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,*,†]

[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 authors. Email: mamunoz@onsager.ugr.es and juan.bonachela@rutgers.edu

†These authors contributed equally to this work.

April 30, 2025

Imbalances in the human gut microbiome (dysbioses) are linked to multiple diseases but remain poorly understood. Current biomarkers to identify dysbiosis are inconsistent and fail to capture the ecological mechanisms differentiating healthy from diseased microbiomes. We propose a general dysbiosis biomarker, inspired by phenomenology observed in a gut-microbiome theoretical model introduced here. The emergent communities show complex interaction networks and

two distinct collective states, corresponding to healthy and dysbiotic microbiomes. Our robust metric for dysbiosis, by quantifying the balance between cooperation and competition, differentiates these states in both simulated and real datasets across diverse diseases. Moreover, it reveals that dysbiosis results from a shift toward greater cooperation in the community. Our metric further correlates with disease progression, highlighting its potential as a diagnostic tool.

The human gut microbiome is a dynamic and complex system that plays a crucial role in maintaining health through its interactions with the host and environment (*1–6*). Innovations in profiling techniques over the past two decades, have greatly improved our ability to quantify both microbiome composition and functional capabilities (*7–9*). These studies have linked imbalances in the gut microbiome, or dysbioses, to numerous 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, among others (*1, 2, 7, 10–16*). Consequently, novel treatments such as fecal microbiota transplantation, diet-based interventions, and probiotics (*17–19*), aim to restore microbial balance and reverse dysbiosis by altering community composition to eliminate such imbalances. However, these treatments show highly variable and unpredictable success rates, which slows clinical implementation. A key reason for this unpredictability is the limited understanding of the mechanisms underlying gut microbiome dynamics and structure. Particularly, the metabolic and ecological interactions among microorganisms, and their relationships with the host and environment, remain poorly understood (*1, 14, 20, 21*). This knowledge gap has hindered the identification of robust microbiome biomarkers that correlate with, and accurately differentiate, health and disease. Our work seeks to address both of these fundamental challenges.

1 The search for a universal indicator of dysbiosis

Dysbiosis is a broad concept that encompasses various compositional and functional attributes across diseases, which hinders its unambiguous characterization (*12*). Early efforts focused on identifying single bacterial species or functions whose altered levels could correlate with a particular disease (*10, 12, 22, 23*). However, it soon became clear that dysbiosis is a collective state, typically

affecting multiple species and functions, which can only be captured by systemic, whole-community approaches (*16, 24*). A common claim across studies and conditions is that disease correlates with reduced microbial diversity, often suggested as a universal marker of dysbiosis (*7, 13, 15, 25–27*). We re-examined empirical data from several diseases, and our analysis reveals a more nuanced, disease-specific pattern (Fig. 1), in line with recent studies (*12, 16, 24, 28*). Specifically, while IBD and CDI (Fig. 1A-B) consistently show reduced diversity compared to healthy microbiomes (*16, 29–31*), conditions like IBS and CRC (Fig. 1C-D) yield mixed results, with decreases, increases, or no statistically significant change in diversity (*16, 32–37*). These findings clearly refute that diversity universally declines in diseased states, and raise the fundamental question of whether there are conserved mechanisms driving dysbiosis, which therefore would enable the identification of robust disease indicators.

The dichotomy between health and disease has fuelled the extensively explored hypothesis that dysbiosis generally emerges from shifts between alternative stable states, that is, between distinct taxonomic or functional configurations of the gut microbiome that can occur for identical conditions (*50–62*). However, this idea remains under debate due to the varied methodologies used to define and measure such states, the context-dependent nature of the microbiome, and the challenges of controlling external factors such as diet, medication, and lifestyle (*7, 10, 63*). All these factors influence gut composition and function, complicating the identification of stable states and consistent disease markers.

The challenges above, however, do not apply to theoretical models, where definitions and environmental factors can be fine-tuned (*21, 63–67*). For example, models have identified microbial taxa capable of inhibiting *C. difficile* (CD), a prediction subsequently validated experimentally to confirm that the suggested taxa indeed conferred CD colonization resistance (*68*). As another example, machine learning models have systematically tested competing hypotheses regarding the success of FMT (*69*). Despite these advances, few models designed to study gut microbiome dynamics predict or can even generate alternative stable states, limiting our understanding of their emergence and fundamental properties (*64, 70*). Therefore, there is need for novel theoretical models that can capture gut microbiome dynamics and also reveal the ecological mechanisms driving the formation of diverse collective states.

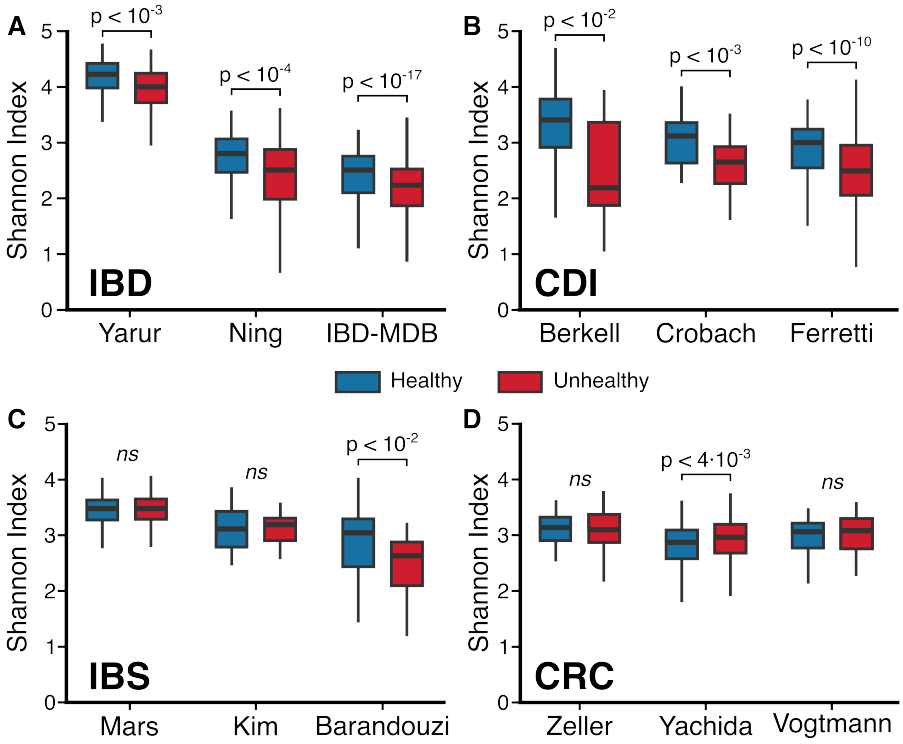

Figure 1: **α-diversity in healthy and unhealthy individuals.** Comparison of the Shannon index (see Methods) obtained using existing metagenomic analyses from healthy individuals, in blue, whereas the red color represents individuals affected by : (**A**) inflammatory bowel disease (IBD), using data from Yarur et al., Ning et al., and the Human Microbiome Project (*38–40*); (**B**) *C. difficile* infection (CDI), using data from Berkell et al., Crobach et al., and Ferreti et al. (*41–43*); (**C**) irritable bowel syndrome (IBS), using data from Mars et al., Kim et al., and Barandouzi et al. (*44–46*); and (**D**) colorectal cancer (CRC), using data from Zeller et al., Yachida et al., and Vogtmann et al. (*37, 47–49*). A significant decrease in diversity is evident for IBD and CDI, but not for IBS or CRC (*ns* = not significant).

2 A new model for the dynamics of the gut microbiome

We introduce here a mathematical model specifically designed to explore the ecological dynamics of the gut microbiome (Fig.2A and Supplementary Text). Our consumer-resource model captures bacterial and nutrient dynamics and incorporates the possibility of cross-feeding as new metabolites are produced by different bacterial strains. 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 time.

A key strength of our framework is its emphasis on the crucial role that metabolic pathways play in shaping gut communities (*71–79*). To this end, our model incorporates metabolic pathways that define all possible metabolic routes between nutrients and metabolites (hereon, substances), along with the metabolic preferences specific to each bacterial strain. While we explicitly represent only catabolic reactions for simplicity, anabolism is indirectly accounted for through a trade-off function linking enzymatic cost to energy invested in constructive metabolism. 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 1) the free energy ($E_{\alpha\beta} > 0$) released during the process; and 2) the enzymatic cost ($C_{\alpha\beta} > 0$), reflecting the number and complexity of enzymes involved. Each bacterial type i is defined by a unique pathway matrix, $P^i_{\alpha\beta}$ (binary matrix with the metabolic routes available to that bacterial type) and a coefficient γ_i that controls the fraction of free energy used for biomass growth. See Supplementary Text for a detailed description of the model, including the definition of species, nutrients, metabolites, the implementation of bacteria-specific metabolic preferences, pathways, trade-off function, and chosen parametrization. We examine the limitations associated with the modeling assumptions above in Discussion.

Our model successfully reproduces, without the need to rely on parameter fitting, well-known macroecological patterns of the gut microbiome that are observed empirically across multiple systems. Using a biologically plausible parametrization (see Supplementary Text), the 95% confidence interval obtained across replicates of the model captures the distributions of species abundance mean and fluctuations, Taylor's law (i.e. relationship between mean and variance), species prevalence (the distribution of the fraction of samples in which a species is present), and the dependence of preva-

lence on mean abundance (Supplementary Fig.S1) (*80, 81*). Additionally, Supplementary Fig. S2 shows that our model replicates the functional redundancy observed in the gut microbiome, where multiple taxonomic compositions encode the same metabolic functions (*82–85*). In our framework, functions are represented by metabolites grouped into coarse-grained energy levels (see Methods (*86*)). These broader functional categories are akin to Clusters of Orthologous Genes (COG) or Kyoto Encyclopedia of Genes and Genomes (KEGG) classifications for empirical data (*87, 88*). Functional redundancy becomes apparent at sufficiently coarse-grained resolutions for both our model and observed data. All the results above thus validate our modeling approach.

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

As shown in Figure 2B, the dynamics of our simulated gut microbiome community suggest two qualitatively distinct classes of states: one characterized by the rapid turnover of many bacterial types; and another one where only a few types dominate with quasi-stationary abundances (see definition in Methods (*86*)). Let us label these two classes of states as *healthy* and *dysbiotic*, respectively, for reasons that will become apparent in the comparison with real data below. 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 parametrizations (see Supplementary Fig. S3 and S4).

Existing community-level theoretical models, such as the generalized Lotka-Volterra (GLV) (*89–91*) and consumer-resource models with random interactions (*92, 93*), have reported dynamics reminiscent of either one of our two states (often referred to as ”chaotic fluctuations” and ”stable fixed points” in the literature). However, these models do not capture multistability between the different states. The fact that our model captures diverse alternative states and dynamic transitions between them highlights the importance of explicitly incorporating metabolic pathways, and of letting the structure of the community emerge from the dynamic interaction in the system, elements overlooked in previous theoretical efforts. As shown in Fig. 2C-D, the dysbiotic state exhibits a significant reduction in α-diversity (measured by the Shannon Index), the number of enzymes (total

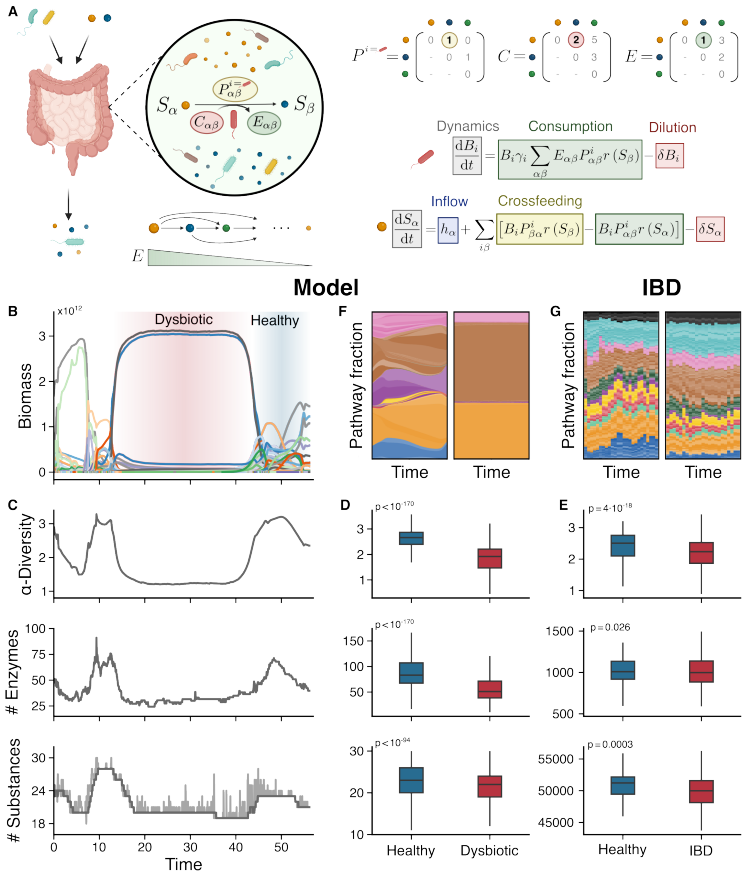

Figure 2: **Gut microbiome trait-based consumer-resource model and the spontaneous emergence of alternative stable states.** (**A**) Schematic representation of the model. The gut microbiome is conceptualized as living in a chemostat-like environment that includes both bacterial species and substances. 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 realization highlighted in the larger panel B, calculated α-diversity, number of enzymes, and number of substances over time for a single realization of the model; dark grey represents rolling average, and light grey the actual values. (**D**) Comparison, for the same metrics in panel C, of healthy and dysbiotic states. (**E**) Corresponding metrics 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)** Similary plot for healthy/IBD, where colors represent (coarse-grain) functions, and shades the various paths able to perform the given function. See Methods for details (*86*).

enzymatic cost of all active pathways), and the number of secreted metabolites (substances with non-zero concentrations) compared to the healthy state; see Methods for definitions (*86*). These findings suggest that the microbial community in the dysbiotic state operates with greater efficiency than the community in the healthy state, as fewer bacterial types and metabolic pathways are required to form consortia able to convert a larger proportion of the available energy into microbial biomass (Supplementary Fig. S5).

