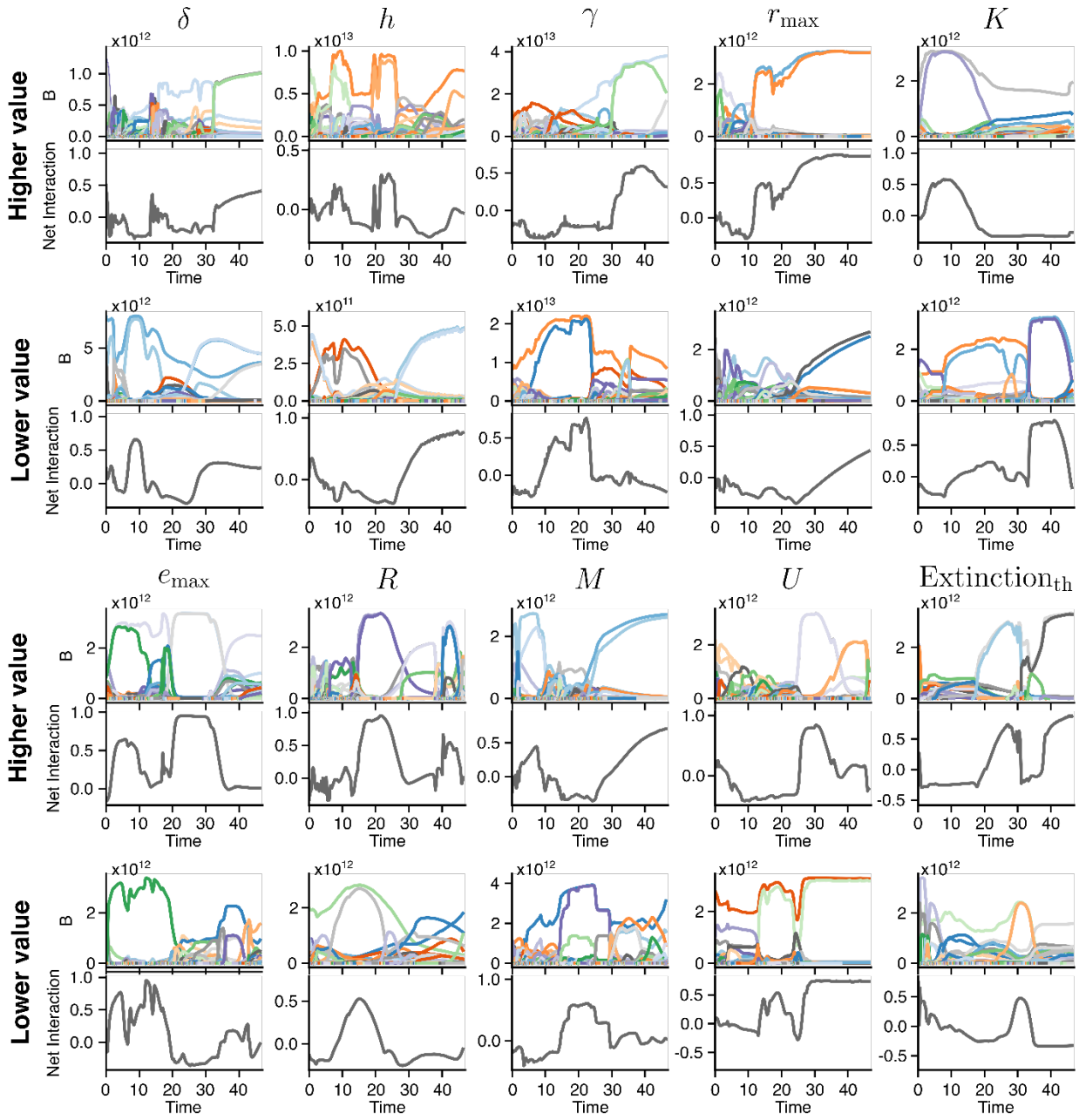


Figure S5: ***Alternative stable states are robust to parameter variations.*** *Time-series (in years) of total biomass (top row in each subpanel) and net interaction (bottom row in each subpanel) for different parameter settings in the model. Each column represents a change from the original parameter values reported in Table S1.* ***(A)*** *Increased values: from left to right,* $\delta = 2.5/24$, $h = 10$, $\gamma = 1.25 \cdot 10^{14}$, $r_{max} = 10^{-9}$, $K = 10^{-3}$. ***(B)*** *Decreased values: from left to right,* $\delta = 0.4/24$, $h = 0.5$, $\gamma = 10^{13}$, $r_{max} = 10^{-11}$, $K = 10^{-5}$. ***(C)*** *Increased values: from left to right,* $e_{max} = 4$, $R = 40$, $M = 120$, $U = 4$, *extinction threshold* $= 5 \cdot 10^{-4}$. ***(D)*** *Decreased values: from left to right,*

$e_{max} = 2$, $R = 20$, $M = 80$, $U = 1$, *extinction threshold* $= 5 \cdot 10^{-5}$. *In all cases, the model continues to exhibit both healthy and dysbiotic states, demonstrating that the emergence of alternative stable states is qualitatively robust to wide parameter changes.*

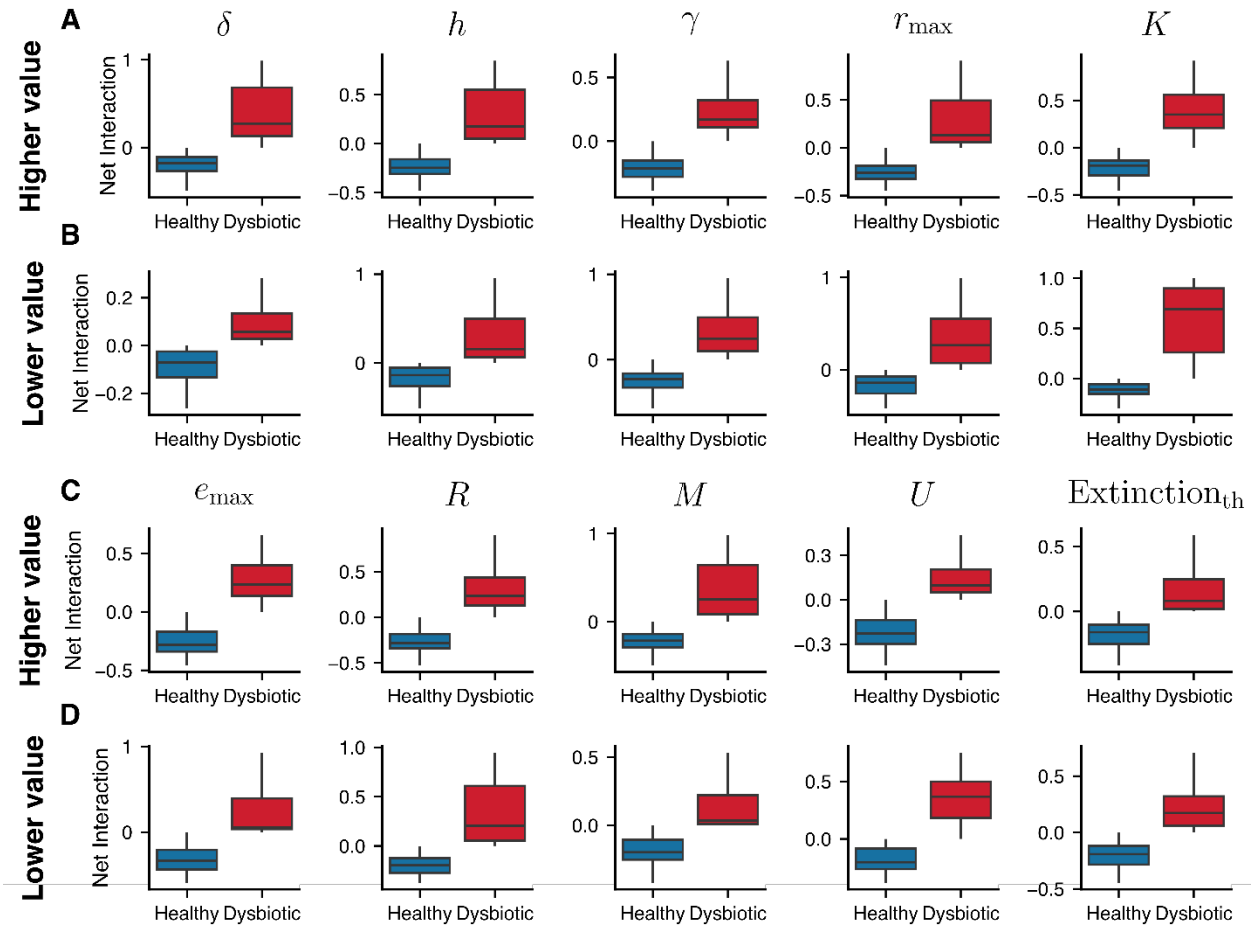


Figure S6: ***Robustness of alternative stable states and dysbiosis-related increase in net interaction against parameter variations.*** *Net interaction values for healthy and dysbiotic states across 10 independent simulations under variations of model parameters. Each boxplot corresponds to a distinct parameter setting, modified from the baseline values in Table S1. Higher values* ***(A)*** $\delta = 2.5/24$; $h = 10$; $\gamma = 1.25 \cdot 10^{14}$; $r_{max} = 10^{-9}$; $K = 10^{-3}$ *versus lower values* ***(B)*** $\delta = 0.4/24$; $h = 0.5$; $\gamma = 10^{13}$; $r_{max} = 10^{-11}$; $K = 10^{-5}$. *Higher values* ***(C)*** $e_{max} = 4$; $R = 40$; $M = 120$; $U = 4$; *extinction threshold* $= 5 \cdot 10^{-4}$ *versus lower values* ***(D)*** $e_{max} = 2$; $R = 20$; $M = 80$; $U = 1$; *extinction threshold* $= 5 \cdot 10^{-5}$. *In all cases, the model robustly generates two distinct ecological states, healthy and dysbiotic, with consistently higher net interaction in the dysbiotic state.*

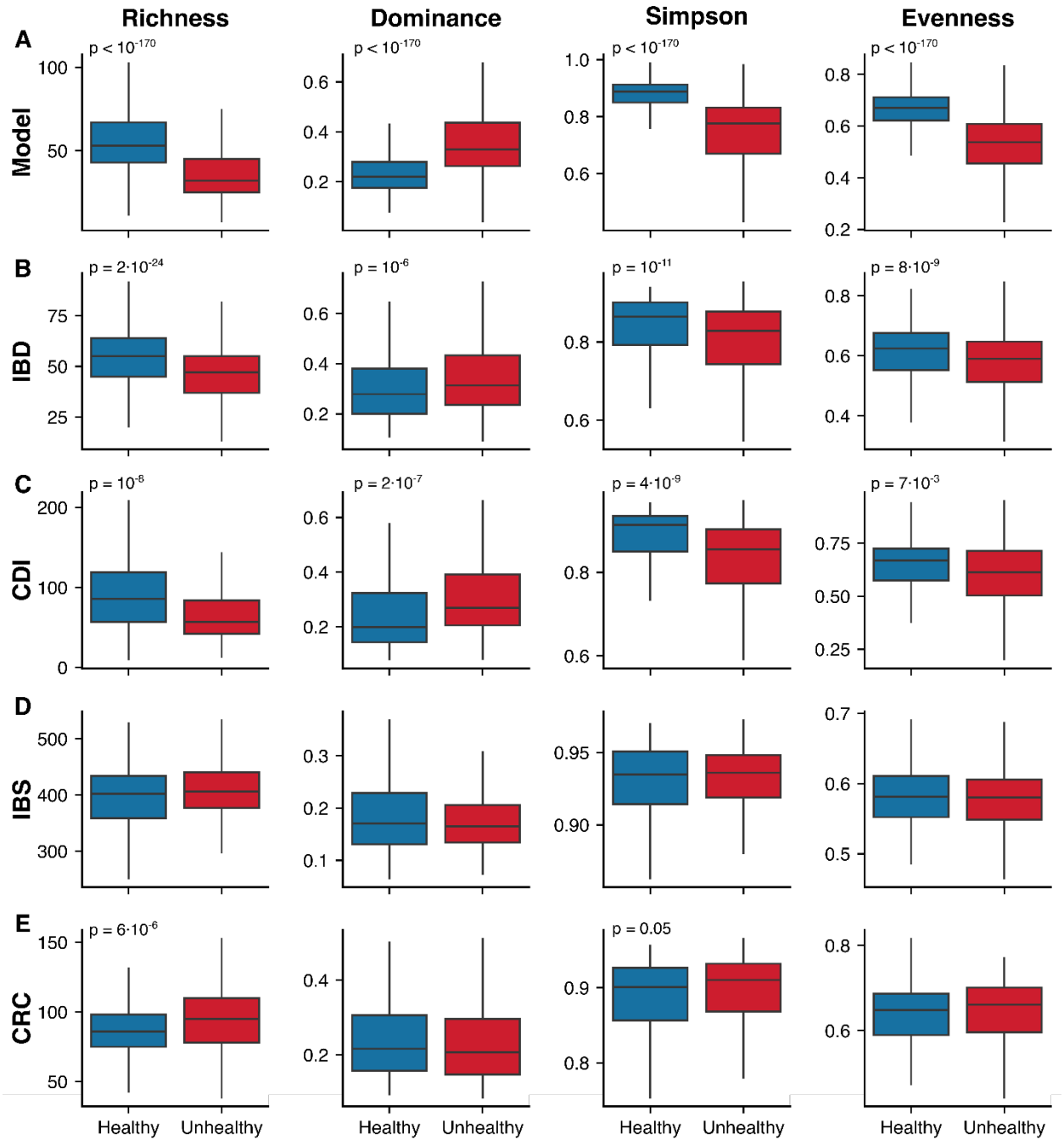


Figure S7: ***Diversity metrics in the model and real-world datasets.*** *Comparison of four diversity metrics (richness, dominance, Simpson's index, and Pielou's evenness) for the model* ***(A)*** *and four real-world datasets: IBD* ***(B)****, CDI* ***(C)****, IBS* ***(D)****, and CRC* ***(E)****. Definitions and computation details for each metric are provided in Methods (45).*