Given the dichotomy between these two distinct states, reminiscent of the healthy vs dysbiotic gut microbiome, we also measured the biomarkers above in healthy individuals and patients with IBD. As Fig. 2E shows, the resulting pattern mirrors our model predictions, thereby supporting our designation of the two alternative states as *healthy* and *dysbiotic*. The parallelism of model states with real healthy/dysbiotic states is reinforced by the observation that, from a qualitative perspective, both 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, while others become overrepresented, see Fig. 2F,G) (*94, 95*). Moreover, for our model as well as a range of diseases, β-diversity across healthy states was significantly lower than for dysbiotic states (Supplementary Fig. S6). These patterns, however, are subtle or inconsistent (see, for example, β-diversity for CRC, Supplementary Fig. S6D, or number of metabolites and diversity of secreted metabolites for IBS and CRC, Supplementary Fig. S7). Beyond this lack of consistency, these biomarkers merely reflect the consequences of the shift from one state to the other and therefore do not explain the mechanisms underlying the transitions, nor can be used as early-warning indicators. Thus, we further explored with our model the differences between healthy and dysbiotic states, aiming to develop a robust quantitative metric able to reliably distinguish between healthy and dysbiotic states and that can also reveal the mechanisms driving the emergence of these states and transitions.

4 The balance of community interactions is disrupted in dysbiotic states and disease

In our model, bacterial interactions are not assumed but rather 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 (*64, 96, 97*).

There is currently no consensus on whether competition or cooperation primarily drive gut microbiota dynamics. While some empirical studies suggest that competitive interactions dominate due to their role in promoting stability (*21, 98, 99*), others emphasize the critical role of cross-feeding interactions in maintaining a healthy microbial community (*100–103*). Thus, the relative contributions of these interactions to gut microbiota function in health and disease remain unclear.

To quantify this relative contribution, we introduce a *net interaction* metric, ρ, defined as the difference between the total cross-feeding and competition interactions within the community, normalized by the sum of all interaction strengths (see Eq. S8 in Methods (*86*)). Since this metric is normalized, the information it provides remains independent of community size and, consequently, diversity.

Our analysis reveals that the healthy state in the model is characterized by a negative net interaction (i.e. community dominated by competition), whereas the dysbiotic state is characterized by positive net interactions (cooperation-dominated community, Fig. 3A), which we use to rigorously identify the two states in the model (see Methods). This observation leads us to hypothesize that in IBD and potentially in other diseases, the gut microbiome may shift toward greater cooperation compared to the dynamics seen in healthy controls. While the net interaction metric robustly captures this transition, diversity metrics such as the Shannon index are broadly distributed for any given value of ρ (Supplementary Fig. S8), which may explain why it fails to consistently distinguish healthy from diseased states across conditions.

To test our transition hypothesis, we used an algorithm for empirical co-abundance networks to infer effective interaction networks (see Methods (*86*) and (*104*)) and calculated the associated net interaction metric, ρ_{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.

As Fig. 3A and Supplementary Fig. S9A show, the net interaction ρ_{inferred} inferred for model snapshots successfully captured the qualitative behavior of the true underlying network, as *i)* simulated communities from dysbiotic states showed a higher ρ_{inferred} than those from healthy states; and *ii)* dysbiotic state communities with higher true-ρ values showed higher values of ρ_{inferred}. Importantly, Fig. S9A shows that the inference method remains robust across different subsampling strategies. Building on this validation, we then applied this methodology across different diseases (IBD, CDI, IBS, and CRC). Supplementary Fig. S9B-E shows that, in all these cases, there is a significant increase in ρ_{inferred} compared to their respective healthy control groups. The qualitative results remained robust across additional network inference methods, one designed for cross-sectional data (*105*) (Supplementary Fig. S10) and another one, specifically developed for longitudinal data (*106*) (Supplementary Fig. S11). The agreement reinforces the association of dysbiosis with a shift toward a more cooperative microbial network.

With this in mind, we used the healthy/control case as a baseline for each dataset and computed the difference between the ρ_{inferred} values of the dysbiotic and control inferred networks, which we termed the "*Ecological Network Balance Index*" (ENBI) Eq. (S10). This index is negative when the community exhibits more competitive dynamics than the control (ENBI < 0) and positive when the community shows more cooperative interactions than in the control (ENBI > 0).

Across all diseases analyzed, as well as in our model, the dysbiotic state consistently exhibits a significantly positive ENBI value, reflecting a shift toward increased cooperative interactions compared to the controls (Fig. 3B-D). Furthermore, in the case of CRC, for which long-term data on disease progression is available, the ENBI increases (and thus community dynamics become increasingly cooperative compared to those in healthy individuals) as the disease advances (Supplementary Fig. S9B and Fig.3D). This consistent rise in the ENBI mirrors that observed in the model as the simulated community becomes more and more dominated by the few strains that compose the dysbiotic state.

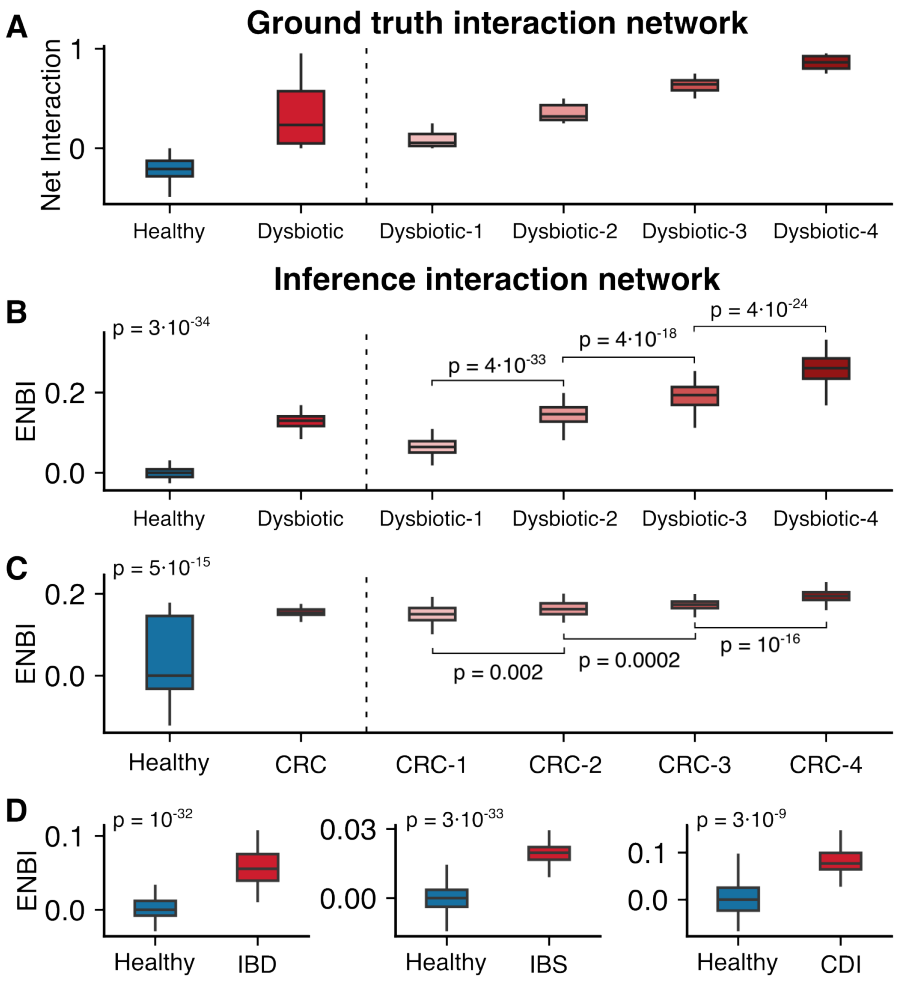

Figure 3: **The Ecological Network Balance Index (ENBI) discerns healthy from pathological communities, and correlates with disease progression.** (**A**) Net interaction calculated from the actual interaction network in the model, for healthy and dysbiotic states separately (left side of the dashed line) and different consecutive snapshots of unhealthy states (right side). For the latter, snapshots were grouped and ordered in 4 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 cooperative microbial community compared to healthy patients; the ENBI increases with disease progression (right side), supporting our hypothesis of a progressive shift towards more cooperative network structures. (**D**) ENBI for IBD, IBS, and CDI also show that diseased states tend to be more cooperative than their respective controls.

5 Discussion

Whether substantial changes in the gut microbiome cause or result from the onset of disease remains debated. Regardless of causality, our study sheds light on the ecological dynamics and community-level effects shaping the gut microbiome in health and disease. The network ecological balance index (ENBI), our newly introduced metric to quantify the dominant interaction dynamics, offers deeper and more nuanced insights into the structure of microbial communities and underlying mechanisms than existing biomarkers. Unlike traditional metrics (e.g. Shannon index, which is abundance-based), the ENBI consistently reveals a reproducible pattern across diseases and methodologies, highlighting a microbial shift toward more cooperative dynamics in dysbiotic states compared to healthy controls.

In particular, our model suggests that diseased states are associated with a shift in microbial interaction networks, by which cooperative dynamics increasingly dominate over the competitive interactions. This relative increase does not necessarily reflect an absolute rise in cross-feeding; it could instead result from a sharper decline in competitive interactions. Thus our observation is not in conflict with previous reports of reduced cross-feeding and increased metabolic self-sufficiency in dysbiosis (*102, 103, 107*). Furthermore, prior studies typically focus on specific metabolic pathways that get altered in disease, whereas our analysis assesses the overall balance between interaction types.

A detailed analysis of the resulting microbial communities indicates that this shift stems from the formation of mutualistic consortia, typically comprising a subset of species from the more heterogeneous healthy state. These consortia are more efficient at exploiting available resources and thus outcompete other species, potentially depleting metabolites, all of which may lead to the alteration or even loss of functions essential for host health. This exclusion often, but not always, results in reduced α-diversity, which explains why some diseases do not lead to a reduced diversity with respect to the control case.

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 significant clinical implications, particularly because our metric can be quantified non-invasively from stool metagenomic data. Therefore, monitoring the ENBI of a patient can help

devise early-warning indicators for disease onset, especially important for conditions such as CRC for which early detection is key to the success of the treatment and improvement of survival rates (*108*). A potential detection protocol could involve an initial ENBI calibration through regular stool samples over a set period to establish a patient-specific baseline interaction network, as part of annual health checkups. Subsequent, less-frequent sampling could then track significant ENBI deviations, signaling the emergence of dysbiosis. Even if symptoms are already present when monitoring starts, the rate of ENBI change would indicate disease progression. Nonetheless, despite the broad applicability of the ENBI, the biomarker 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 (e.g., for CRC, computed tomography colonography, endoscopy, or specific biochemical markers).

Beyond its diagnostic potential, this framework offers a tool to explore how factors such as diet, probiotics, lifestyle, antibiotics, or fecal microbiota transplantation influence the ecological interactions that shape the microbiome. In doing so, it opens up intriguing questions and research directions: Is the success of FMT in conditions like CDI linked to its ability to restore competition-driven microbial dynamics? Do probiotics work by re-establishing competitive interactions, or by preventing the emergence of dominant pathological consortia? How do diet and lifestyle contribute to shaping microbial communities? Can insights into ecological interaction networks obtained with our framework be leveraged to develop targeted interventions aiming to restore dysbiotic communities to a healthy state?

Our model follows a metabolically explicit approach but, unlike genome-scale metabolic models and kinetic models (*64*), which require detailed genome annotations and well-defined objective functions with extensive parameter fitting, our simpler framework focuses on essential ecological and metabolic principles and a biologically plausible parametrization. Our bottom-up approach is thus particularly well-suited for unveiling mechanistic underpinnings of gut microbiome emergent properties such as diversity, multistability, and metabolic organization (*65*). 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 from each other; Supplementary Fig. S6): on the one hand, healthy states show similar functional composition, even though the communities able to produce those functions may be different; on

the other hand, the loss of metabolic pathways disrupts partially or totally some of those functions in dysbiotic states, with disease data suggesting that the disruption may be disease-specific (*12*). Moreover, the model suggests that, for the loss of a function, not only the identity of the missing bacterial types matter but also the consequential changes in network structure. The subtle behavior for β-diversity observed in the model, and lack of consistency observed in our analysis of the data, agree with the lack of generality of the AKP across diseases (*109*).

Nonetheless, our framework is not without limitations. The model includes simplifications such as the indirect (as opposed to explicit) representation of anabolic reactions; the representation of host influence solely through resource input and dilution; and the reduced parametrization for energy and enzymatic cost matrices. These are in addition to other factors traditionally overlooked in models such as the dynamic interaction with phages or chemical warfare between bacteria. 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 (*63, 110, 111*). Other limitations are inherent to the heterogeneity of the methods used across studies. For example, cross-disease comparisons —which could provide valuable insights into microbiome dynamics for different conditions— are currently hindered by variability in sampling and taxonomic profiling methods across studies. Such methodological differences directly impact the inference of the ecological interaction network, making quantitative comparisons across diseases difficult. For that reason, we focused here on identifying qualitative trends rather than absolute numerical differences. Establishing standarized unified sampling and profiling methods would enable more robust cross-disease analyses in the future.

Addressing the challenges above will require a multidisciplinary effort that integrates empirical studies with theoretical approaches like the one we presented here. Connecting microbial ecology with clinical research, our framework contributes to precision medicine. By uncovering links between ecological interactions, disease progression, and therapeutic outcomes, we hope to inspire further research and more personalized strategies for maintaining or restoring gut health.

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. *Nutrition Reviews* **70**, S45–S56 (2012), doi:10.1111/j.1753-4887.2012.00505.x, http://dx.doi.org/10.1111/j.1753-4887.2012.00505.x.

2. W. M. de Vos, H. Tilg, M. Van Hul, P. D. Cani, Gut microbiome and health: mechanistic insights. *Gut* **71** (5), 1020–1032 (2022), doi:10.1136/gutjnl-2021-326789, http://dx.doi.org/10.1136/gutjnl-2021-326789.

3. C. Prosty, *et al.*, Causal role of the gut microbiome in certain human diseases: a narrative review. *eGastroenterology* **2** (3), e100086 (2024), doi:10.1136/egastro-2024-100086, http://dx.doi.org/10.1136/egastro-2024-100086.

4. J. A. Gilbert, *et al.*, Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature* **535** (7610), 94–103 (2016), doi:10.1038/nature18850, http://dx.doi.org/10.1038/nature18850.

5. S. N. Chaudhari, M. D. McCurry, A. S. Devlin, Chains of evidence from correlations to causal molecules in microbiome-linked diseases. *Nature Chemical Biology* **17** (10), 1046–1056 (2021), doi:10.1038/s41589-021-00861-z, http://dx.doi.org/10.1038/s41589-021-00861-z.

6. K. Hou, *et al.*, Microbiota in health and diseases. *Signal Transduction and Targeted Therapy* **7** (1) (2022), doi:10.1038/s41392-022-00974-4, http://dx.doi.org/10.1038/s41392-022-00974-4.

7. The Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome. *Nature* **486** (7402), 207–214 (2012), doi:10.1038/nature11234, http://dx.doi.org/10.1038/nature11234.

8. A. Zhernakova, *et al.*, Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* **352** (6285), 565–569 (2016).

9. G. Vrancken, A. C. Gregory, G. R. B. Huys, K. Faust, J. Raes, Synthetic ecology of the human gut microbiota. *Nature Reviews Microbiology* **17** (12), 754–763 (2019), doi:10.1038/s41579-019-0264-8, http://dx.doi.org/10.1038/s41579-019-0264-8.

10. I. Cho, M. J. Blaser, The human microbiome: at the interface of health and disease. *Nature Reviews Genetics* **13** (4), 260–270 (2012), doi:10.1038/nrg3182, http://dx.doi.org/10.1038/nrg3182.

11. J. Jovel, L. A. Dieleman, D. Kao, A. L. Mason, E. Wine, The human gut microbiome in health and disease. *Metagenomics* pp. 197–213 (2018).

12. Y. Fan, O. Pedersen, Gut microbiota in human metabolic health and disease. *Nature Reviews Microbiology* **19** (1), 55–71 (2020), doi:10.1038/s41579-020-0433-9, http://dx.doi.org/10.1038/s41579-020-0433-9.

13. F. Sommer, J. M. Anderson, R. Bharti, J. Raes, P. Rosenstiel, The resilience of the intestinal microbiota influences health and disease. *Nature Reviews Microbiology* **15** (10), 630–638 (2017), doi:10.1038/nrmicro.2017.58.

14. Z. Huang, *et al.*, The gut microbiome in human health and disease—Where are we and where are we going? A bibliometric analysis. *Frontiers in Microbiology* **13** (2022), doi:10.3389/fmicb.2022.1018594, http://dx.doi.org/10.3389/fmicb.2022.1018594.

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

16. 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. *Nature Communications* **8** (1) (2017), doi:10.1038/s41467-017-01973-8, http://dx.doi.org/10.1038/s41467-017-01973-8.

17. Y. Wang, S. Zhang, T. J. Borody, F. Zhang, Encyclopedia of fecal microbiota transplantation: a review of effectiveness in the treatment of 85 diseases. *Chinese Medical Jour-*

nal (2022), doi:10.1097/cm9.0000000000002339, http://dx.doi.org/10.1097/CM9.0000000000002339.

18. Q. Yang, *et al.*, Role of Dietary Nutrients in the Modulation of Gut Microbiota: A Narrative Review. *Nutrients* **12** (2), 381 (2020), doi:10.3390/nu12020381, http://dx.doi.org/10.3390/nu12020381.

19. O.-A. Petrariu, *et al.*, Role of probiotics in managing various human diseases, from oral pathology to cancer and gastrointestinal diseases. *Frontiers in Microbiology* **14** (2024), doi:10.3389/fmicb.2023.1296447, http://dx.doi.org/10.3389/fmicb.2023.1296447.

20. T. S. Schmidt, J. Raes, P. Bork, The Human Gut Microbiome: From Association to Modulation. *Cell* **172** (6), 1198–1215 (2018), doi:10.1016/j.cell.2018.02.044, http://dx.doi.org/10.1016/j.cell.2018.02.044.

21. M. Kumar, B. Ji, K. Zengler, J. Nielsen, Modelling approaches for studying the microbiome. *Nature Microbiology* **4** (8), 1253–1267 (2019), doi:10.1038/s41564-019-0491-9, http://dx.doi.org/10.1038/s41564-019-0491-9.

22. I. Sekirov, S. L. Russell, L. C. M. Antunes, B. B. Finlay, Gut Microbiota in Health and Disease. *Physiological Reviews* **90** (3), 859–904 (2010), doi:10.1152/physrev.00045.2009, http://dx.doi.org/10.1152/physrev.00045.2009.

23. A. B. Shreiner, J. Y. Kao, V. B. Young, The gut microbiome in health and in disease. *Current Opinion in Gastroenterology* **31** (1), 69–75 (2015), doi:10.1097/mog.0000000000000139, http://dx.doi.org/10.1097/MOG.0000000000000139.

24. V. K. Gupta, *et al.*, A predictive index for health status using species-level gut microbiome profiling. *Nature Communications* **11** (1) (2020), doi:10.1038/s41467-020-18476-8, http://dx.doi.org/10.1038/s41467-020-18476-8.

25. A. Mosca, M. Leclerc, J. P. Hugot, Gut Microbiota Diversity and Human Diseases: Should We Reintroduce Key Predators in Our Ecosystem? *Frontiers in Microbiology* **7** (2016), doi:10.3389/fmicb.2016.00455, http://dx.doi.org/10.3389/fmicb.2016.00455.

26. M. Kriss, K. Z. Hazleton, N. M. Nusbacher, C. G. Martin, C. A. Lozupone, Low diversity gut microbiota dysbiosis: drivers, functional implications and recovery. *Current Opinion in Microbiology* **44**, 34–40 (2018), doi:10.1016/j.mib.2018.07.003, http://dx.doi.org/10.1016/j.mib.2018.07.003.

27. C. Petersen, J. L. Round, Defining dysbiosis and its influence on host immunity and disease. *Cellular Microbiology* **16** (7), 1024–1033 (2014), doi:10.1111/cmi.12308, http://dx.doi.org/10.1111/cmi.12308.

28. Z. Ma, L. Li, N. J. Gotelli, Diversity-disease relationships and shared species analyses for human microbiome-associated diseases. *The ISME Journal* **13** (8), 1911–1919 (2019).

29. L. I. H. Abdel-Rahman, X. C. Morgan, Searching for a Consensus Among Inflammatory Bowel Disease Studies: A Systematic Meta-Analysis. *Inflammatory Bowel Diseases* **29** (1), 125–139 (2022), doi:10.1093/ibd/izac194, http://dx.doi.org/10.1093/ibd/izac194.

30. L. Aldars-García, M. Chaparro, J. P. Gisbert, Systematic Review: The Gut Microbiome and Its Potential Clinical Application in Inflammatory Bowel Disease. *Microorganisms* **9** (5), 977 (2021), doi:10.3390/microorganisms9050977, http://dx.doi.org/10.3390/microorganisms9050977.

31. E. Martinez, B. Taminiau, C. Rodriguez, G. Daube, Gut Microbiota Composition Associated with Clostridioides difficile Colonization and Infection. *Pathogens* **11** (7), 781 (2022), doi:10.3390/pathogens11070781, http://dx.doi.org/10.3390/pathogens11070781.

32. R. Pittayanon, *et al.*, Gut Microbiota in Patients With Irritable Bowel Syndrome—A Systematic Review. *Gastroenterology* **157** (1), 97–108 (2019), doi:10.1053/j.gastro.2019.03.049, http://dx.doi.org/10.1053/j.gastro.2019.03.049.

33. R. Duan, S. Zhu, B. Wang, L. Duan, Alterations of Gut Microbiota in Patients With Irritable Bowel Syndrome Based on 16S rRNA-Targeted Sequencing: A Systematic Review. *Clinical and Translational Gastroenterology* **10** (2), e00012 (2019), doi:10.14309/ctg.0000000000000012, http://dx.doi.org/10.14309/ctg.0000000000000012.

34. L. Wang, *et al.*, Gut Microbial Dysbiosis in the Irritable Bowel Syndrome: A Systematic Review and Meta-Analysis of Case-Control Studies. *Journal of the Academy of Nutrition and Dietetics* **120** (4), 565–586 (2020), doi:10.1016/j.jand.2019.05.015, http://dx.doi.org/10.1016/j.jand.2019.05.015.

35. Z. Dai, *et al.*, Multi-cohort analysis of colorectal cancer metagenome identified altered bacteria across populations and universal bacterial markers. *Microbiome* **6** (1) (2018), doi:10.1186/s40168-018-0451-2, http://dx.doi.org/10.1186/s40168-018-0451-2.

36. M. Song, A. T. Chan, J. Sun, Influence of the Gut Microbiome, Diet, and Environment on Risk of Colorectal Cancer. *Gastroenterology* **158** (2), 322–340 (2020), doi:10.1053/j.gastro.2019.06.048, http://dx.doi.org/10.1053/j.gastro.2019.06.048.