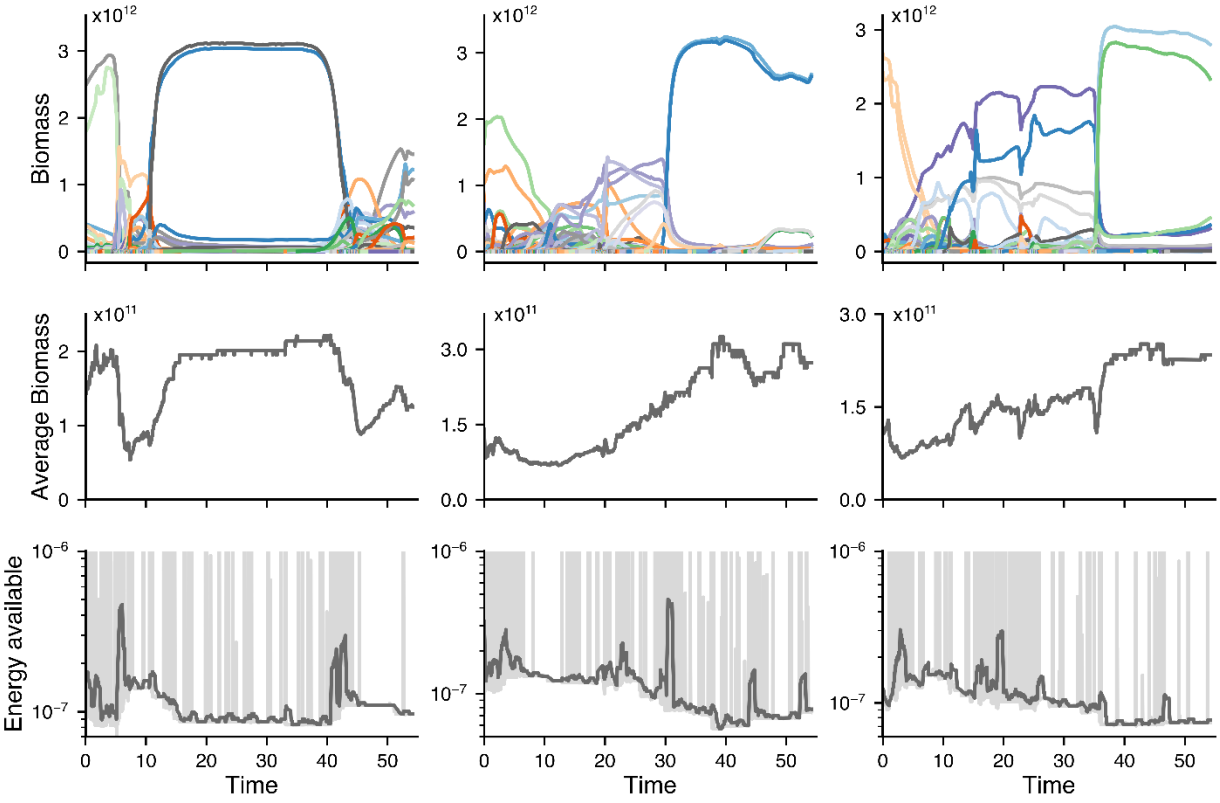


Figure S8: ***Dysbiotic state shows greater efficiency than the healthy state.*** *(**A**-**C**) Time-series (in years) of bacterial biomasses (**A**), average biomass across the system (**B**), and available energy within the system (**C**) for three different realizations of the model. In (**C**), dark grey lines represent the rolling average, while light grey lines depict the actual values. dysbiotic states show higher average biomass while depleting more available energy in the system, indicating a more efficient microbial community.*

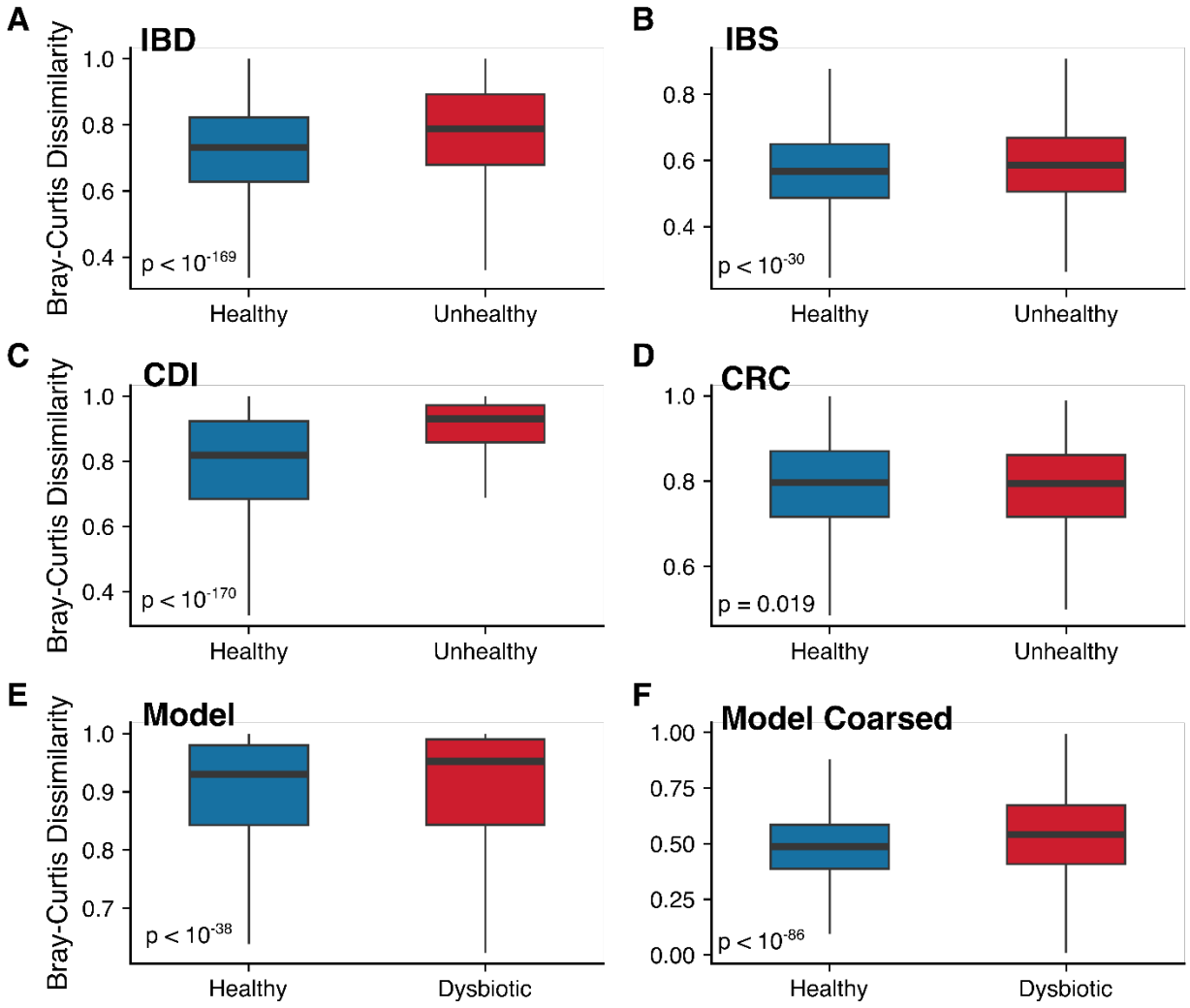


Figure S9: ***The Anna Karenina Principle holds across different diseases, but not all, and in the model.*** *Beta diversity calculated with the Bray-Curtis dissimilarity of healthy and diseased states for (**A**) IBD, (**B**) IBS, (**C**) CDI, (**D**) CRC, as well as for healthy and collapsed states in the model (**E**) and in the model with coarse-grained pathways (**F**) (see Methods (45)). Except for CRC, all diseases and the model follow the Anna Karenina Principle, which means that healthy communities are more similar to each other than dysbiotic ones. The presence of exceptions and the small magnitude of the increases in dissimilarity do not allow for strong conclusions from this observation (see (82)).* P-values (two-sided Mann–Whitney U test) were calculated as described in Methods (45).

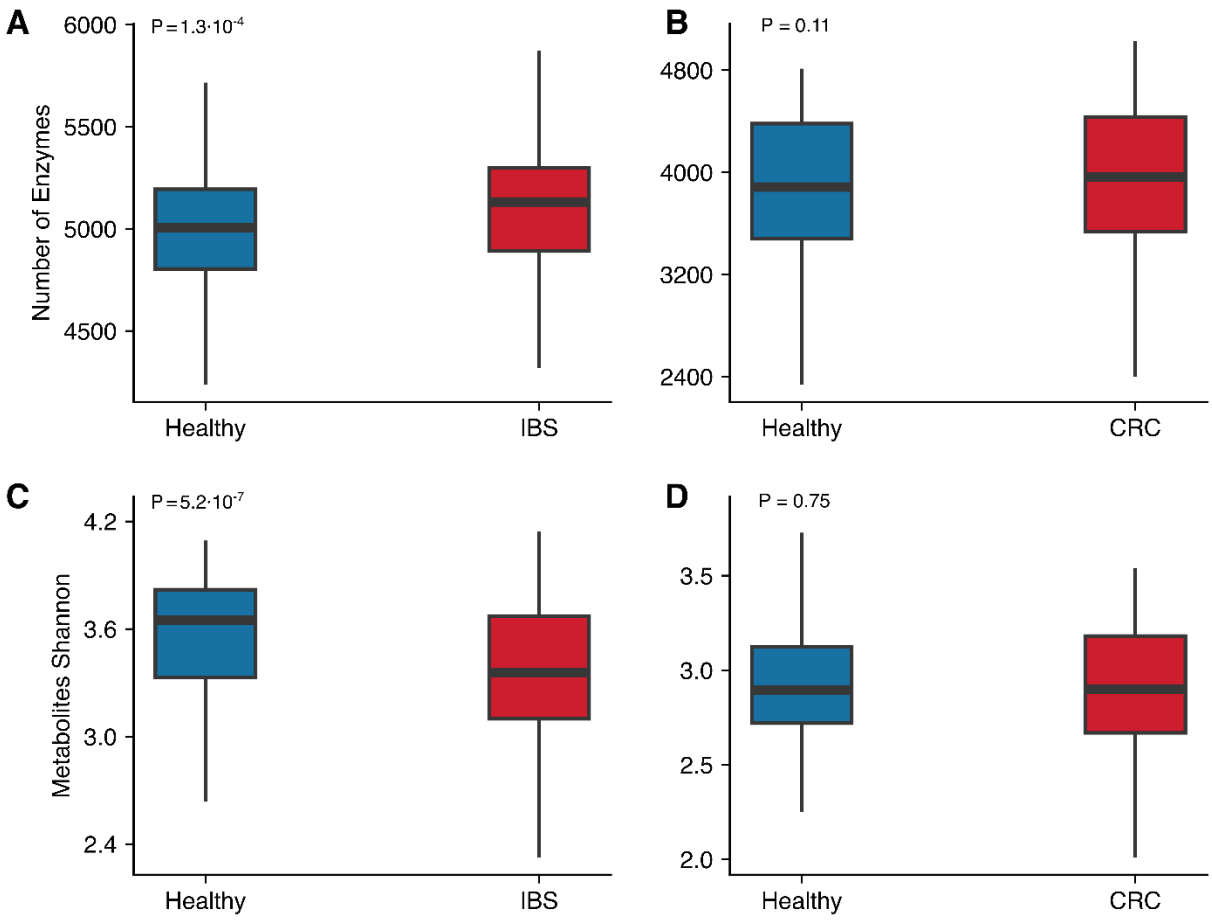


Figure S10: ***Analyzed biomarkers are not consistently reduced across diseases.*** *(**A-B**) Number of enzymes in Healthy vs. IBS (**A**) and Healthy vs. CRC (**B**). (**C-D**) Shannon index of secreted metabolites in Healthy vs. IBS (**C**) and Healthy vs. CRC (**D**). The inconsistencies observed across diseases indicate that these biomarkers alone are not reliable indicators of dysbiosis.* P-values (two-sided Mann–Whitney U test) were calculated as described in Methods (45).

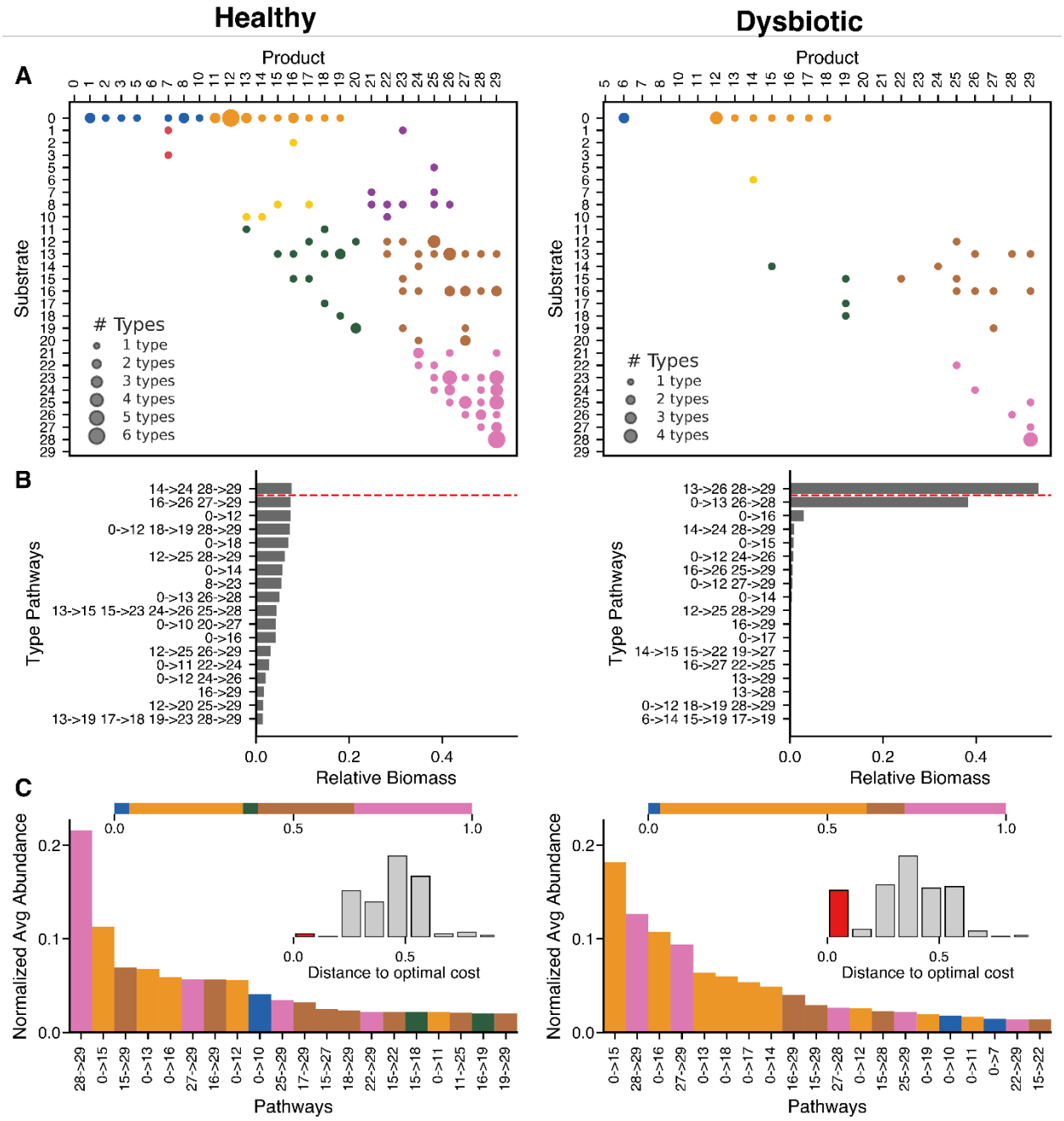


Figure S11: ***Pathway structure in healthy and dysbiotic states. (A)*** *Pathway usage matrix showing the number of bacterial lineages expressing each pathway (from substrate [row] to product [column]) at a representative snapshot of the healthy (left) and dysbiotic (right) states. Pathways are color-coded by functional category, as defined in Methods (45).* ***(B)*** *Relative biomass distribution for each pathway at the same snapshots as in panel A, for healthy (left) and dysbiotic (right) states. The red dashed line separates the 1% most abundant lineages.* ***(C)*** *Normalized average pathway abundance, focusing on the most abundant pathways among the top 1% most abundant bacterial*