37. A. M. Thomas, *et al.*, Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nature Medicine* **25** (4), 667–678 (2019), doi:10.1038/s41591-019-0405-7, http://dx.doi.org/10.1038/s41591-019-0405-7.

38. A. J. Yarur, *et al.*, 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** (4), 963–975.e5 (2023), doi:10.1053/j.gastro.2023.06.036, http://dx.doi.org/10.1053/j.gastro.2023.06.036.

39. L. Ning, *et al.*, Microbiome and metabolome features in inflammatory bowel disease via multi-omics integration analyses across cohorts. *Nature Communications* **14** (1) (2023), doi:10.1038/s41467-023-42788-0, http://dx.doi.org/10.1038/s41467-023-42788-0.

40. H. M. Project, The Inflammatory Bowel Disease Multiomics Database, https://ibdmdb.org/, accessed on January 17th, 2025.

41. M. Berkell, *et al.*, Microbiota-based markers predictive of development of Clostridioides difficile infection. *Nature Communications* **12** (1) (2021), doi:10.1038/s41467-021-22302-0, http://dx.doi.org/10.1038/s41467-021-22302-0.

42. M. J. T. Crobach, *et al.*, The Bacterial Gut Microbiota of Adult Patients Infected, Colonized or Noncolonized by Clostridioides difficile. *Microorganisms* **8** (5), 677 (2020), doi:10.3390/microorganisms8050677, http://dx.doi.org/10.3390/microorganisms8050677.

43. P. Ferretti, *et al.*, C. difficile may be overdiagnosed in adults and is a prevalent commensal in infants. *eLife* (2023), doi:10.7554/elife.90111.1, http://dx.doi.org/10.7554/eLife.90111.1.

44. R. A. Mars, *et al.*, Longitudinal Multi-omics Reveals Subset-Specific Mechanisms Underlying Irritable Bowel Syndrome. *Cell* **182** (6), 1460–1473.e17 (2020), doi:10.1016/j.cell.2020.08.007, http://dx.doi.org/10.1016/j.cell.2020.08.007.

45. 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. *Microbiology Spectrum* **11** (1) (2023), doi:10.1128/spectrum.02125-22, http://dx.doi.org/10.1128/spectrum.02125-22.

46. Z. A. Barandouzi, *et al.*, Associations of neurotransmitters and the gut microbiome with emotional distress in mixed type of irritable bowel syndrome. *Scientific Reports* **12** (1) (2022), doi:10.1038/s41598-022-05756-0, http://dx.doi.org/10.1038/s41598-022-05756-0.

47. G. Zeller, *et al.*, Potential of fecal microbiota for early-stage detection of colorectal cancer. *Molecular Systems Biology* **10** (11) (2014), doi:10.15252/msb.20145645, http://dx.doi.org/10.15252/msb.20145645.

48. S. Yachida, *et al.*, Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nature Medicine* **25** (6), 968–976 (2019), doi:10.1038/s41591-019-0458-7, http://dx.doi.org/10.1038/s41591-019-0458-7.

49. E. Vogtmann, *et al.*, Colorectal Cancer and the Human Gut Microbiome: Reproducibility with Whole-Genome Shotgun Sequencing. *PLOS ONE* **11** (5), e0155362 (2016), doi:10.1371/journal.pone.0155362, http://dx.doi.org/10.1371/journal.pone.0155362.

50. 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 Microbiology Reviews* **41** (2), 182–199 (2017), doi:10.1093/femsre/fuw045, http://dx.doi.org/10.1093/femsre/fuw045.

51. D. A. Relman, The human microbiome: ecosystem resilience and health. *Nutrition Reviews* **70**, S2–S9 (2012), doi:10.1111/j.1753-4887.2012.00489.x, http://dx.doi.org/10.1111/j.1753-4887.2012.00489.x.

52. M. Van de Guchte, *et al.*, Alternative stable states in the intestinal ecosystem: proof of concept in a rat model and a perspective of therapeutic implications. *Microbiome* **8** (1) (2020), doi:10.1186/s40168-020-00933-7, http://dx.doi.org/10.1186/s40168-020-00933-7.

53. M. Arumugam, *et al.*, Enterotypes of the human gut microbiome. *Nature* **473** (7346), 174–180 (2011), doi:10.1038/nature09944, http://dx.doi.org/10.1038/nature09944.

54. P. I. Costea, *et al.*, Enterotypes in the landscape of gut microbial community composition. *Nature Microbiology* **3** (1), 8–16 (2017), doi:10.1038/s41564-017-0072-8, http://dx.doi.org/10.1038/s41564-017-0072-8.

55. L. Lahti, J. Salojärvi, A. Salonen, M. Scheffer, W. M. de Vos, Tipping elements in the human intestinal ecosystem. *Nature Communications* **5** (1) (2014), doi:10.1038/ncomms5344, http://dx.doi.org/10.1038/ncomms5344.

56. S. Zaoli, J. Grilli, A macroecological description of alternative stable states reproduces intra- and inter-host variability of gut microbiome. *Science Advances* **7** (43) (2021), doi:10.1126/sciadv.abj2882, http://dx.doi.org/10.1126/sciadv.abj2882.

57. A. L. Hartman, *et al.*, Human gut microbiome adopts an alternative state following small bowel transplantation. *Proceedings of the National Academy of Sciences* **106** (40), 17187–17192 (2009), doi:10.1073/pnas.0904847106, http://dx.doi.org/10.1073/pnas.0904847106.

58. J. Tap, E. Ruppé, M. Derrien, *The Human Gut Microbiota in all its States: From Disturbance to Resilience* (Elsevier), p. 161–178 (2022), doi:10.1016/b978-0-12-819265-8.00039-5, http://dx.doi.org/10.1016/B978-0-12-819265-8.00039-5.

59. J. Tap, *et al.*, Global branches and local states of the human gut microbiome define associations with environmental and intrinsic factors. *Nature Communications* **14** (1) (2023), doi:10.1038/s41467-023-38558-7, http://dx.doi.org/10.1038/s41467-023-38558-7.

60. H. Fujita, *et al.*, Alternative stable states, nonlinear behavior, and predictability of microbiome dynamics. *Microbiome* **11** (1) (2023), doi:10.1186/s40168-023-01474-5, http://dx.doi.org/10.1186/s40168-023-01474-5.

61. W. K. Chang, D. VanInsberghe, L. Kelly, Topological analysis reveals state transitions in human gut and marine bacterial communities. *npj Biofilms and Microbiomes* **6** (1) (2020), doi:10.1038/s41522-020-00145-9, http://dx.doi.org/10.1038/s41522-020-00145-9.

62. D. Garza, *et al.*, Discovery of alternative stable states in a synthetic human gut microbial community (2024), doi:10.1101/2024.11.28.625814, http://dx.doi.org/10.1101/2024.11.28.625814.

63. K. Faust, J. Raes, Microbial interactions: from networks to models. *Nature Reviews Microbiology* **10** (8), 538–550 (2012), doi:10.1038/nrmicro2832, http://dx.doi.org/10.1038/nrmicro2832.

64. D. Rios Garza, D. Gonze, H. Zafeiropoulos, B. Liu, K. Faust, Metabolic models of human gut microbiota: Advances and challenges. *Cell Systems* **14** (2), 109–121 (2023), doi:10.1016/j.cels.2022.11.002, http://dx.doi.org/10.1016/j.cels.2022.11.002.

65. N. I. van den Berg, *et al.*, Ecological modelling approaches for predicting emergent properties in microbial communities. *Nature Ecology and Evolution* **6** (7), 855–865 (2022), doi:10.1038/s41559-022-01746-7, http://dx.doi.org/10.1038/s41559-022-01746-7.

66. Y.-Y. Liu, Controlling the human microbiome. *Cell Systems* **14** (2), 135–159 (2023), doi:10.1016/j.cels.2022.12.010, http://dx.doi.org/10.1016/j.cels.2022.12.010.

67. E. Lange, L. Kranert, J. Krüger, D. Benndorf, R. Heyer, Microbiome modeling: a beginner's guide. *Frontiers in Microbiology* **15** (2024), doi:10.3389/fmicb.2024.1368377, http://dx.doi.org/10.3389/fmicb.2024.1368377.

68. C. G. Buffie, *et al.*, Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. *Nature* **517** (7533), 205–208 (2014), doi:10.1038/nature13828, http://dx.doi.org/10.1038/nature13828.

69. T. S. B. Schmidt, *et al.*, Drivers and determinants of strain dynamics following fecal microbiota transplantation. *Nature Medicine* **28** (9), 1902–1912 (2022), doi:10.1038/s41591-022-01913-0, http://dx.doi.org/10.1038/s41591-022-01913-0.

70. G. Aguadé-Gorgorió, J. Arnoldi, M. Barbier, S. Kéfi, A taxonomy of multiple stable states in complex ecological communities. *Ecology Letters* **27** (4) (2024), doi:10.1111/ele.14413, http://dx.doi.org/10.1111/ele.14413.

71. S. Krishnan, N. Alden, K. Lee, Pathways and functions of gut microbiota metabolism impacting host physiology. *Current Opinion in Biotechnology* **36**, 137–145 (2015), doi:10.1016/j.copbio.2015.08.015, http://dx.doi.org/10.1016/j.copbio.2015.08.015.

72. K. Oliphant, E. Allen-Vercoe, Macronutrient metabolism by the human gut microbiome: major fermentation by-products and their impact on host health. *Microbiome* **7** (1) (2019), doi:10.1186/s40168-019-0704-8, http://dx.doi.org/10.1186/s40168-019-0704-8.

73. M. A. Fischbach, J. L. Sonnenburg, Eating For Two: How Metabolism Establishes Interspecies Interactions in the Gut. *Cell Host and Microbe* **10** (4), 336–347 (2011), doi:10.1016/j.chom.2011.10.002, http://dx.doi.org/10.1016/j.chom.2011.10.002.

74. P. Portincasa, *et al.*, Gut Microbiota and Short Chain Fatty Acids: Implications in Glucose Homeostasis. *International Journal of Molecular Sciences* **23** (3), 1105 (2022), doi:10.3390/ijms23031105, http://dx.doi.org/10.3390/ijms23031105.

75. K. A. Krautkramer, J. Fan, F. Bäckhed, Gut microbial metabolites as multi-kingdom intermediates. *Nature Reviews Microbiology* **19** (2), 77–94 (2020), doi:10.1038/s41579-020-0438-4, http://dx.doi.org/10.1038/s41579-020-0438-4.

76. P.-J. Lahtvee, *et al.*, Multi-omics approach to study the growth efficiency and amino acid metabolism in Lactococcus lactis at various specific growth rates. *Microbial Cell*

Factories **10** (1) (2011), doi:10.1186/1475-2859-10-12, http://dx.doi.org/10.1186/1475-2859-10-12.

77. I. Rowland, *et al.*, Gut microbiota functions: metabolism of nutrients and other food components. *European Journal of Nutrition* **57** (1), 1–24 (2017), doi:10.1007/s00394-017-1445-8, http://dx.doi.org/10.1007/s00394-017-1445-8.

78. Z.-N. Ling, *et al.*, Amino acid metabolism in health and disease. *Signal Transduction and Targeted Therapy* **8** (1) (2023), doi:10.1038/s41392-023-01569-3, http://dx.doi.org/10.1038/s41392-023-01569-3.

79. T. Wang, A. Goyal, V. Dubinkina, S. Maslov, Evidence for a multi-level trophic organization of the human gut microbiome. *PLOS Computational Biology* **15** (12), e1007524 (2019), doi:10.1371/journal.pcbi.1007524, http://dx.doi.org/10.1371/journal.pcbi.1007524.

80. J. Grilli, Macroecological laws describe variation and diversity in microbial communities. *Nature Communications* **11** (1) (2020), doi:10.1038/s41467-020-18529-y, http://dx.doi.org/10.1038/s41467-020-18529-y.

81. 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.

82. A. Moya, M. Ferrer, Functional Redundancy-Induced Stability of Gut Microbiota Subjected to Disturbance. *Trends in Microbiology* **24** (5), 402–413 (2016), doi:10.1016/j.tim.2016.02.002, http://dx.doi.org/10.1016/j.tim.2016.02.002.

83. S. Louca, *et al.*, Function and functional redundancy in microbial systems. *Nature Ecology and Evolution* **2** (6), 936–943 (2018), doi:10.1038/s41559-018-0519-1, http://dx.doi.org/10.1038/s41559-018-0519-1.

84. L. Fant, I. Macocco, J. Grilli, Eco-evolutionary dynamics lead to functionally robust and redundant communities. *bioRxiv* (2021), doi:10.1101/2021.04.02.438173, https://www.biorxiv.org/content/early/2021/08/23/2021.04.02.438173.

85. A. B. George, T. Wang, S. Maslov, Functional convergence in slow-growing microbial communities arises from thermodynamic constraints. *The ISME Journal* **17** (9), 1482–1494 (2023), doi:10.1038/s41396-023-01455-4, http://dx.doi.org/10.1038/s41396-023-01455-4.

86. Materials and methods are available as supplementary material.

87. R. L. Tatusov, The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research* **28** (1), 33–36 (2000), doi:10.1093/nar/28.1.33, http://dx.doi.org/10.1093/nar/28.1.33.

88. H. Ogata, *et al.*, KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **27** (1), 29–34 (1999), doi:10.1093/nar/27.1.29, http://dx.doi.org/10.1093/nar/27.1.29.

89. G. Bunin, Ecological communities with Lotka-Volterra dynamics. *Physical Review E* **95** (4), 042414 (2017).

90. F. Roy, M. Barbier, G. Biroli, G. Bunin, Complex interactions can create persistent fluctuations in high-diversity ecosystems. *PLoS computational biology* **16** (5), e1007827 (2020).

91. E. Mallmin, A. Traulsen, S. De Monte, Chaotic turnover of rare and abundant species in a strongly interacting model community. *Proceedings of the National Academy of Sciences* **121** (11), e2312822121 (2024).

92. I. Dalmedigos, G. Bunin, Dynamical persistence in high-diversity resource-consumer communities. *PLoS computational biology* **16** (10), e1008189 (2020).

93. E. Blumenthal, J. W. Rocks, P. Mehta, Phase Transition to Chaos in Complex Ecosystems with Nonreciprocal Species-Resource Interactions. *Physical Review Letters* **132** (12), 127401 (2024).

94. X. C. Morgan, *et al.*, Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biology* **13** (9) (2012), doi:10.1186/gb-2012-13-9-r79, http://dx.doi.org/10.1186/gb-2012-13-9-r79.

95. M. Schirmer, A. Garner, H. Vlamakis, R. J. Xavier, Microbial genes and pathways in inflammatory bowel disease. *Nature Reviews Microbiology* **17** (8), 497–511 (2019), doi: 10.1038/s41579-019-0213-6, http://dx.doi.org/10.1038/s41579-019-0213-6.

96. K. Z. Coyte, S. Rakoff-Nahoum, Understanding Competition and Cooperation within the Mammalian Gut Microbiome. *Current Biology* **29** (11), R538–R544 (2019), doi:10.1016/j.cub.2019.04.017, http://dx.doi.org/10.1016/j.cub.2019.04.017.

97. J. Sung, *et al.*, Global metabolic interaction network of the human gut microbiota for context-specific community-scale analysis. *Nature Communications* **8** (1) (2017), doi: 10.1038/ncomms15393, http://dx.doi.org/10.1038/ncomms15393.

98. K. Z. Coyte, J. Schluter, K. R. Foster, The ecology of the microbiome: Networks, competition, and stability. *Science* **350** (6261), 663–666 (2015), doi:10.1126/science.aad2602, http://dx.doi.org/10.1126/science.aad2602.

99. 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. *Nature Microbiology* **9** (4), 1036–1048 (2024), doi:10.1038/s41564-024-01625-w, http://dx.doi.org/10.1038/s41564-024-01625-w.

100. E. J. Culp, A. L. Goodman, Cross-feeding in the gut microbiome: Ecology and mechanisms. *Cell Host and Microbe* **31** (4), 485–499 (2023), doi:10.1016/j.chom.2023.03.016, http://dx.doi.org/10.1016/j.chom.2023.03.016.

101. A. Goyal, T. Wang, V. Dubinkina, S. Maslov, Ecology-guided prediction of cross-feeding interactions in the human gut microbiome. *Nature Communications* **12** (1) (2021), doi:10.1038/s41467-021-21586-6, http://dx.doi.org/10.1038/s41467-021-21586-6.

102. V. R. Marcelino, *et al.*, Disease-specific loss of microbial cross-feeding interactions in the human gut. *Nature Communications* **14** (1) (2023), doi:10.1038/s41467-023-42112-w, http://dx.doi.org/10.1038/s41467-023-42112-w.

103. A. R. Watson, *et al.*, Metabolic independence drives gut microbial colonization and resilience in health and disease. *Genome Biology* **24** (1) (2023), doi:10.1186/s13059-023-02924-x, http://dx.doi.org/10.1186/s13059-023-02924-x.

104. 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 Systems* **9** (3), 286–296.e8 (2019), doi:10.1016/j.cels.2019.08.002, http://dx.doi.org/10.1016/j.cels.2019.08.002.

105. C. Li, *et al.*, BEEM-Static: Accurate inference of ecological interactions from cross-sectional microbiome data. *PLOS Computational Biology* **17** (9), e1009343 (2021), doi:10.1371/journal.pcbi.1009343, http://dx.doi.org/10.1371/journal.pcbi.1009343.

106. C. K. Fisher, P. Mehta, Identifying Keystone Species in the Human Gut Microbiome from Metagenomic Timeseries Using Sparse Linear Regression. *PLoS ONE* **9** (7), e102451 (2014), doi:10.1371/journal.pone.0102451, http://dx.doi.org/10.1371/journal.pone.0102451.

107. I. Veseli, *et al.*, Microbes with higher metabolic independence are enriched in human gut microbiomes under stress (2024), doi:10.7554/elife.89862.2, http://dx.doi.org/10.7554/eLife.89862.2.

108. S. S. Gude, *et al.*, Colorectal Cancer Diagnostic Methods: The Present and Future. *Cureus* (2023), doi:10.7759/cureus.37622, http://dx.doi.org/10.7759/cureus.37622.

109. S. D. Jurburg, S. A. Blowes, A. Shade, N. Eisenhauer, J. M. Chase, Synthesis of recovery patterns in microbial communities across environments. *Microbiome* **12** (1), 79 (2024).

110. S. Weiss, *et al.*, Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *The ISME Journal* **10** (7), 1669–1681 (2016), doi:10.1038/ismej.2015.235, http://dx.doi.org/10.1038/ismej.2015.235.

111. M. Cappellato, G. Baruzzo, I. Patuzzi, B. Di Camillo, Modeling Microbial Community Networks: Methods and Tools. *Current Genomics* **22** (4),

267–290 (2021), doi:10.2174/1389202921999200905133146, http://dx.doi.org/10.2174/1389202921999200905133146.

112. P. D. Karp, *et al.*, The BioCyc collection of microbial genomes and metabolic pathways. *Briefings in Bioinformatics* **20** (4), 1085–1093 (2017), doi:10.1093/bib/bbx085, http://dx.doi.org/10.1093/bib/bbx085.

113. J. P. Jacobs, *et al.*, Multi-omics profiles of the intestinal microbiome in irritable bowel syndrome and its bowel habit subtypes. *Microbiome* **11** (1) (2023), doi:10.1186/s40168-022-01450-5, http://dx.doi.org/10.1186/s40168-022-01450-5.

114. J. B. Russell, G. M. Cook, Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiological Reviews* **59** (1), 48–62 (1995), doi:10.1128/mr.59.1.48-62.1995, http://dx.doi.org/10.1128/mr.59.1.48-62.1995.

115. J. H. de Queiroz, J.-L. Uribelarrea, A. Pareilleux, Estimation of the energetic biomass yield and efficiency of oxidative phosphorylation in cell-recycle cultures of Schizosaccharomyces pombe. *Applied Microbiology and Biotechnology* **39** (4–5), 609–614 (1993), doi:10.1007/bf00205061, http://dx.doi.org/10.1007/BF00205061.

116. K. Adamberg, G. Raba, S. Adamberg, Use of Changestat for Growth Rate Studies of Gut Microbiota. *Frontiers in Bioengineering and Biotechnology* **8** (2020), doi:10.3389/fbioe.2020.00024, http://dx.doi.org/10.3389/fbioe.2020.00024.

117. D. A. Southgate, How much and what classes of carbohydrate reach the colon. *European Journal of Cancer Prevention* **7** (2), S81–82 (1998), doi:10.1097/00008469-199805000-00014, http://dx.doi.org/10.1097/00008469-199805000-00014.

118. M. Stipanuk, M. Caudill, *Biochemical, Physiological, and Molecular Aspects of Human Nutrition - E-Book: Biochemical, Physiological, and Molecular Aspects of Human Nutrition - E-Book* (Elsevier Health Sciences) (2012), https://books.google.es/books?id=XVNPAQAAQBAJ.

119. C. Mulder, A. J. Hendriks, Half-saturation constants in functional responses. *Global Ecology and Conservation* **2**, 161–169 (2014), doi:10.1016/j.gecco.2014.09.006, http://dx.doi.org/10.1016/j.gecco.2014.09.006.

120. J. W. Fink, N. A. Held, M. Manhart, Microbial population dynamics decouple growth response from environmental nutrient concentration. *Proceedings of the National Academy of Sciences* **120** (2) (2023), doi:10.1073/pnas.2207295120, http://dx.doi.org/10.1073/pnas.2207295120.

121. R. Sender, S. Fuchs, R. Milo, Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* **164** (3), 337–340 (2016), doi:10.1016/j.cell.2016.01.013, http://dx.doi.org/10.1016/j.cell.2016.01.013.

122. D. T. Truong, *et al.*, MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nature Methods* **12** (10), 902–903 (2015), doi:10.1038/nmeth.3589, http://dx.doi.org/10.1038/nmeth.3589.

123. S. Sunagawa, *et al.*, Metagenomic species profiling using universal phylogenetic marker genes. *Nature Methods* **10** (12), 1196–1199 (2013), doi:10.1038/nmeth.2693, http://dx.doi.org/10.1038/nmeth.2693.

Acknowledgments

The authors would like to thank Jacopo Grilli, Matteo Sireci, and the members of the Bonachela lab 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 ERDF/EU, 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 #826106) and NSF grant DMS-2052616 to J.A.B.

Author contributions: JAB and MAM designed and supervised the research. RCL, JAB, and MAM developed the model, with input of SAL, MJB, MM, and MGDB. RCL collected the data and performed statistical and modeling analyses. All authors wrote and reviewed the manuscript.

Competing interests: There are no competing interests to declare.

Data and materials availability: No empirical data were produced in this study. All the publicly available datasets are cited in the section ”Datasets” in (*86*). Our code will be available on either zenodo or github.

Supplementary Materials for
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,*,†]

[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 Robert Wood Johnson Medical School, Piscataway, NJ, USA.[7] Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA.*Corresponding authors. Email: mamunoz@onsager.ugr.es and juan.bonachela@rutgers.edu

[†]These authors contributed equally to this work.

This PDF file includes:

Materials and Methods

Supplementary Text

Figures S1 to S11

Table S1

Materials and Methods

Measures

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

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.

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

$$BC_{a,d} = 1 - \frac{2\sum_i \min(b_i^a, b_i^d)}{\sum_i (b_i^a + b_i^d)} \tag{S2}$$

where b_i^a and b_i^d represent, for samples a and d, the relative abundance of type i (model) 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:

$$\text{Total Enzymatic Cost} = \sum_{i\alpha\beta} P^i_{\alpha\beta} C_{\alpha\beta} \tag{S3}$$

where $P^i_{\alpha\beta}$ 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

$$\text{Number of Metabolites} = \sum_\alpha \theta(S_\alpha) \tag{S4}$$

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:

$$\text{Pathway Fraction}_{\alpha\beta} = \frac{\sum_i P^i_{\alpha\beta} C_{\alpha\beta} B_i}{\sum_{i,\alpha,\beta} P^i_{\alpha\beta} C_{\alpha\beta} B_i} \tag{S5}$$

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

$$c^+_{ij} = \sum_{\alpha\beta\delta} P^i_{\alpha\beta} P^j_{\beta\delta} r_i(S_\alpha) r_j(S_\beta) B_i B_j \tag{S6}$$

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

$$c^-_{ij} = \sum_{\alpha\beta\delta} P^i_{\alpha\beta} P^j_{\alpha\delta} r_i(S_\alpha) r_j(S_\alpha) B_i B_j \tag{S7}$$

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

$$\rho = \frac{\sum_{i \neq j} (c^+_{ij} - c^-_{ij})}{\sum_{i \neq j} (c^+_{ij} + c^-_{ij})} \tag{S8}$$

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

$$\rho_{\text{inferred}} = \frac{\sum_{i \neq j}(\omega_{ij}^{+} - \omega_{ij}^{-})}{\sum_{i \neq j}(\omega_{ij}^{+} + \omega_{ij}^{-})} \tag{S9}$$

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

$$\text{ENBI} = \rho_{\text{inferred}}^{\text{U}} - \rho_{\text{inferred}}^{\text{H}} \tag{S10}$$

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).