lineages in each realization. Pathways are color-coded by functional category, with the horizontal bar in the inset showing the distribution of functional groups. The inset histograms depict the abundance of these top 1% bacterial lineages as a function of their distance to the theoretical optimum (given by the maximum of Eq. S18), and show that the most abundant lineages in the dysbiotic state are closer to the optimum than those in the healthy state (see red bars).

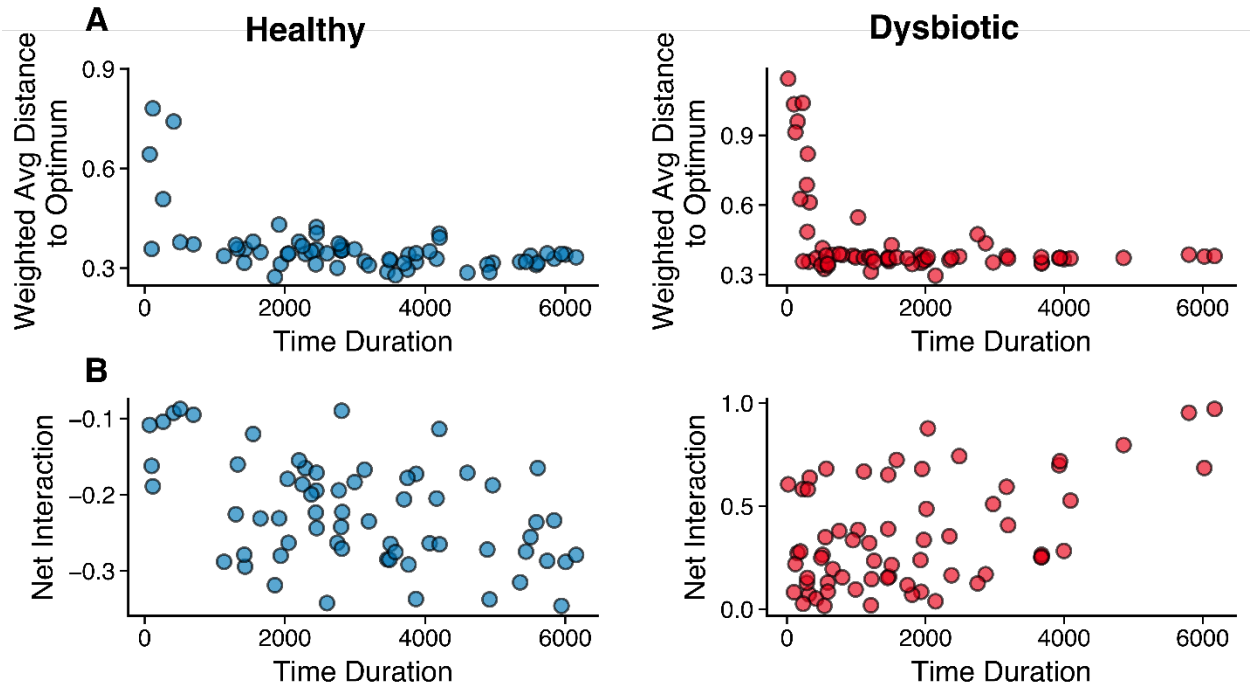


Figure S12: ***Duration of a state correlates with average metabolic efficiency and interaction structure. (A)*** *Relationship between the duration of a community state (healthy or dysbiotic) and the average distance to the optimum of the enzymatic cost–growth trade-off across its bacterial lineages, weighted by their relative abundance and duration. Each point represents the longest-lasting state for each simulation. Both healthy and dysbiotic states show a negative correlation (Spearman correlation of -0.51 with* $p - value = 1.8 \cdot 10^{-5}$ *and -0.31 with* $p - value = 0.013$*): communities composed of lineages closer to the theoretical optimum (i.e., more metabolically efficient) are more ecologically stable.* ***(B)*** *Relationship between the duration of a state and its average net interaction. In healthy states, longer durations correspond to more negative net interactions (i.e., stronger competition; Spearman correlation of -0.4 with* $p - value = 0.001$*), whereas in dysbiotic states, longer durations are associated with more*

positive net interactions (i.e., enhanced cross-feeding; Spearman correlation of 0.38 with $p - value = 0.0016$). Note that these durations (in hours), although illustrative, cannot be immediately linked to real-world timescales, since details influencing timing (e.g. initial state) differ from real-world situations.

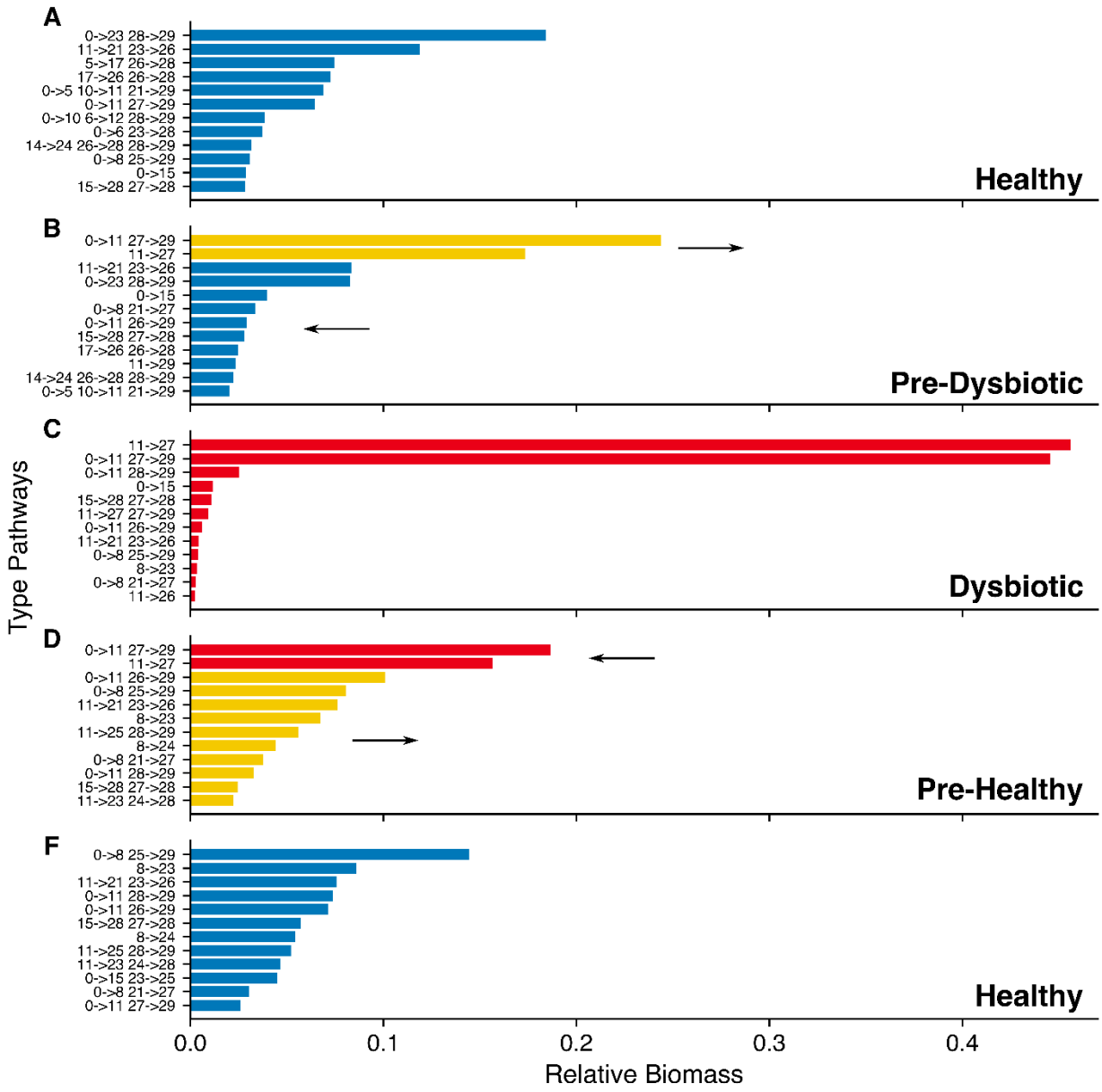


Figure S13: ***Representative example of transition between healthy and dysbiotic states.*** *Snapshots from a single realization illustrating a full transition cycle: from healthy to dysbiotic state and back. Each panel shows the relative biomass of the most abundant bacterial lineages at different stages:* ***(A)*** *healthy state,* ***(B)*** *pre-dysbiotic (i.e.,*

transition from healthy to dysbiotic), ***(C)*** *dysbiotic state,* ***(D)*** *pre-healthy (i.e., transition from dysbiotic to healthy), and* ***(E)*** *recovered healthy state. Colors represent lineages that dominate in healthy states (blue), dysbiotic states (red), or lineages that increase during transitions (yellow). This figure illustrates how (stochastic) arrivals of specific lineages and pathways can trigger or reverse dysbiosis.*

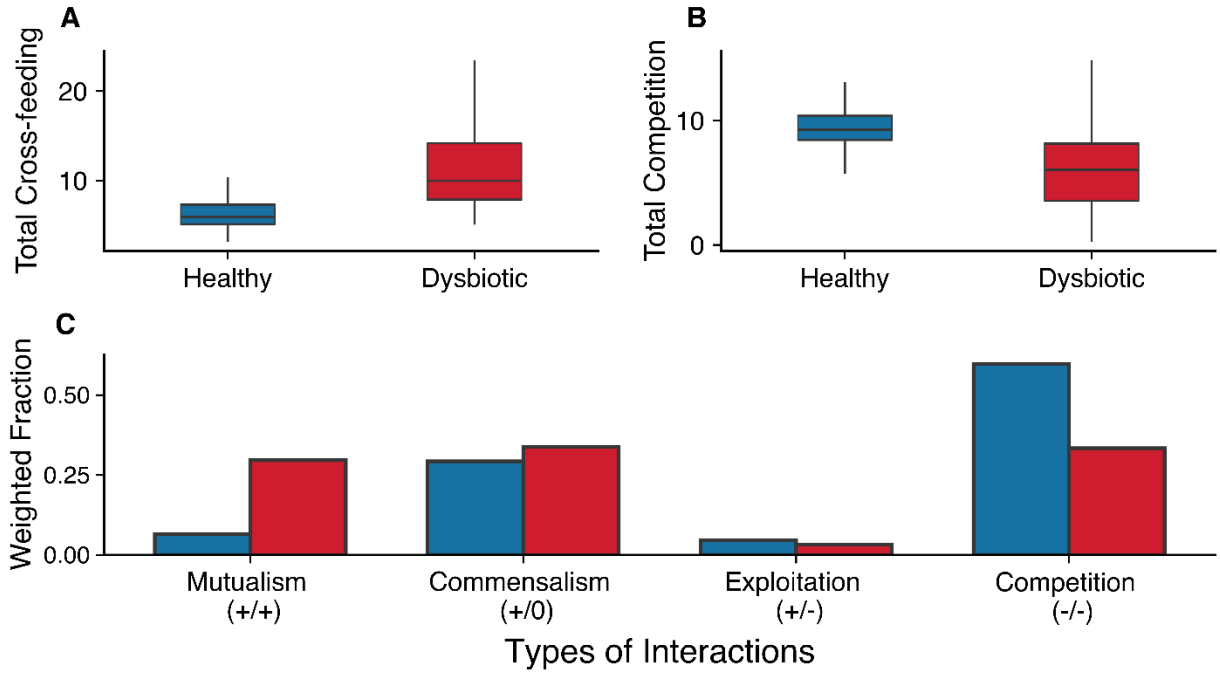


Figure S14: ***Interaction patterns in healthy and dysbiotic states in the model. (A)*** *Total cross-feeding,* $c_T^+ = \sum_{ij} c_{ij}^+$*,* *and* ***(B)*** *total competition,* $c_T^- = \sum_{ij} c_{ij}^-$*, where* c_{ij}^+ *and* c_{ij}^- *are computed using Eqs. 8 and S9, respectively, for healthy and dysbiotic states.* ***(C)*** *Relative frequency of interaction lineages between bacterial pairs* (i, j)*, classified by the signs of* $C_{ij} = c_{ij}^+ - c_{ij}^-$ *and its reciprocal* C_{ji}*. Frequencies are weighted by the absolute value* $|C_{ij}| + |C_{ji}|$*. In agreement with real data (64) our model shows that negative interactions are predominant. Besides, diseased states in the model show an increase of positive interactions (specially mutualism) and decrease of competition.*

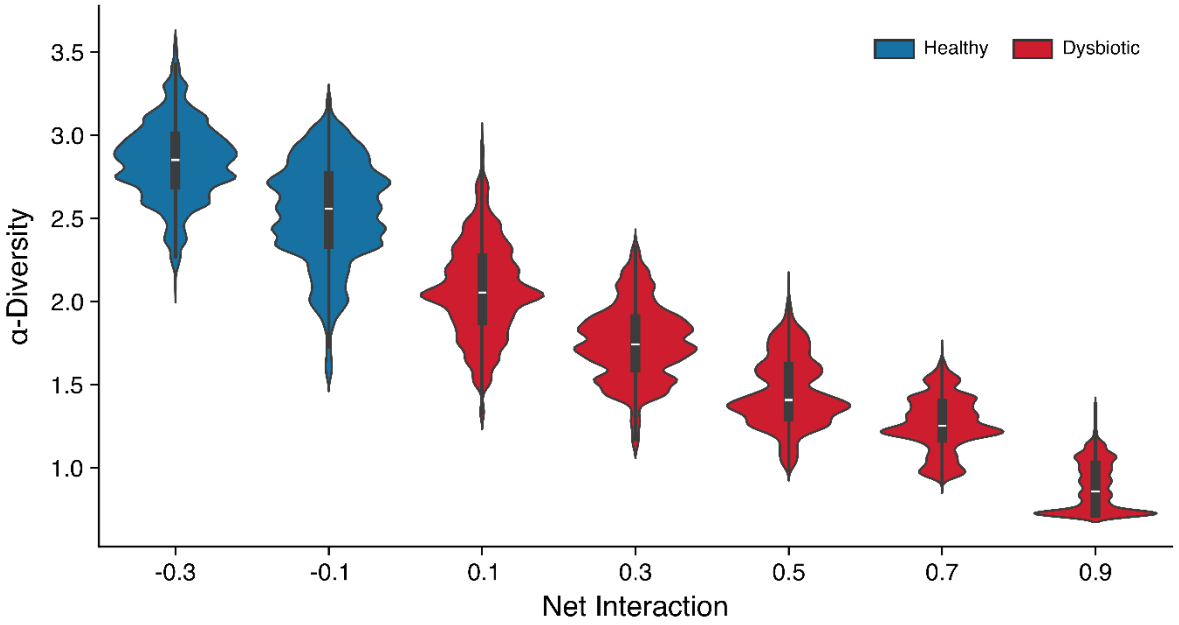


Figure S15: ***Diversity exhibits high variability across net interaction values.*** *The mean and median of α-diversity for a given value of the net interaction shows a negative correlation between the Shannon index and net interaction. Diversity is, however, widely distributed, which suggests that diversity alone is not a reliable indicator of the underlying interaction structure. For example, for the data in the figure, a value of α-diversity of 1.5 could be observed for any network with net interaction in $[-0.1,0.7]$. Taken together, these results support the claim that a switch to a dysbiotic state (i.e. to a network with more positive interactions) typically, but not always, correlates with a decrease in diversity.*

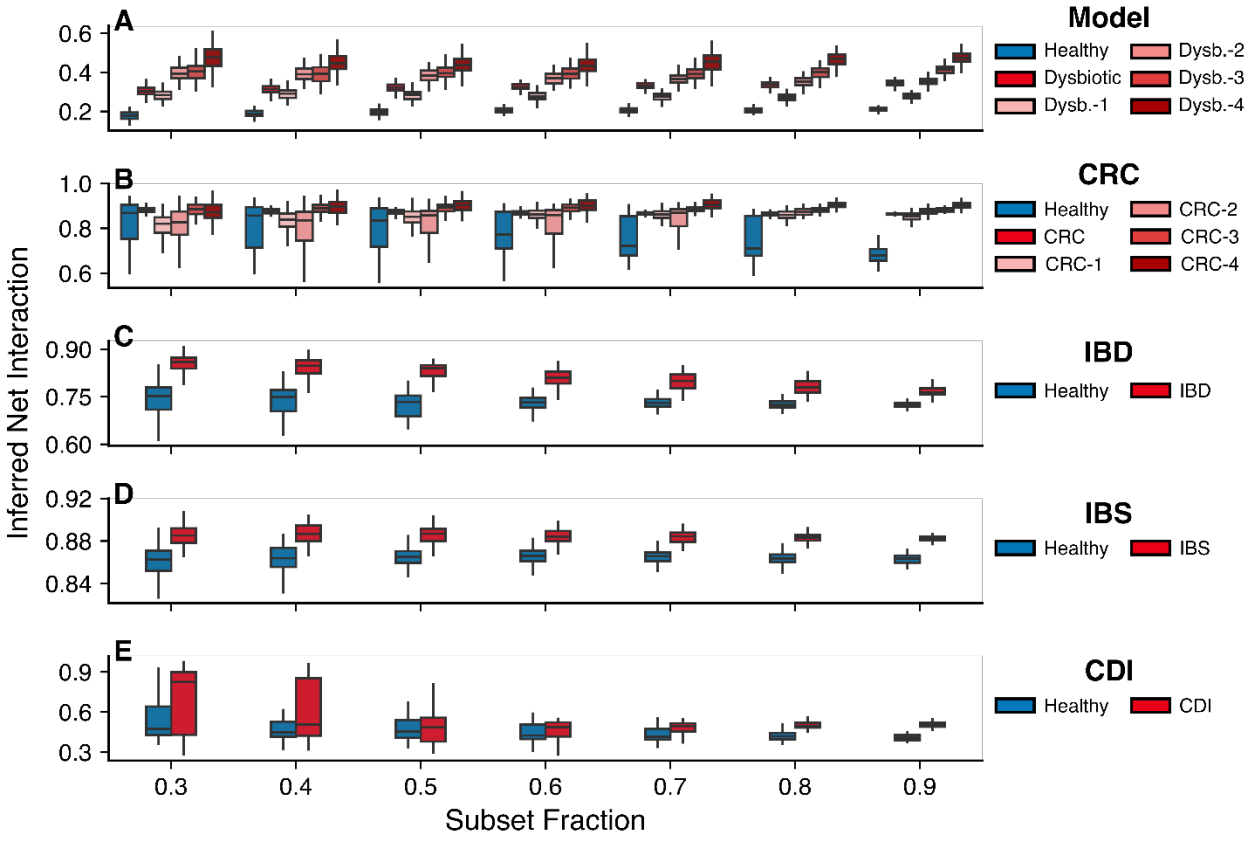


Figure S16: ***The inferred net interaction is robust under bootstrapping.*** *Inferred net interaction calculated for different subsampling fractions of the original dataset for the model (**A**), CRC (**B**), IBD (**C**), IBS (**D**), and CDI (**E**). The inferred net interaction reliably captures the qualitative trends of the true underlying net interaction, ρ, in the model across all subsampled datasets. Healthy states consistently show lower inferred net interaction compared to diseased (dysbiotic) states across all datasets. Additionally, for CRC, the inferred net interaction correlates with disease progression, with more advanced disease stages displaying systematically higher inferred net interaction across all subsampled sets. For clarity and better visualization, we opted not to include p-values and Cliff's delta values in the figure. However, all comparisons are statistically highly significant (based on 500 bootstrap-generated samples), with medium to large effect sizes as measured by Cliff's delta.*

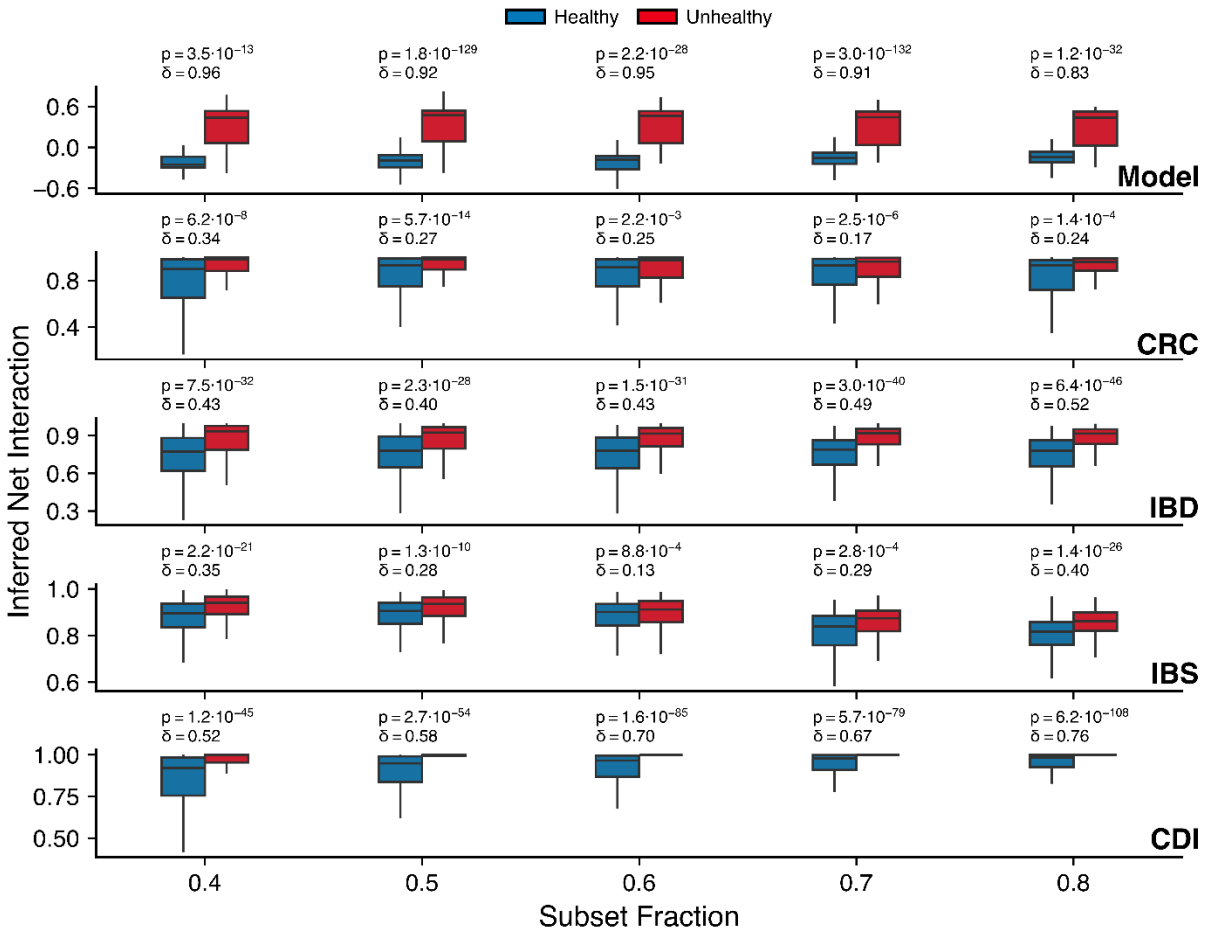


Figure S17: ***Inferred net interaction using BEEM-Static.*** *Inferred net interaction estimated with the BEEM-Static algorithm (79) (see Methods (45)) for different subsampling fractions of the original dataset for the model (**A**), CRC (**B**), IBD (**C**), IBS (**D**), and CDI (**E**). Across all cases, dysbiotic states in the model and disease states in real datasets show a consistently higher inferred net interaction, suggesting a shift toward more cooperative microbial communities compared to healthy states and healthy controls, respectively. All results are highly statistically significant (p-value calculated with two-sided Mann-Whitney U test; 500 bootstraping-generated samples) with medium to large effect sizes (Cliff's deltas).*

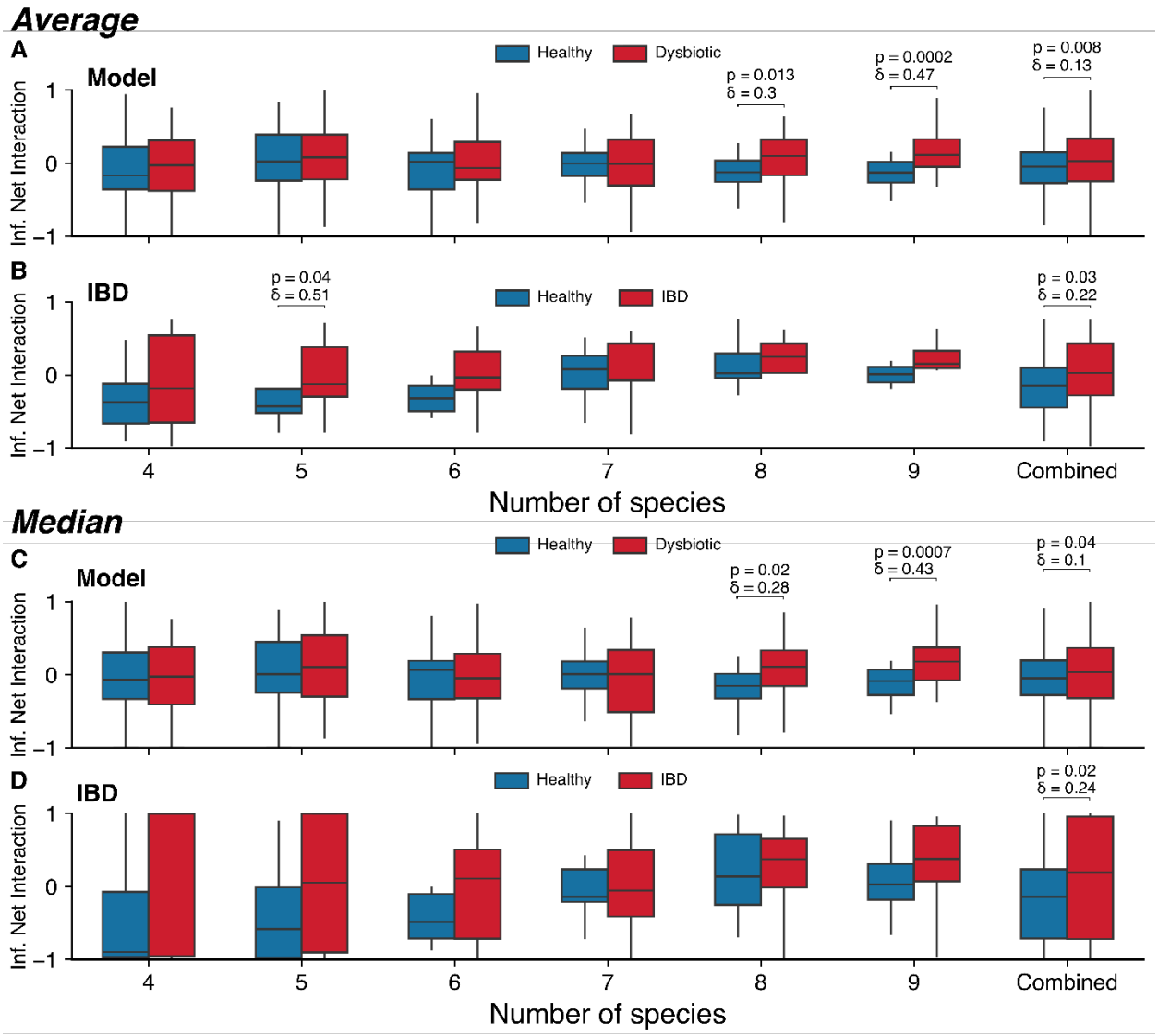


Figure S18: ***Inferred net interaction using LIMITS.*** *Net interaction inferred using the LIMITS algorithm (80) for model simulations (**A, C**) and the IBD dataset (**B, D**), using either the mean (average, top panels) or median (bottom panels) to construct the interaction networks. Either choice does not affect the conclusion: in all cases dysbiotic (model) and diseased (IBD) communities exhibit higher net interaction compared to their healthy counterparts. However, these results should be interpreted with caution, as not all comparisons reach statistical significance, likely due to the limited number of patients with at least 20 longitudinal samples.* Effect sizes (Cliff's delta) and p-values (two-sided Mann–Whitney U test) were calculated as described in Methods (45).