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

To infer the underlying interaction network of the microbial community, we used three complementary methods: 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).

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 (*104*) 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. S9 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 (*104*). 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 pub-

lication (*105*). 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. Supplementary Fig.S10 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 (*106*).

For each subject, we selected 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 averaging the outcomes from all iterations corresponding to each error threshold, and subsequently averaging the results across all error thresholds. In our implementation, we applied a slight modification to the original algorithm by using the average, rather than the median, of all iterations to construct the final network. This modification allowed us to obtain more consistent results.

Model simulation details

We integrated numerically the model equations described in the Supplementary Text 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 Supplementary Text Eqs. S14 and S15). The trade-off function used for the simulations presented in the main text used 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 Supplementary Text, 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 (*112*). 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 (data used as obtained from original sources):

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 (*44*). Metabolomic data obtained from (*113*).

CDI: Metagenomic data obtained from (*43*).

CRC: Metagenomic, enzymatic, and metabolomic data were obtained from (*48*).

F4 and M3 gut data: Metagenomic data obtained from https://github.com/twbattaglia/MicrobeDS.

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.

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	(*114, 115*)
δ	Outflow rate	$\frac{1}{24}$ 1/h	(*116*)
h	Inflow rate of substances	2 g/(L h)	(*72, 117, 118*)
K	Half-saturation constant	10^{-4} g/L	(*119, 120*)
$r^{\max}$	Maximum uptake rate	10^{-10} g/(L cell h)	(*120*)
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	(*121*)
R	Total number of substances	30	-
B_T	Total biomass in a typical gut microbiome	$\sim 3.9 \times 10^{13}$ cell	(*121*)

Supplementary Text

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 types (e.g., species, strains, or another suitable taxonomical group) with abundances $B_i(t)$, where i identifies the bacterial type. 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 types 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 types— is positive. If the invasion fitness is negative and the initial biomass of the invading bacteria is sufficiently small, the new bacterial type is declared extinct in the next temporal step without affecting the system. Therefore, only viable invader types 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. First,a random number of pathways is assigned to the bacterium, ranging between 1 and a specified maximum. 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. 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:

$$\frac{dB_i}{dt} = B_i \left(\gamma_i \sum_{\alpha\beta}^{R} r_{i,\alpha}^{\max} \frac{S_\alpha}{K_{i,\alpha} + S_\alpha} P^i_{\alpha\beta} E_{\alpha\beta} - \delta \right) \tag{S11}$$

$$\frac{dS_\alpha}{dt} = h_\alpha + \sum_{i}^{N} \sum_{\beta}^{R} B_i \left(P^i_{\beta\alpha} r_{i,\beta}^{\max} \frac{S_\beta}{K_{i,\beta} + S_\beta} - P^i_{\alpha\beta} r_{i,\alpha}^{\max} \frac{S_\alpha}{K_{i,\alpha} + S_\alpha} \right) - \delta S_\alpha \tag{S12}$$

where γ_i is the yield or conversion constant from energy to biomass for bacterium i; $r_{i,\alpha}^{max}$ 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. 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:

$$\gamma_i \propto \gamma \cdot g \left(\sum_{\alpha,\beta} C_{\alpha\beta} P^i_{\alpha\beta} \right) \tag{S13}$$

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

$$g(x) = \frac{b}{1 + c \cdot x^\nu} \tag{S14}$$

$$g(x) = b e^{-c \cdot x} \tag{S15}$$

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.

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 tradeoff (see above), certain parametrizations 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 type 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 type 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. S11 and S13):

$$\text{Growth rate} \propto g\left(\sum_{\alpha,\beta} C_{\alpha\beta} P^i_{\alpha\beta}\right) \sum_{\alpha\beta}^{R} \frac{S_\alpha}{K_{i,\alpha} + S_\alpha} P^i_{\alpha\beta} e_{\alpha\beta} \tag{S16}$$

where $e_{\alpha\beta} = E_{\alpha\beta}\frac{M_\alpha}{V_{\text{colon}}}$ represents the energy of path $S_\alpha \rightarrow S_\beta$ 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:

$$\text{Growth rate} \propto g\left(C^i_{\text{T}}\right) e^i_{\text{T}} \tag{S17}$$

where $C^i_{\text{T}} = \sum_{\alpha,\beta} C_{\alpha\beta} P^i_{\alpha\beta}$ represents the total enzymatic cost per bacterial type, and $e^i_{\text{T}} = \sum_{\alpha\beta}^{R} P^i_{\alpha\beta} e_{\alpha\beta}$ 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:

$$\text{Growth rate} \propto g\left(C^i_{\text{T}}\right) C^i_{\text{T}} \tag{S18}$$

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.(S18) 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. S18 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. S18 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.

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 type able to metabolize this substance (and thus transform it into a random, different substance).

From that initial conditions, the simulations proceed according to the following algorithm:

0. Run the ecological dynamics, given by Eqs. S11-S12 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 type whose pathways are randomly selected (see above).
3. Repeat until a selected ending time.

Comparison to existing, relevant work

Recent studies have highlighted the importance of metabolic interactions in shaping microbial communities in health and disease. Marcelino et al. (*102*) introduced a quantitative framework for microbial cross-feeding, demonstrating that specific metabolite exchanges are disrupted in diseased states. Similarly, Watson et al. (*103*) and Veseli et al. (*107*) showed that microbes in dysbiotic

conditions exhibit increased metabolic self-sufficiency, indicating reduced dependence on cross-feeding for certain anabolic pathways.

While these observations might initially appear contradictory to our results, they are in fact complementary. Our approach quantifies net metabolic interactions, incorporating both cross-feeding and competition. Unlike previous studies, which emphasize the loss of specific cross-feeding dependencies, our findings indicate that cross-feeding interactions become more significant relative to competition during dysbiosis. Importantly, this relative increase does not necessarily reflect an absolute rise in cross-feeding; it could instead result from a sharper decline in competitive interactions. Furthermore, prior studies typically focus on specific metabolic pathways that get altered in disease, whereas our analysis assesses the overall balance between interaction types.

These studies also emphasized how metabolic self-sufficiency provides an advantage under stressful conditions, aligning with our observation that disease states are characterized by fewer available metabolites, favoring taxa less reliant on external metabolic inputs. Marcelino et al. specifically suggest that high species diversity is associated with increased net metabolite production, creating more opportunities for consumer species to thrive. This aligns with our simulations, in which higher diversity consistently corresponds with greater metabolite availability.

Therefore, our results complement existing literature by illustrating that the balance of microbial interactions shifts during disease, notably elevating the relative importance of cross-feeding compared to competition. These findings highlight the value of integrative approaches that simultaneously consider both specific metabolic dependencies and the broader dynamics of microbial interactions.

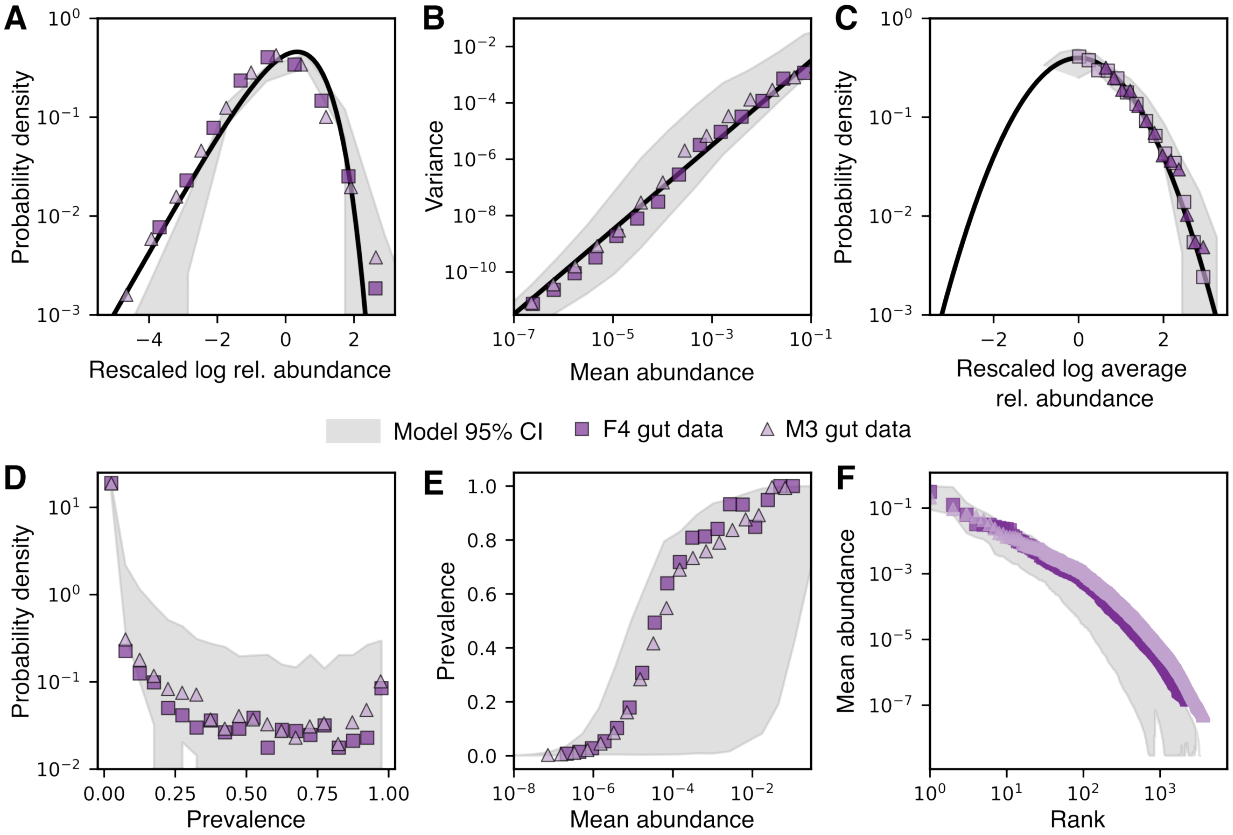

Figure S1: **Model reproduces well-established macroecological patterns of the gut microbiome without any parameter fitting.**(**A-C**) Macroecological patterns described in (*80*). (**A**) The Abundance Fluctuation Distribution (AFD), which characterizes the distribution of a species' abundances across 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 (*81*). (**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.