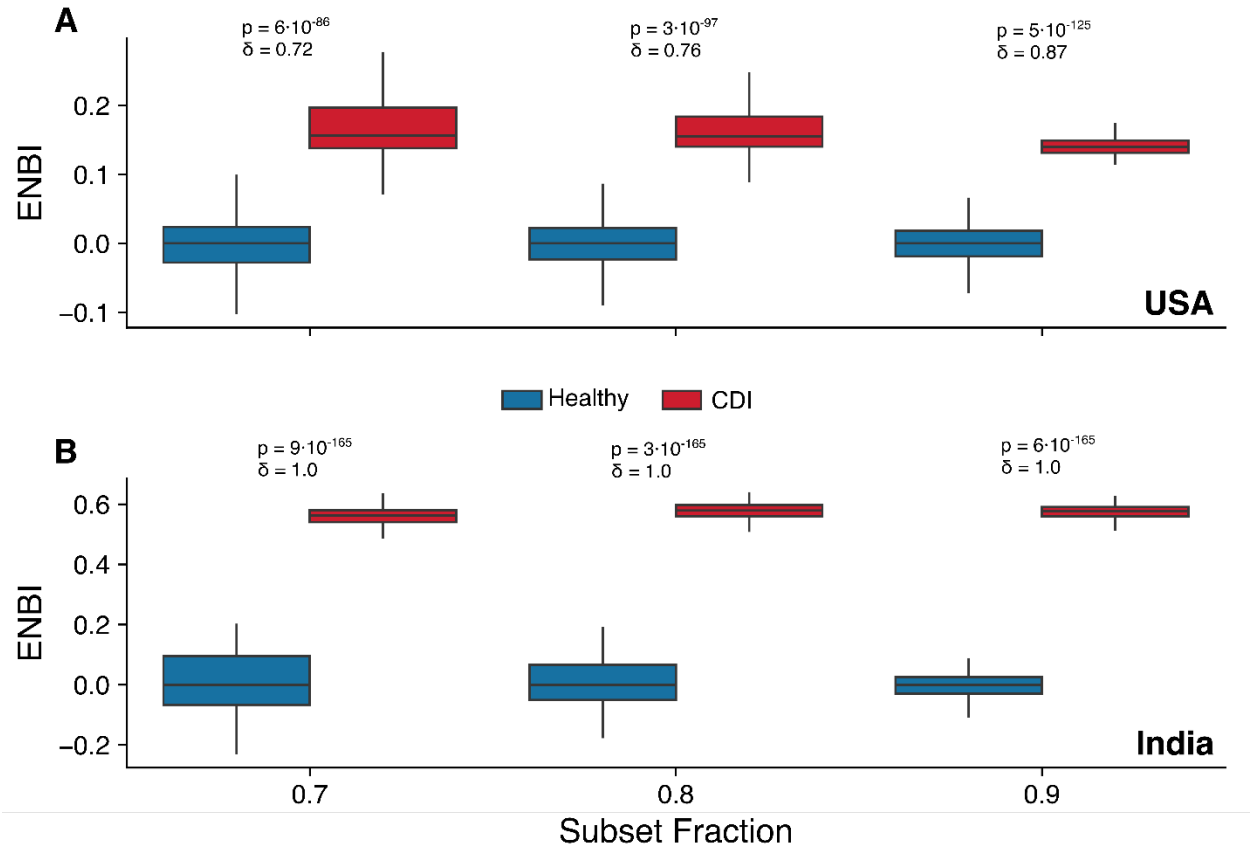


Figure S19: ***ENBI distinguishes healthy and diseased states across regions.*** *ENBI computed for Clostridioides difficile infection (CDI) patients and healthy controls from two sub-cohorts (USA* ***(A)*** *and India* ***(B)****) in the same dataset. In both regions, CDI samples show higher ENBI, indicating consistent shifts toward more positive net interactions regardless of geographic or dietary background. Effect sizes (δ, Cliff's delta) are large, indicating that the statistically significant differences (p-value calculated via two-sidedMann-Whitney U test) likely reflect biologically meaningful shifts.*

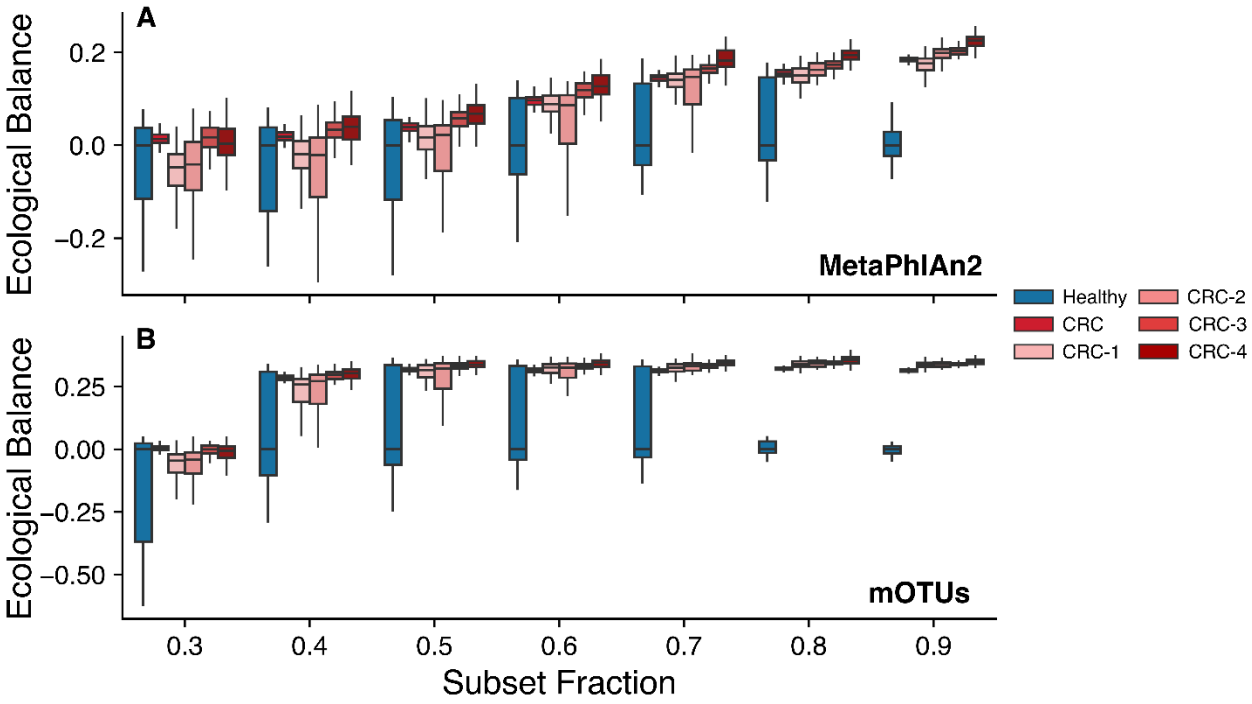


Figure S20: ***The ecological network balance index (ENBI) is robust across different taxonomic profiling methodologies.*** *ENBI calculated for different subsampling fractions of the original CRC datasets using two distinct taxonomic profiling methodologies: MetaPhlAn2 (111) (**A**) and mOTUs (112) (**B**). In both cases, disease states consistently show a positive ENBI value, indicating a shift toward more cooperative microbial communities compared to controls. Furthermore, ENBI correlates with disease progression across all subsampled datasets, reinforcing its robustness as a biomarker irrespective of the taxonomic profiling method used. All results are highly statistically significant (500 bootstraping-generated samples).*

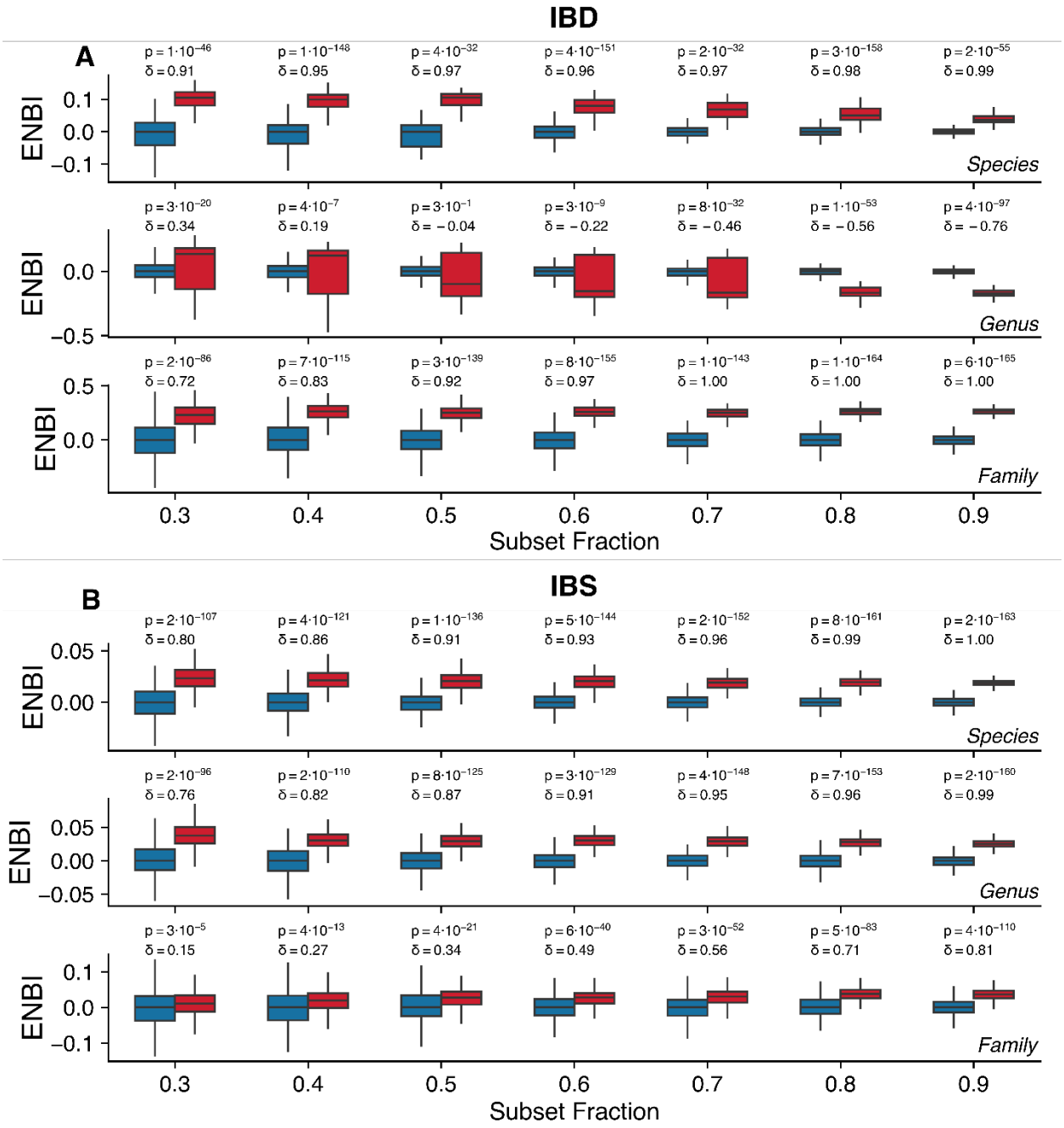


Figure S21: ***Robustness of the ENBI to taxonomic resolution.*** *ENBI computed at different taxonomic levels—species, genus, and family—for* ***(A)*** *IBD and* ***(B)*** *IBS datasets. In both cases, the ENBI robustly separates healthy and diseased states across taxonomic resolutions, with the exception of IBD at the genus level which, despite being statistically significant (p-value calculated via two-sided Mann-Whitney U test), shows smaller effect sizes (Cliff's delta), indicating a less meaningful signal.*