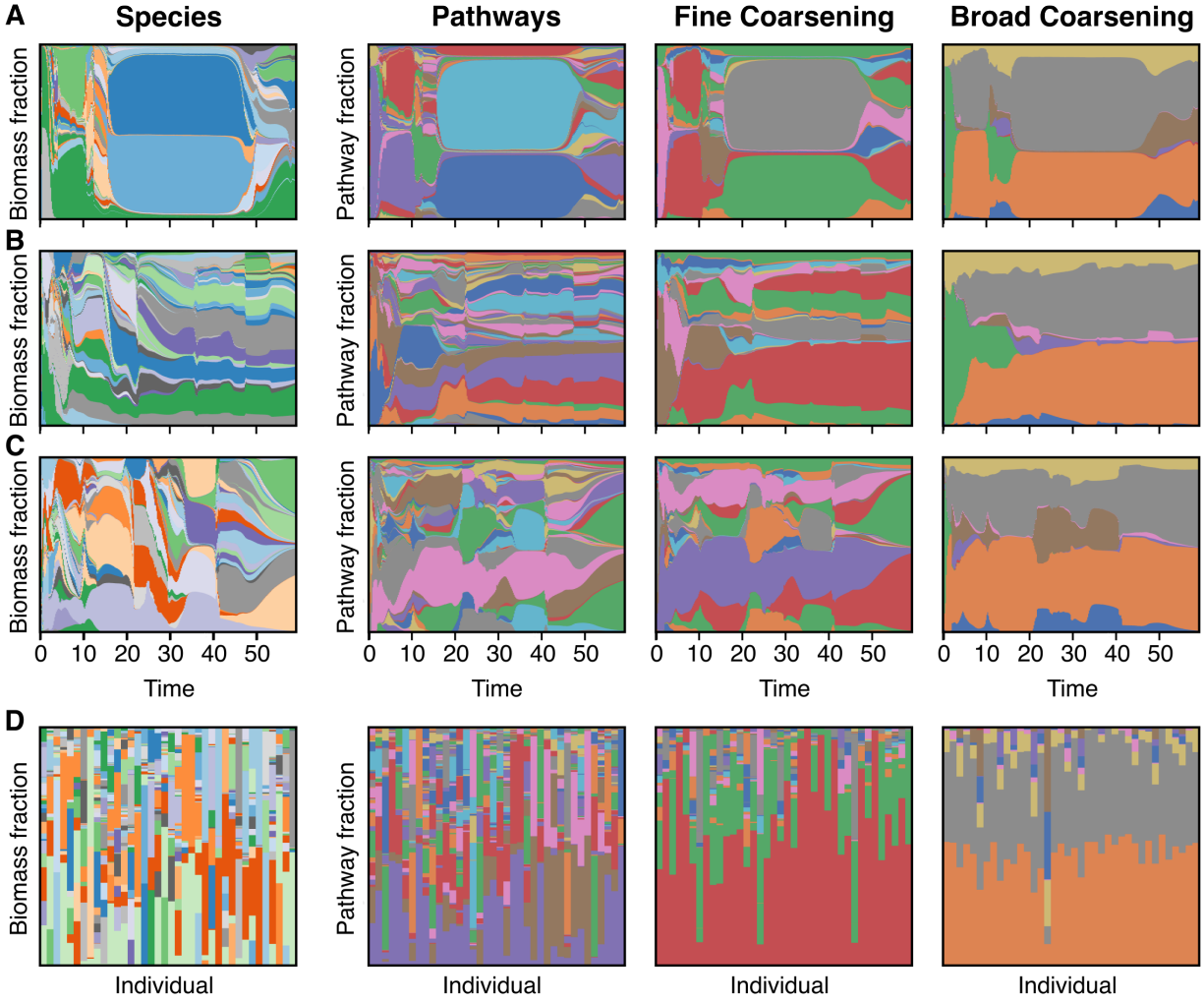

Figure S2: **Model data reproduces functional redundancy.** (**A-C**) Time-series of relative species abundance (left), enzymatic cost fraction for all pathways (number of resulting functional groups, $n_{\mathrm{fg}} = 435$; middle left), for a fine coarsening ($n_{\mathrm{fg}} = 43$; middle right), and for a broad coarsening ($n_{\mathrm{fg}} = 9$; 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.

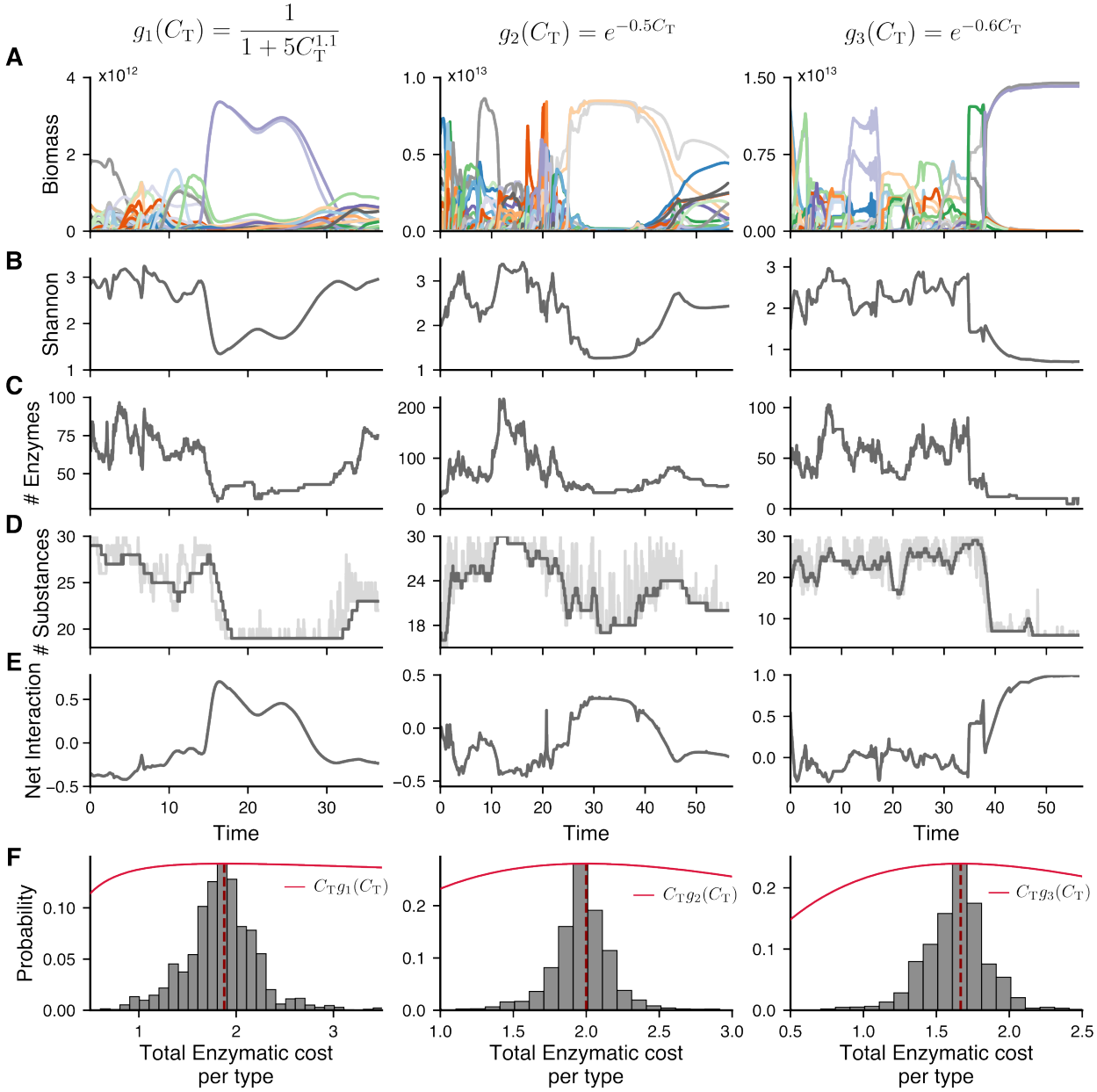

Figure S3: **Alternative stable states are robust against trade-off changes.** For three different trade-off functions, g_1 (left), g_2 (middle), and g_3 (right), the following are shown: (**A-E**) Time-series of bacterial abundances (**A**), Shannon index (**B**), enzymatic cost (**C**), number of substances (**D**), and net interaction (**E**). (**F**) The total enzymatic cost per type (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.

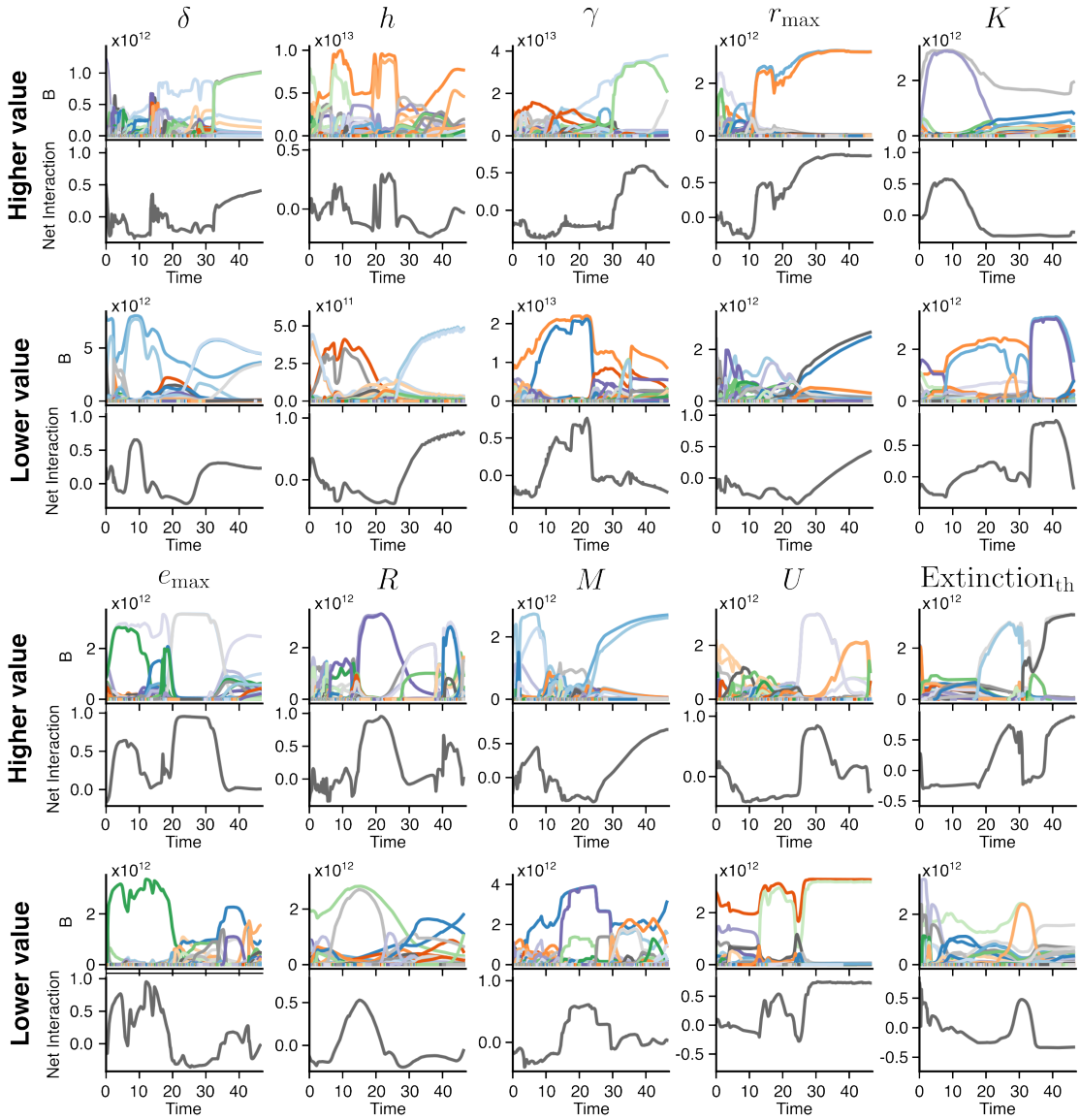

Figure S4: **Alternative stable states are robust to parameter variations.** Time-series of total biomass (top row in each subpanel) and net interaction (bottom row in each subpanel) for different parameter settings in the model. Each column represents a change from the original parameter values reported 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.

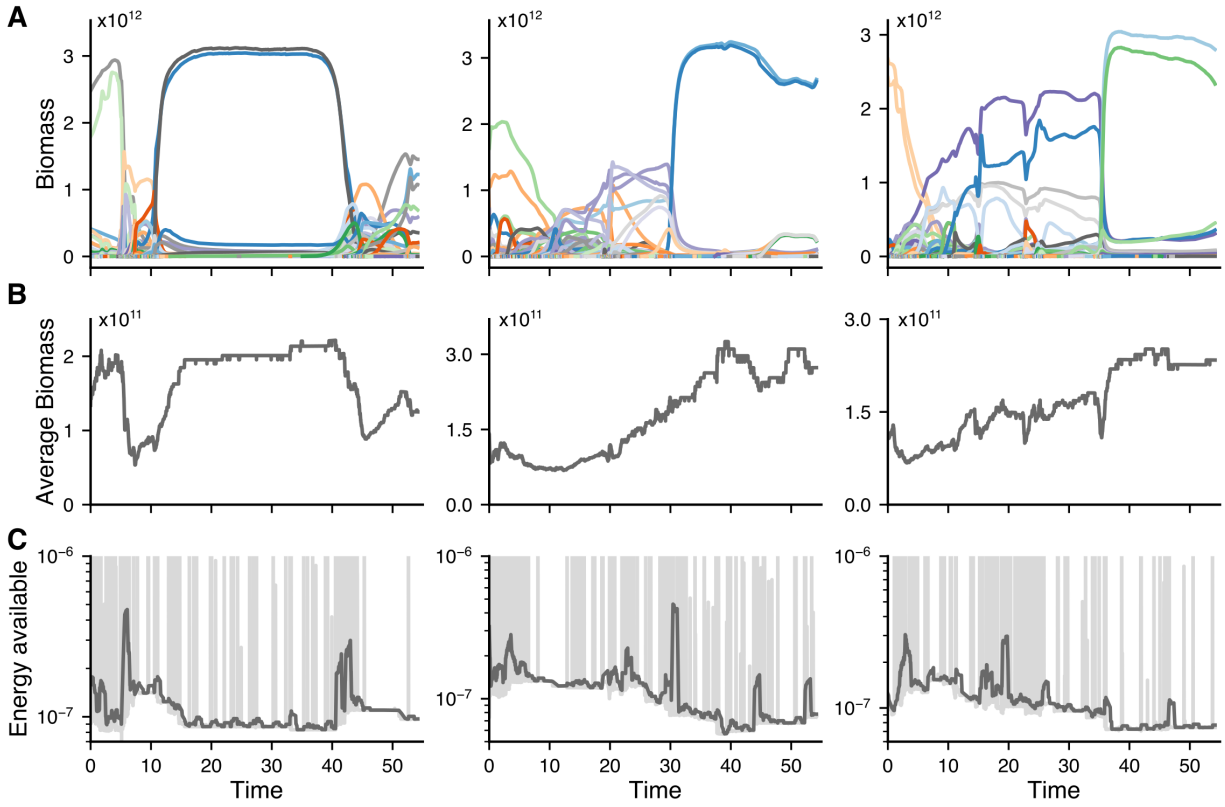

Figure S5: **dysbiotic state shows greater efficiency than the healthy state.** (**A-C**) Time-series of bacterial biomasses (**A**), average biomass across the system (**B**), and available energy within the system (**C**) for three different realizations of the model. In (**C**), dark grey lines represent 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.