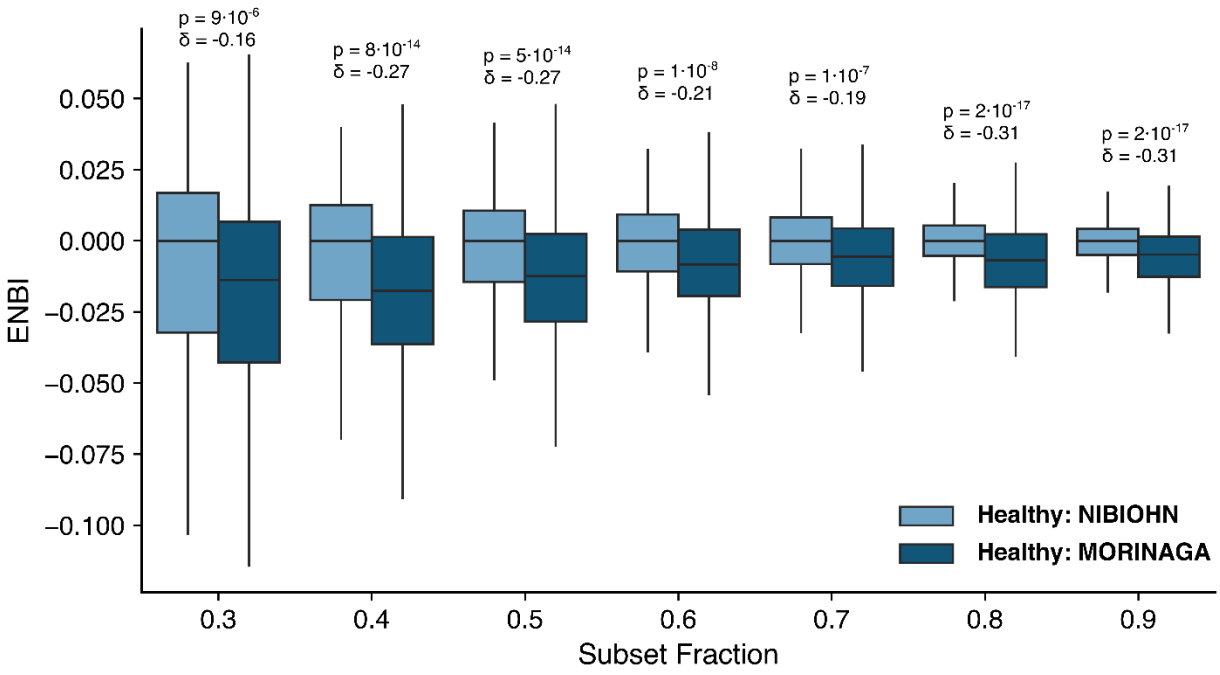


Figure S22: ***ENBI does not differentiate healthy cohorts.*** *ENBI computed for two large Japanese healthy cohorts (MORINAGA and NIBIOHN) from (104), using NIBIOHN as reference. Despite differences in geography, diet, and microbiome composition, ENBI values remain comparable, supporting its robustness. While statistical tests yield significant p-values (calculated via two-sided Mann-Whitney U test), the effect size (δ, Cliff's delta) is small, suggesting that the significance likely reflects the large sample size (500 bootstrap samples per group) rather than a biologically meaningful difference.*

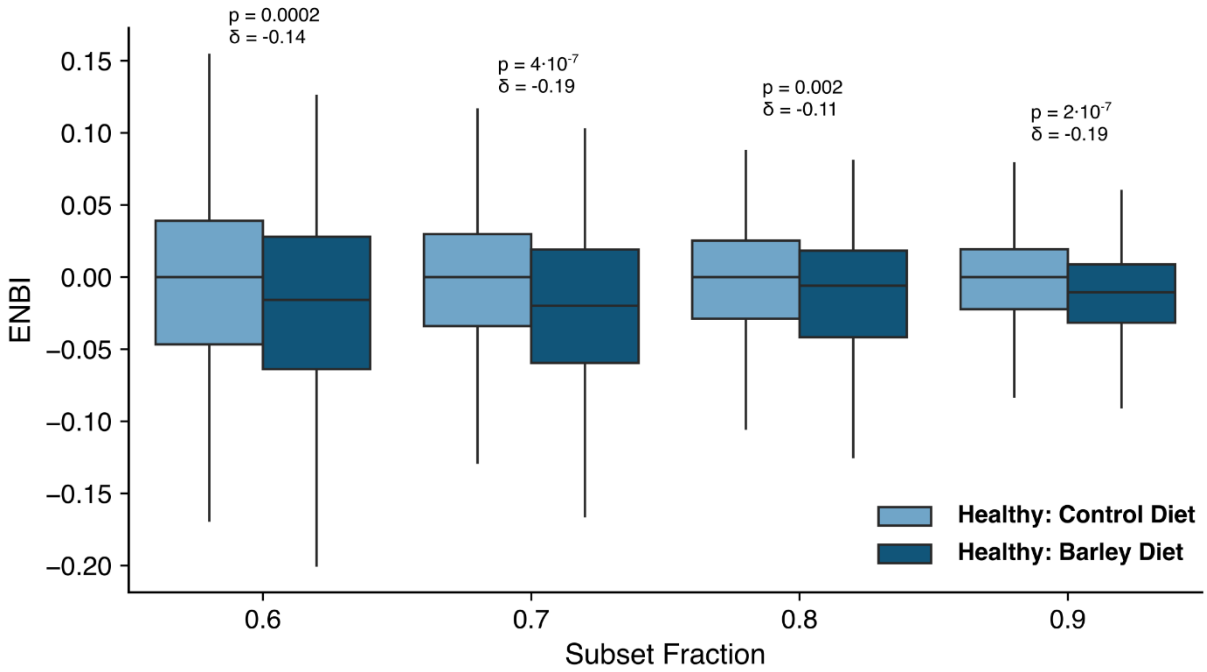


Figure S23: ***ENBI is unaffected by mild dietary interventions.*** *ENBI computed from a crossover dietary study in healthy individuals (105), comparing a barley-rich versus a controlled diet, using the latter as a reference case. ENBI values remain nearly identical, supporting its robustness to moderate diet shifts. As in the previous figure, while statistical tests yield significant p-values (calculated via two-sided Mann-Whitney U test), the effect size (δ, Cliff's delta) is small, suggesting that the significance likely reflects the large sample size (500 bootstrap samples per group) rather than a biologically meaningful difference.*

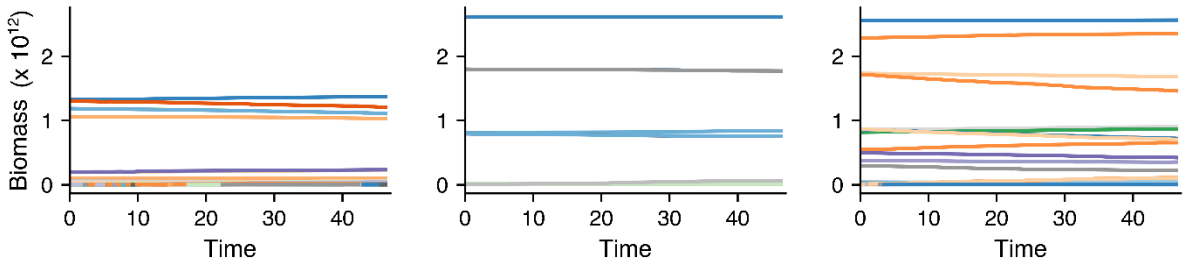


Figure S24: ***Linear trade-offs yield biologically implausible dynamics.*** *Three realizations of the model using a linear trade-off function $g(C_T) = 1/C_T$, which lacks an optimum for Eq. S21. In all cases, the system quickly becomes trapped in a single community state and exhibits negligible temporal variability, which is biologically unrealistic (100).*

References and Notes

1. W. M. de Vos, E. A. de Vos, Role of the intestinal microbiome in health and disease: From correlation to causation. *Nutr. Rev.* **70**, S45–S56 (2012). doi:10.1111/j.1753-4887.2012.00505.x Medline

2. W. M. de Vos, H. Tilg, M. Van Hul, P. D. Cani, Gut microbiome and health: Mechanistic insights. *Gut* **71**, 1020–1032 (2022). doi:10.1136/gutjnl-2021-326789 Medline

3. J. A. Gilbert, R. A. Quinn, J. Debelius, Z. Z. Xu, J. Morton, N. Garg, J. K. Jansson, P. C. Dorrestein, R. Knight, Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature* **535**, 94–103 (2016). doi:10.1038/nature18850 Medline

4. K. Hou, Z.-X. Wu, X.-Y. Chen, J.-Q. Wang, D. Zhang, C. Xiao, D. Zhu, J. B. Koya, L. Wei, J. Li, Z.-S. Chen, Microbiota in health and diseases. *Signal Transduct. Target. Ther.* **7**, 135 (2022). doi:10.1038/s41392-022-00974-4 Medline

5. Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012). doi:10.1038/nature11234 Medline

6. A. Zhernakova, A. Kurilshikov, M. J. Bonder, E. F. Tigchelaar, M. Schirmer, T. Vatanen, Z. Mujagic, A. V. Vila, G. Falony, S. Vieira-Silva, J. Wang, F. Imhann, E. Brandsma, S. A. Jankipersadsing, M. Joossens, M. C. Cenit, P. Deelen, M. A. Swertz, LifeLines cohort study, R. K. Weersma, E. J. M. Feskens, M. G. Netea, D. Gevers, D. Jonkers, L. Franke, Y. S. Aulchenko, C. Huttenhower, J. Raes, M. H. Hofker, R. J. Xavier, C. Wijmenga, J. Fu, Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* **352**, 565–569 (2016). doi:10.1126/science.aad3369 Medline

7. G. Vrancken, A. C. Gregory, G. R. B. Huys, K. Faust, J. Raes, Synthetic ecology of the human gut microbiota. *Nat. Rev. Microbiol.* **17**, 754–763 (2019). doi:10.1038/s41579-019-0264-8 Medline

8. I. Cho, M. J. Blaser, The human microbiome: At the interface of health and disease. *Nat. Rev. Genet.* **13**, 260–270 (2012). doi:10.1038/nrg3182 Medline

9. J. Jovel, L. A. Dieleman, D. Kao, A. L. Mason, E. Wine, "The Human Gut Microbiome in Health and Disease," in *Metagenomics*, M. Nagarajan, Ed. (Academic Press, 2018), pp. 197–213.

10. Y. Fan, O. Pedersen, Gut microbiota in human metabolic health and disease. *Nat. Rev. Microbiol.* **19**, 55–71 (2021). doi:10.1038/s41579-020-0433-9 Medline

11. F. Sommer, J. M. Anderson, R. Bharti, J. Raes, P. Rosenstiel, The resilience of the intestinal microbiota influences health and disease. *Nat. Rev. Microbiol.* **15**, 630–638 (2017). doi:10.1038/nrmicro.2017.58 Medline

12. A. M. Valdes, J. Walter, E. Segal, T. D. Spector, Role of the gut microbiota in nutrition and health. *BMJ* **361**, k2179 (2018). doi:10.1136/bmj.k2179 Medline

13. C. Duvallet, S. M. Gibbons, T. Gurry, R. A. Irizarry, E. J. Alm, Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat. Commun.* **8**, 1784 (2017). doi:10.1038/s41467-017-01973-8 Medline

14. Y. Wang, S. Zhang, T. J. Borody, F. Zhang, Encyclopedia of fecal microbiota transplantation: a review of effectiveness in the treatment of 85 diseases. *Chin. Med. J.* **135**, 1927–1939 (2022). doi:10.1097/CM9.0000000000002339

15. Q. Yang, Q. Liang, B. Balakrishnan, D. P. Belobrajdic, Q.-J. Feng, W. Zhang, Role of Dietary Nutrients in the Modulation of Gut Microbiota: A Narrative Review. *Nutrients* **12**, 381 (2020). doi:10.3390/nu12020381 Medline

16. O.-A. Petrariu, I. C. Barbu, A.-G. Niculescu, M. Constantin, G. A. Grigore, R.-E. Cristian, G. Mihaescu, C. O. Vrancianu, Role of probiotics in managing various human diseases, from oral pathology to cancer and gastrointestinal diseases. *Front. Microbiol.* **14**, 1296447 (2024). doi:10.3389/fmicb.2023.1296447 Medline

17. T. S. B. Schmidt, J. Raes, P. Bork, The Human Gut Microbiome: From Association to Modulation. *Cell* **172**, 1198–1215 (2018). doi:10.1016/j.cell.2018.02.044 Medline

18. M. Kumar, B. Ji, K. Zengler, J. Nielsen, Modelling approaches for studying the microbiome. *Nat. Microbiol.* **4**, 1253–1267 (2019). doi:10.1038/s41564-019-0491-9 Medline

19. I. Sekirov, S. L. Russell, L. C. M. Antunes, B. B. Finlay, Gut microbiota in health and disease. *Physiol. Rev.* **90**, 859–904 (2010). doi:10.1152/physrev.00045.2009 Medline

20. A. B. Shreiner, J. Y. Kao, V. B. Young, The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.* **31**, 69–75 (2015). doi:10.1097/MOG.0000000000000139 Medline

21. V. K. Gupta, M. Kim, U. Bakshi, K. Y. Cunningham, J. M. Davis 3rd, K. N. Lazaridis, H. Nelson, N. Chia, J. Sung, A predictive index for health status using species-level gut microbiome profiling. *Nat. Commun.* **11**, 4635 (2020). doi:10.1038/s41467-020-18476-8 Medline

22. M. Seppi, J. Pasqualini, S. Facchin, E. V. Savarino, S. Suweis, Emergent functional organization of gut microbiomes in health and diseases. *Biomolecules* **14**, 5 (2023). doi:10.3390/biom14010005 Medline

23. A. Mosca, M. Leclerc, J. P. Hugot, Gut Microbiota Diversity and Human Diseases: Should We Reintroduce Key Predators in Our Ecosystem? *Front. Microbiol.* **7**, 455 (2016). doi:10.3389/fmicb.2016.00455 Medline

24. M. Kriss, K. Z. Hazleton, N. M. Nusbacher, C. G. Martin, C. A. Lozupone, Low diversity gut microbiota dysbiosis: Drivers, functional implications and recovery. *Curr. Opin. Microbiol.* **44**, 34–40 (2018). doi:10.1016/j.mib.2018.07.003 Medline

25. C. Petersen, J. L. Round, Defining dysbiosis and its influence on host immunity and disease. *Cell. Microbiol.* **16**, 1024–1033 (2014). doi:10.1111/cmi.12308 Medline

26. Z. S. Ma, L. Li, N. J. Gotelli, Diversity-disease relationships and shared species analyses for human microbiome-associated diseases. *ISME J.* **13**, 1911–1919 (2019). doi:10.1038/s41396-019-0395-y Medline

27. L. Aldars-García, M. Chaparro, J. P. Gisbert, Systematic Review: The Gut Microbiome and Its Potential Clinical Application in Inflammatory Bowel Disease. *Microorganisms* **9**, 977 (2021). doi:10.3390/microorganisms9050977 Medline

28. E. Martinez, B. Taminiau, C. Rodriguez, G. Daube, Gut Microbiota Composition Associated with *Clostridioides difficile* Colonization and Infection. *Pathogens* **11**, 781 (2022). doi:10.3390/pathogens11070781 Medline

29. R. Pittayanon, J. T. Lau, Y. Yuan, G. I. Leontiadis, F. Tse, M. Surette, P. Moayyedi, Gut Microbiota in Patients With Irritable Bowel Syndrome-A Systematic Review. *Gastroenterology* **157**, 97–108 (2019). doi:10.1053/j.gastro.2019.03.049 Medline

30. A. M. Thomas, P. Manghi, F. Asnicar, E. Pasolli, F. Armanini, M. Zolfo, F. Beghini, S. Manara, N. Karcher, C. Pozzi, S. Gandini, D. Serrano, S. Tarallo, A. Francavilla, G. Gallo, M. Trompetto, G. Ferrero, S. Mizutani, H. Shiroma, S. Shiba, T. Shibata, S. Yachida, T. Yamada, J. Wirbel, P. Schrotz-King, C. M. Ulrich, H. Brenner, M. Arumugam, P. Bork, G. Zeller, F. Cordero, E. Dias-Neto, J. C. Setubal, A. Tett, B. Pardini, M. Rescigno, L. Waldron, A. Naccarati, N. Segata, Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat. Med.* **25**, 667–678 (2019). doi:10.1038/s41591-019-0405-7 Medline

31. S. A. Shetty, F. Hugenholtz, L. Lahti, H. Smidt, W. M. de Vos, Intestinal microbiome landscaping: Insight in community assemblage and implications for microbial modulation strategies. *FEMS Microbiol. Rev.* **41**, 182–199 (2017). doi:10.1093/femsre/fuw045 Medline

32. M. Van de Guchte, S. D. Burz, J. Cadiou, J. Wu, S. Mondot, H. M. Blottière, J. Doré, Alternative stable states in the intestinal ecosystem: Proof of concept in a rat model and a perspective of therapeutic implications. *Microbiome* **8**, 153 (2020). doi:10.1186/s40168-020-00933-7 Medline

33. M. Arumugam, J. Raes, E. Pelletier, D. Le Paslier, T. Yamada, D. R. Mende, G. R. Fernandes, J. Tap, T. Bruls, J.-M. Batto, M. Bertalan, N. Borruel, F. Casellas, L. Fernandez, L. Gautier, T. Hansen, M. Hattori, T. Hayashi, M. Kleerebezem, K. Kurokawa, M. Leclerc, F. Levenez, C. Manichanh, H. B. Nielsen, T. Nielsen, N. Pons, J. Poulain, J. Qin, T. Sicheritz-Ponten, S. Tims, D. Torrents, E. Ugarte, E. G. Zoetendal, J. Wang, F. Guarner, O. Pedersen, W. M. de Vos, S. Brunak, J. Doré, MetaHIT Consortium, J. Weissenbach, S. D. Ehrlich, P. Bork, Enterotypes of the human gut microbiome. *Nature* **473**, 174–180 (2011). doi:10.1038/nature09944 Medline

34. L. Lahti, J. Salojärvi, A. Salonen, M. Scheffer, W. M. de Vos, Tipping elements in the human intestinal ecosystem. *Nat. Commun.* **5**, 4344 (2014). doi:10.1038/ncomms5344 Medline

35. S. Zaoli, J. Grilli, A macroecological description of alternative stable states reproduces intra- and inter-host variability of gut microbiome. *Sci. Adv.* **7**, eabj2882 (2021). doi:10.1126/sciadv.abj2882 Medline

36. H. Fujita, M. Ushio, K. Suzuki, M. S. Abe, M. Yamamichi, K. Iwayama, A. Canarini, I. Hayashi, K. Fukushima, S. Fukuda, E. T. Kiers, H. Toju, Alternative stable states, nonlinear behavior, and predictability of microbiome dynamics. *Microbiome* **11**, 63 (2023). doi:10.1186/s40168-023-01474-5 Medline

37. D. Garza, B. Liu, C. van de Velde, P. Saha, X. Zhou, D. Gonze, K. Simoens, K. Bernaerts, K. Faust, Discovery of alternative stable states in a synthetic human gut microbial community. bioRxiv 2024.11.28.625814 (2024). doi:10.1101/2024.11.28.625814

38. K. Faust, J. Raes, Microbial interactions: From networks to models. *Nat. Rev. Microbiol.* **10**, 538–550 (2012). doi:10.1038/nrmicro2832 Medline

39. D. Rios Garza, D. Gonze, H. Zafeiropoulos, B. Liu, K. Faust, Metabolic models of human gut microbiota: Advances and challenges. *Cell Syst.* **14**, 109–121 (2023). doi:10.1016/j.cels.2022.11.002 Medline

40. N. I. van den Berg, D. Machado, S. Santos, I. Rocha, J. Chacón, W. Harcombe, S. Mitri, K. R. Patil, Ecological modelling approaches for predicting emergent properties in microbial communities. *Nat. Ecol. Evol.* **6**, 855–865 (2022). doi:10.1038/s41559-022-01746-7 Medline

41. Y.-Y. Liu, Controlling the human microbiome. *Cell Syst.* **14**, 135–159 (2023). doi:10.1016/j.cels.2022.12.010 Medline

42. C. G. Buffie, V. Bucci, R. R. Stein, P. T. McKenney, L. Ling, A. Gobourne, D. No, H. Liu, M. Kinnebrew, A. Viale, E. Littmann, M. R. M. van den Brink, R. R. Jenq, Y. Taur, C. Sander, J. R. Cross, N. C. Toussaint, J. B. Xavier, E. G. Pamer, Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. *Nature* **517**, 205–208 (2015). doi:10.1038/nature13828 Medline

43. T. S. B. Schmidt, S. S. Li, O. M. Maistrenko, W. Akanni, L. P. Coelho, S. Dolai, A. Fullam, A. M. Glazek, R. Hercog, H. Herrema, F. Jung, S. Kandels, A. Orakov, R. Thielemann, M. von Stetten, T. Van Rossum, V. Benes, T. J. Borody, W. M. de Vos, C. Y. Ponsioen, M. Nieuwdorp, P. Bork, Drivers and determinants of strain dynamics following fecal microbiota transplantation. *Nat. Med.* **28**, 1902–1912 (2022). doi:10.1038/s41591-022-01913-0 Medline

44. G. Aguadé-Gorgorió, J. F. Arnoldi, M. Barbier, S. Kéfi, A taxonomy of multiple stable states in complex ecological communities. *Ecol. Lett.* **27**, e14413 (2024). doi:10.1111/ele.14413 Medline

45. Materials and methods are available as supplementary materials.

46. S. Krishnan, N. Alden, K. Lee, Pathways and functions of gut microbiota metabolism impacting host physiology. *Curr. Opin. Biotechnol.* **36**, 137–145 (2015). doi:10.1016/j.copbio.2015.08.015 Medline

47. K. Oliphant, E. Allen-Vercoe, Macronutrient metabolism by the human gut microbiome: Major fermentation by-products and their impact on host health. *Microbiome* **7**, 91 (2019). doi:10.1186/s40168-019-0704-8 Medline

48. K. A. Krautkramer, J. Fan, F. Bäckhed, Gut microbial metabolites as multi-kingdom intermediates. *Nat. Rev. Microbiol.* **19**, 77–94 (2021). doi:10.1038/s41579-020-0438-4 Medline

49. I. Rowland, G. Gibson, A. Heinken, K. Scott, J. Swann, I. Thiele, K. Tuohy, Gut microbiota functions: Metabolism of nutrients and other food components. *Eur. J. Nutr.* **57**, 1–24 (2018). doi:10.1007/s00394-017-1445-8 Medline

50. Z.-N. Ling, Y.-F. Jiang, J.-N. Ru, J.-H. Lu, B. Ding, J. Wu, Amino acid metabolism in health and disease. *Signal Transduct. Target. Ther.* **8**, 345 (2023). doi:10.1038/s41392-023-01569-3 Medline

51. J. Grilli, Macroecological laws describe variation and diversity in microbial communities. *Nat. Commun.* **11**, 4743 (2020). doi:10.1038/s41467-020-18529-y Medline

52. P.-Y. Ho, B. H. Good, K. C. Huang, Competition for fluctuating resources reproduces statistics of species abundance over time across wide-ranging microbiotas. *eLife* **11**, e75168 (2022). doi:10.7554/eLife.75168 Medline

53. S. Louca, M. F. Polz, F. Mazel, M. B. N. Albright, J. A. Huber, M. I. O'Connor, M. Ackermann, A. S. Hahn, D. S. Srivastava, S. A. Crowe, M. Doebeli, L. W. Parfrey, Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.* **2**, 936–943 (2018). doi:10.1038/s41559-018-0519-1 Medline

54. R. L. Tatusov, M. Y. Galperin, D. A. Natale, E. V. Koonin, The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**, 33–36 (2000). doi:10.1093/nar/28.1.33 Medline

55. H. Ogata, S. Goto, K. Sato, W. Fujibuchi, H. Bono, M. Kanehisa, KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **27**, 29–34 (1999). doi:10.1093/nar/27.1.29

56. G. Bunin, Ecological communities with Lotka-Volterra dynamics. *Phys. Rev. E* **95**, 042414 (2017). doi:10.1103/PhysRevE.95.042414 Medline

57. E. Mallmin, A. Traulsen, S. De Monte, Chaotic turnover of rare and abundant species in a strongly interacting model community. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2312822121 (2024). doi:10.1073/pnas.2312822121 Medline

58. I. Dalmedigos, G. Bunin, Dynamical persistence in high-diversity resource-consumer communities. *PLOS Comput. Biol.* **16**, e1008189 (2020). doi:10.1371/journal.pcbi.1008189 Medline

59. E. Blumenthal, J. W. Rocks, P. Mehta, Phase Transition to Chaos in Complex Ecosystems with Nonreciprocal Species-Resource Interactions. *Phys. Rev. Lett.* **132**, 127401 (2024). doi:10.1103/PhysRevLett.132.127401 Medline

60. M. Schirmer, A. Garner, H. Vlamakis, R. J. Xavier, Microbial genes and pathways in inflammatory bowel disease. *Nat. Rev. Microbiol.* **17**, 497–511 (2019). doi:10.1038/s41579-019-0213-6 Medline

61. E. J. Stevens, K. A. Bates, K. C. King, Host microbiota can facilitate pathogen infection. *PLOS Pathog.* **17**, e1009514 (2021). doi:10.1371/journal.ppat.1009514 Medline

62. J. Lloyd-Price, C. Arze, A. N. Ananthakrishnan, M. Schirmer, J. Avila-Pacheco, T. W. Poon, E. Andrews, N. J. Ajami, K. S. Bonham, C. J. Brislawn, D. Casero, H. Courtney, A. Gonzalez, T. G. Graeber, A. B. Hall, K. Lake, C. J. Landers, H. Mallick, D. R. Plichta, M. Prasad, G. Rahnavard, J. Sauk, D. Shungin, Y. Vázquez-Baeza, R. A. White 3rd, IBDMDB Investigators, J. Braun, L. A. Denson, J. K. Jansson, R. Knight, S. Kugathasan, D. P. B. McGovern, J. F. Petrosino, T. S. Stappenbeck, H. S. Winter, C. B. Clish, E. A. Franzosa, H. Vlamakis, R. J. Xavier, C. Huttenhower, Multi-omics of the gut microbial

ecosystem in inflammatory bowel diseases. *Nature* **569**, 655–662 (2019). doi:10.1038/s41586-019-1237-9 Medline

63. K. Z. Coyte, J. Schluter, K. R. Foster, The ecology of the microbiome: Networks, competition, and stability. *Science* **350**, 663–666 (2015). doi:10.1126/science.aad2602 Medline

64. J. D. Palmer, K. R. Foster, Bacterial species rarely work together. *Science* **376**, 581–582 (2022). doi:10.1126/science.abn5093 Medline

65. K. R. Foster, T. Bell, Competition, not cooperation, dominates interactions among culturable microbial species. *Curr. Biol.* **22**, 1845–1850 (2012). doi:10.1016/j.cub.2012.08.005 Medline

66. O. S. Venturelli, A. C. Carr, G. Fisher, R. H. Hsu, R. Lau, B. P. Bowen, S. Hromada, T. Northen, A. P. Arkin, Deciphering microbial interactions in synthetic human gut microbiome communities. *Mol. Syst. Biol.* **14**, e8157 (2018). doi:10.15252/msb.20178157 Medline

67. W. Lopes, D. R. Amor, J. Gore, Cooperative growth in microbial communities is a driver of multistability. *Nat. Commun.* **15**, 4709 (2024). doi:10.1038/s41467-024-48521-9 Medline

68. D. Machado, O. M. Maistrenko, S. Andrejev, Y. Kim, P. Bork, K. R. Patil, K. R. Patil, Polarization of microbial communities between competitive and cooperative metabolism. *Nat. Ecol. Evol.* **5**, 195–203 (2021). doi:10.1038/s41559-020-01353-4 Medline

69. J. Camacho-Mateu, A. Lampo, M. Sireci, M. A. Muñoz, J. A. Cuesta, Sparse species interactions reproduce abundance correlation patterns in microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2309575121 (2024). doi:10.1073/pnas.2309575121 Medline

70. A. Goyal, L. S. Bittleston, G. E. Leventhal, L. Lu, O. X. Cordero, Interactions between strains govern the eco-evolutionary dynamics of microbial communities. *eLife* **11**, e74987 (2022). doi:10.7554/eLife.74987 Medline

71. C.-Y. Chang, D. Bajić, J. C. C. Vila, S. Estrela, A. Sanchez, Emergent coexistence in multispecies microbial communities. *Science* **381**, 343–348 (2023). doi:10.1126/science.adg0727 Medline

72. K. Z. Coyte, S. Rakoff-Nahoum, Understanding Competition and Cooperation within the Mammalian Gut Microbiome. *Curr. Biol.* **29**, R538–R544 (2019). doi:10.1016/j.cub.2019.04.017 Medline

73. P.-Y. Ho, T. H. Nguyen, J. M. Sanchez, B. C. DeFelice, K. C. Huang, Resource competition predicts assembly of gut bacterial communities in vitro. *Nat. Microbiol.* **9**, 1036–1048 (2024). doi:10.1038/s41564-024-01625-w Medline

74. E. J. Culp, A. L. Goodman, Cross-feeding in the gut microbiome: Ecology and mechanisms. *Cell Host Microbe* **31**, 485–499 (2023). doi:10.1016/j.chom.2023.03.016 Medline

75. A. Goyal, T. Wang, V. Dubinkina, S. Maslov, Ecology-guided prediction of cross-feeding interactions in the human gut microbiome. *Nat. Commun.* **12**, 1335 (2021). doi:10.1038/s41467-021-21586-6 Medline

76. V. R. Marcelino, C. Welsh, C. Diener, E. L. Gulliver, E. L. Rutten, R. B. Young, E. M. Giles, S. M. Gibbons, C. Greening, S. C. Forster, Disease-specific loss of microbial cross-feeding interactions in the human gut. *Nat. Commun.* **14**, 6546 (2023). doi:10.1038/s41467-023-42112-w Medline

77. A. R. Watson, J. Füssel, I. Veseli, J. Z. DeLongchamp, M. Silva, F. Trigodet, K. Lolans, A. Shaiber, E. Fogarty, J. M. Runde, C. Quince, M. K. Yu, A. Söylev, H. G. Morrison, S. T. M. Lee, D. Kao, D. T. Rubin, B. Jabri, T. Louie, A. M. Eren, Metabolic independence drives gut microbial colonization and resilience in health and disease. *Genome Biol.* **24**, 78 (2023). doi:10.1186/s13059-023-02924-x Medline

78. J. Tackmann, J. F. Matias Rodrigues, C. von Mering, Rapid Inference of Direct Interactions in Large-Scale Ecological Networks from Heterogeneous Microbial Sequencing Data. *Cell Syst.* **9**, 286–296.e8 (2019). doi:10.1016/j.cels.2019.08.002 Medline

79. C. Li, T. V. Av-Shalom, J. W. G. Tan, J. S. Kwah, K. R. Chng, N. Nagarajan, BEEM-Static: Accurate inference of ecological interactions from cross-sectional microbiome data. *PLOS Comput. Biol.* **17**, e1009343 (2021). doi:10.1371/journal.pcbi.1009343 Medline

80. C. K. Fisher, P. Mehta, Identifying keystone species in the human gut microbiome from metagenomic timeseries using sparse linear regression. *PLOS ONE* **9**, e102451 (2014). doi:10.1371/journal.pone.0102451 Medline

81. S. S. Gude, R. S. Veeravalli, B. Vejandla, S. S. Gude, T. Venigalla, V. Chintagumpala, Colorectal Cancer Diagnostic Methods: The Present and Future. *Cureus* **15**, e37622 (2023). doi:10.7759/cureus.37622 Medline

82. S. D. Jurburg, S. A. Blowes, A. Shade, N. Eisenhauer, J. M. Chase, Synthesis of recovery patterns in microbial communities across environments. *Microbiome* **12**, 79 (2024). doi:10.1186/s40168-024-01802-3 Medline

83. S. Weiss, W. Van Treuren, C. Lozupone, K. Faust, J. Friedman, Y. Deng, L. C. Xia, Z. Z. Xu, L. Ursell, E. J. Alm, A. Birmingham, J. A. Cram, J. A. Fuhrman, J. Raes, F. Sun, J. Zhou, R. Knight, Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *ISME J.* **10**, 1669–1681 (2016). doi:10.1038/ismej.2015.235 Medline

84. A. J. Yarur, A. Bruss, A. Moosreiner, P. Beniwal-Patel, L. Nunez, B. Berens, J. F. Colombel, S. R. Targan, C. Fox, G. Y. Melmed, M. T. Abreu, P. Deepak, Higher Intra-Abdominal Visceral Adipose Tissue Mass Is Associated With Lower Rates of Clinical and Endoscopic Remission in Patients With Inflammatory Bowel Diseases Initiating Biologic Therapy: Results of the Constellation Study. *Gastroenterology* **165**, 963–975.e5 (2023). doi:10.1053/j.gastro.2023.06.036 Medline

85. L. Ning, Y.-L. Zhou, H. Sun, Y. Zhang, C. Shen, Z. Wang, B. Xuan, Y. Zhao, Y. Ma, Y. Yan, T. Tong, X. Huang, M. Hu, X. Zhu, J. Ding, Y. Zhang, Z. Cui, J.-Y. Fang, H. Chen, J. Hong, Microbiome and metabolome features in inflammatory bowel disease via multi-omics integration analyses across cohorts. *Nat. Commun.* **14**, 7135 (2023). doi:10.1038/s41467-023-42788-0 Medline

86. Human Microbiome Project, The Inflammatory Bowel Disease Multiomics Database; https://ibdmdb.org/.

87. M. Berkell, M. Mysara, B. B. Xavier, C. H. van Werkhoven, P. Monsieurs, C. Lammens, A. Ducher, M. J. G. T. Vehreschild, H. Goossens, J. de Gunzburg, M. J. M. Bonten, S. Malhotra-Kumar, ANTICIPATE study group, Microbiota-based markers predictive of development of *Clostridioides difficile* infection. *Nat. Commun.* **12**, 2241 (2021). doi:10.1038/s41467-021-22302-0 Medline

88. M. J. T. Crobach, Q. R. Ducarmon, E. M. Terveer, C. Harmanus, I. M. J. G. Sanders, K. M. Verduin, E. J. Kuijper, R. D. Zwittink, The Bacterial Gut Microbiota of Adult Patients Infected, Colonized or Noncolonized by *Clostridioides difficile*. *Microorganisms* **8**, 677 (2020). doi:10.3390/microorganisms8050677 Medline

89. P. Ferretti, J. Wirbel, O. M. Maistrenko, T. Van Rossum, R. Alves, A. Fullam, W. Akanni, C. Schudoma, A. Schwarz, R. Thielemann, L. Thomas, S. Kandels, R. Hercog, A. Telzerow, I. Letunic, M. Kuhn, G. Zeller, T. S. B. Schmidt, P. Bork, *C. difficile* may be overdiagnosed in adults and is a prevalent commensal in infants. *eLife* **12**, RP90111 (2023). doi:10.7554/eLife.90111.1

90. R. A. T. Mars, Y. Yang, T. Ward, M. Houtti, S. Priya, H. R. Lekatz, X. Tang, Z. Sun, K. R. Kalari, T. Korem, Y. Bhattarai, T. Zheng, N. Bar, G. Frost, A. J. Johnson, W. van Treuren, S. Han, T. Ordog, M. Grover, J. Sonnenburg, M. D'Amato, M. Camilleri, E. Elinav, E. Segal, R. Blekhman, G. Farrugia, J. R. Swann, D. Knights, P. C. Kashyap, Longitudinal Multi-omics Reveals Subset-Specific Mechanisms Underlying Irritable Bowel Syndrome. *Cell* **182**, 1460–1473.e17 (2020). doi:10.1016/j.cell.2020.08.007 Medline

91. G.-H. Kim, K. Lee, J. O. Shim, Gut Bacterial Dysbiosis in Irritable Bowel Syndrome: A Case-Control Study and a Cross-Cohort Analysis Using Publicly Available Data Sets. *Microbiol. Spectr.* **11**, e0212522 (2023). doi:10.1128/spectrum.02125-22 Medline

92. Z. A. Barandouzi, J. Lee, M. Del Carmen Rosas, J. Chen, W. A. Henderson, A. R. Starkweather, X. S. Cong, Associations of neurotransmitters and the gut microbiome with emotional distress in mixed type of irritable bowel syndrome. *Sci. Rep.* **12**, 1648 (2022). doi:10.1038/s41598-022-05756-0 Medline

93. G. Zeller, J. Tap, A. Y. Voigt, S. Sunagawa, J. R. Kultima, P. I. Costea, A. Amiot, J. Böhm, F. Brunetti, N. Habermann, R. Hercog, M. Koch, A. Luciani, D. R. Mende, M. A. Schneider, P. Schrotz-King, C. Tournigand, J. Tran Van Nhieu, T. Yamada, J. Zimmermann, V. Benes, M. Kloor, C. M. Ulrich, M. von Knebel Doeberitz, I. Sobhani, P. Bork, Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol. Syst. Biol.* **10**, 766 (2014). doi:10.15252/msb.20145645 Medline

94. S. Yachida, S. Mizutani, H. Shiroma, S. Shiba, T. Nakajima, T. Sakamoto, H. Watanabe, K. Masuda, Y. Nishimoto, M. Kubo, F. Hosoda, H. Rokutan, M. Matsumoto, H. Takamaru, M. Yamada, T. Matsuda, M. Iwasaki, T. Yamaji, T. Yachida, T. Soga, K. Kurokawa, A. Toyoda, Y. Ogura, T. Hayashi, M. Hatakeyama, H. Nakagama, Y. Saito, S. Fukuda, T. Shibata, T. Yamada, Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat. Med.* **25**, 968–976 (2019). doi:10.1038/s41591-019-0458-7 Medline

95. E. Vogtmann, X. Hua, G. Zeller, S. Sunagawa, A. Y. Voigt, R. Hercog, J. J. Goedert, J. Shi, P. Bork, R. Sinha, Colorectal Cancer and the Human Gut Microbiome: Reproducibility

with Whole-Genome Shotgun Sequencing. *PLOS ONE* **11**, e0155362 (2016). doi:10.1371/journal.pone.0155362 Medline

96. R. Corral Lopez, Code for article: Imbalance in gut microbial interactions as a marker of health and disease, Zenodo (2025). doi:10.5281/zenodo.15202034

97. K. Meissel, E. S. Yao, Using Cliff's Delta as a Non-Parametric Effect Size Measure: An Accessible Web App and R Tutorial. *Pract. Assess. Res. Eval.* **29**, 2 (2024). doi:10.7275/pare.1977

98. A. Carr, C. Diener, N. S. Baliga, S. M. Gibbons, Use and abuse of correlation analyses in microbial ecology. *ISME J.* **13**, 2647–2655 (2019). doi:10.1038/s41396-019-0459-z Medline

99. K. Faust, Open challenges for microbial network construction and analysis. *ISME J.* **15**, 3111–3118 (2021). doi:10.1038/s41396-021-01027-4 Medline

100. D. Vandeputte, L. De Commer, R. Y. Tito, G. Kathagen, J. Sabino, S. Vermeire, K. Faust, J. Raes, Temporal variability in quantitative human gut microbiome profiles and implications for clinical research. *Nat. Commun.* **12**, 6740 (2021). doi:10.1038/s41467-021-27098-7 Medline

101. P. D. Karp, R. Billington, R. Caspi, C. A. Fulcher, M. Latendresse, A. Kothari, I. M. Keseler, M. Krummenacker, P. E. Midford, Q. Ong, W. K. Ong, S. M. Paley, P. Subhraveti, The BioCyc collection of microbial genomes and metabolic pathways. *Brief. Bioinform.* **20**, 1085–1093 (2019). doi:10.1093/bib/bbx085 Medline

102. J. P. Jacobs, V. Lagishetty, M. C. Hauer, J. S. Labus, T. S. Dong, R. Toma, M. Vuyisich, B. D. Naliboff, J. M. Lackner, A. Gupta, K. Tillisch, E. A. Mayer, Multi-omics profiles of the intestinal microbiome in irritable bowel syndrome and its bowel habit subtypes. *Microbiome* **11**, 5 (2023). doi:10.1186/s40168-022-01450-5 Medline

103. J. G. Caporaso, C. L. Lauber, E. K. Costello, D. Berg-Lyons, A. Gonzalez, J. Stombaugh, D. Knights, P. Gajer, J. Ravel, N. Fierer, J. I. Gordon, R. Knight, Moving pictures of the human microbiome. *Genome Biol.* **12**, R50 (2011). doi:10.1186/gb-2011-12-5-r50 Medline

104. J. Park, K. Kato, H. Murakami, K. Hosomi, K. Tanisawa, T. Nakagata, H. Ohno, K. Konishi, H. Kawashima, Y.-A. Chen, A. Mohsen, J. Z. Xiao, T. Odamaki, J. Kunisawa, K. Mizuguchi, M. Miyachi, Comprehensive analysis of gut microbiota of a healthy population and covariates affecting microbial variation in two large Japanese cohorts. *BMC Microbiol.* **21**, 151 (2021). doi:10.1186/s12866-021-02215-0 Medline

105. Y. Goto, Y. Nishimoto, S. Murakami, T. Nomaguchi, Y. Mori, M. Ito, R. Nakaguro, T. Kudo, T. Matsuoka, T. Yamada, T. Kobayashi, S. Fukuda, Metabologenomic Approach Reveals Intestinal Environmental Features Associated with Barley-Induced Glucose Tolerance Improvements in Japanese: A Randomized Controlled Trial. *Nutrients* **14**, 3468 (2022). doi:10.3390/nu14173468 Medline

106. J. B. Russell, G. M. Cook, Energetics of bacterial growth: Balance of anabolic and catabolic reactions. *Microbiol. Rev.* **59**, 48–62 (1995). doi:10.1128/mr.59.1.48-62.1995 Medline

107. K. Adamberg, G. Raba, S. Adamberg, Use of Changestat for Growth Rate Studies of Gut Microbiota. *Front. Bioeng. Biotechnol.* **8**, 24 (2020). doi:10.3389/fbioe.2020.00024 Medline

108. M. H. Stipanuk, M. A. Caudill, *Biochemical, Physiological, and Molecular Aspects of Human Nutrition* (Elsevier Health Sciences, 2013).

109. J. W. Fink, N. A. Held, M. Manhart, Microbial population dynamics decouple growth response from environmental nutrient concentration. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2207295120 (2023). doi:10.1073/pnas.2207295120 Medline

110. R. Sender, S. Fuchs, R. Milo, Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* **164**, 337–340 (2016). doi:10.1016/j.cell.2016.01.013 Medline

111. D. T. Truong, E. A. Franzosa, T. L. Tickle, M. Scholz, G. Weingart, E. Pasolli, A. Tett, C. Huttenhower, N. Segata, MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* **12**, 902–903 (2015). doi:10.1038/nmeth.3589 Medline

112. S. Sunagawa, D. R. Mende, G. Zeller, F. Izquierdo-Carrasco, S. A. Berger, J. R. Kultima, L. P. Coelho, M. Arumugam, J. Tap, H. B. Nielsen, S. Rasmussen, S. Brunak, O. Pedersen, F. Guarner, W. M. de Vos, J. Wang, J. Li, J. Doré, S. D. Ehrlich, A. Stamatakis, P. Bork, Metagenomic species profiling using universal phylogenetic marker genes. *Nat. Methods* **10**, 1196–1199 (2013). doi:10.1038/nmeth.2693 Medline