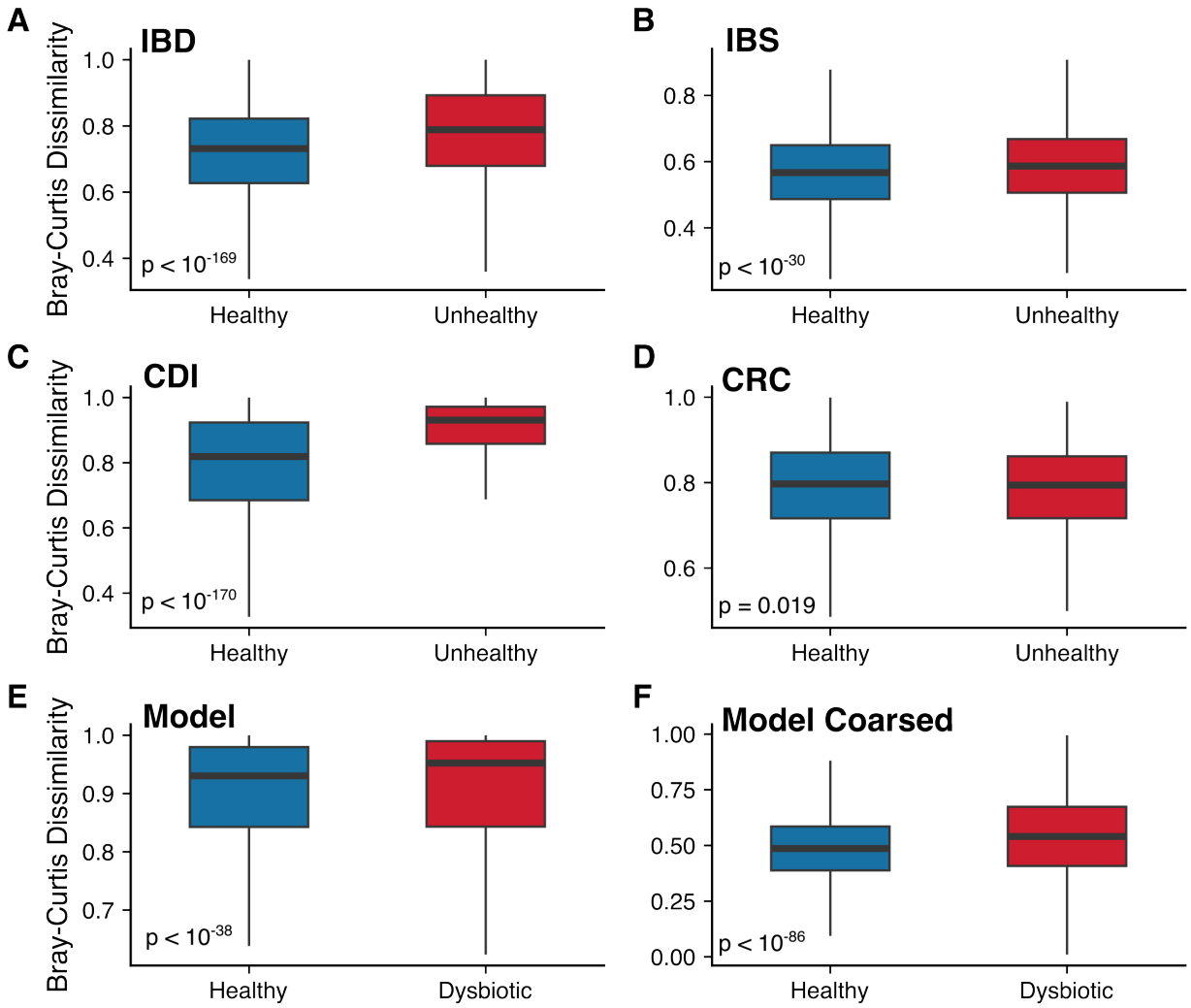

Figure S6: **The Anna Karenina Principle holds across different diseases, but not all, and in the model.** Beta diversity calculated with the Bray-Curtis dissimilarity of healthy and diseased states for (**A**) IBD, (**B**) IBS, (**C**) CDI, (**D**) CRC, as well as for healthy and collapsed states in the model (**E**) and in the model with coarse-grained pathways (**F**) (see Methods (*86*)). 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 (*109*)).

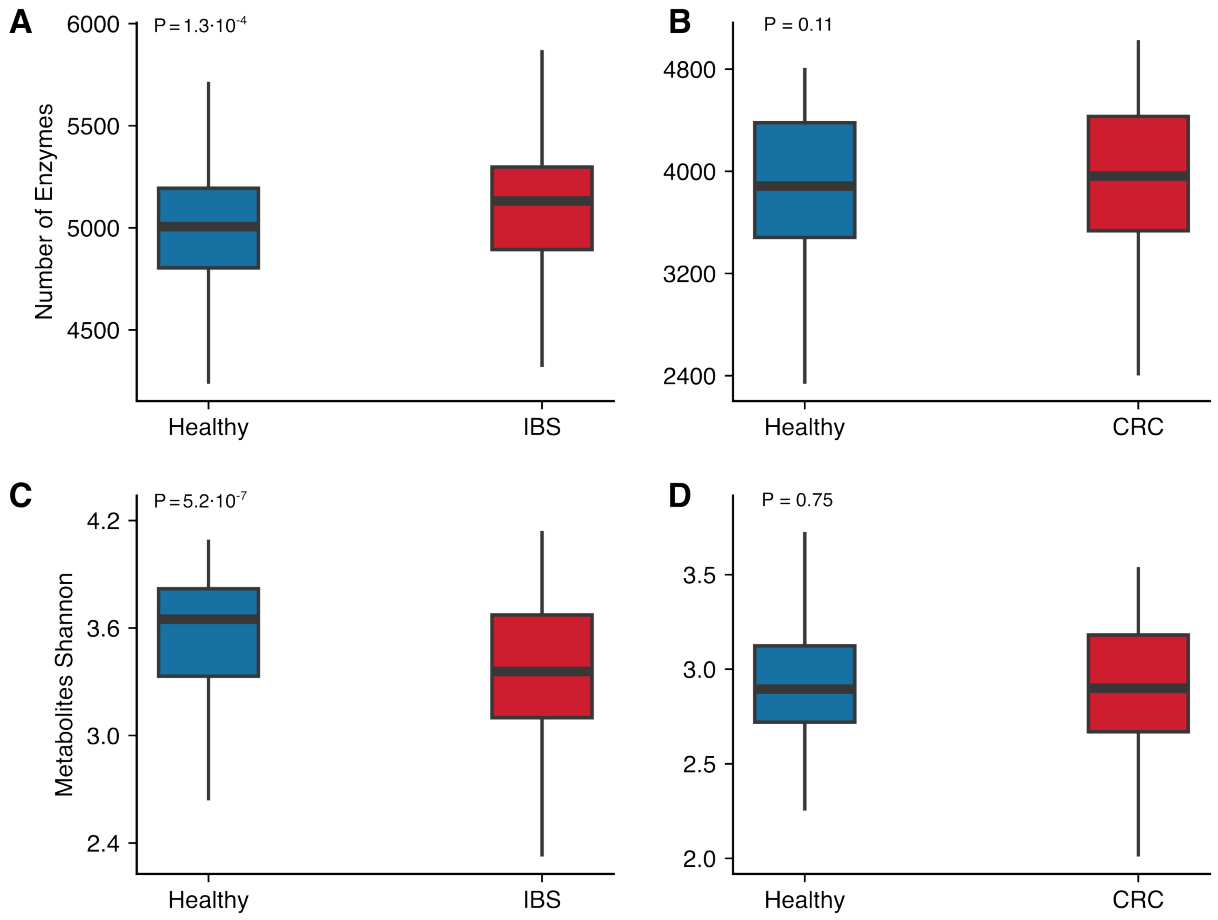

Figure S7: **Analyzed biomarkers are not consistently reduced across diseases.** (**A-B**) Number of enzymes in Healthy vs. IBS (**A**) and Healthy vs. CRC (**B**). (**C-D**) Shannon index of secreted metabolites in Healthy vs. IBS (**C**) and Healthy vs. CRC (**D**). The inconsistencies observed across diseases indicate that these biomarkers alone are not reliable indicators of dysbiosis.

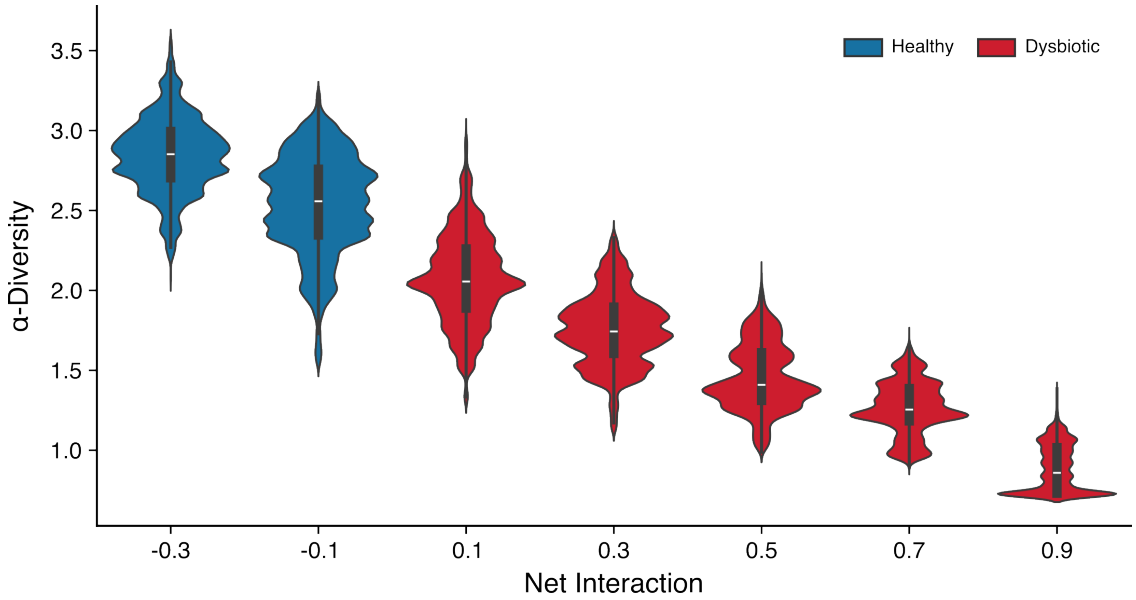

Figure S8: **Diversity exhibits high variability across net interaction values.** The mean and median of α-diversity for a given value of the net interaction shows a negative correlation between the Shannon index and ρ. Diversity is, however, widely distributed, which suggests that diversity alone is 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 $[-0.1, 0.7]$. All together supports the claim that a switch to dysbiotic (i.e. to a cooperative network) typically but not always correlates with a decrease in diversity.

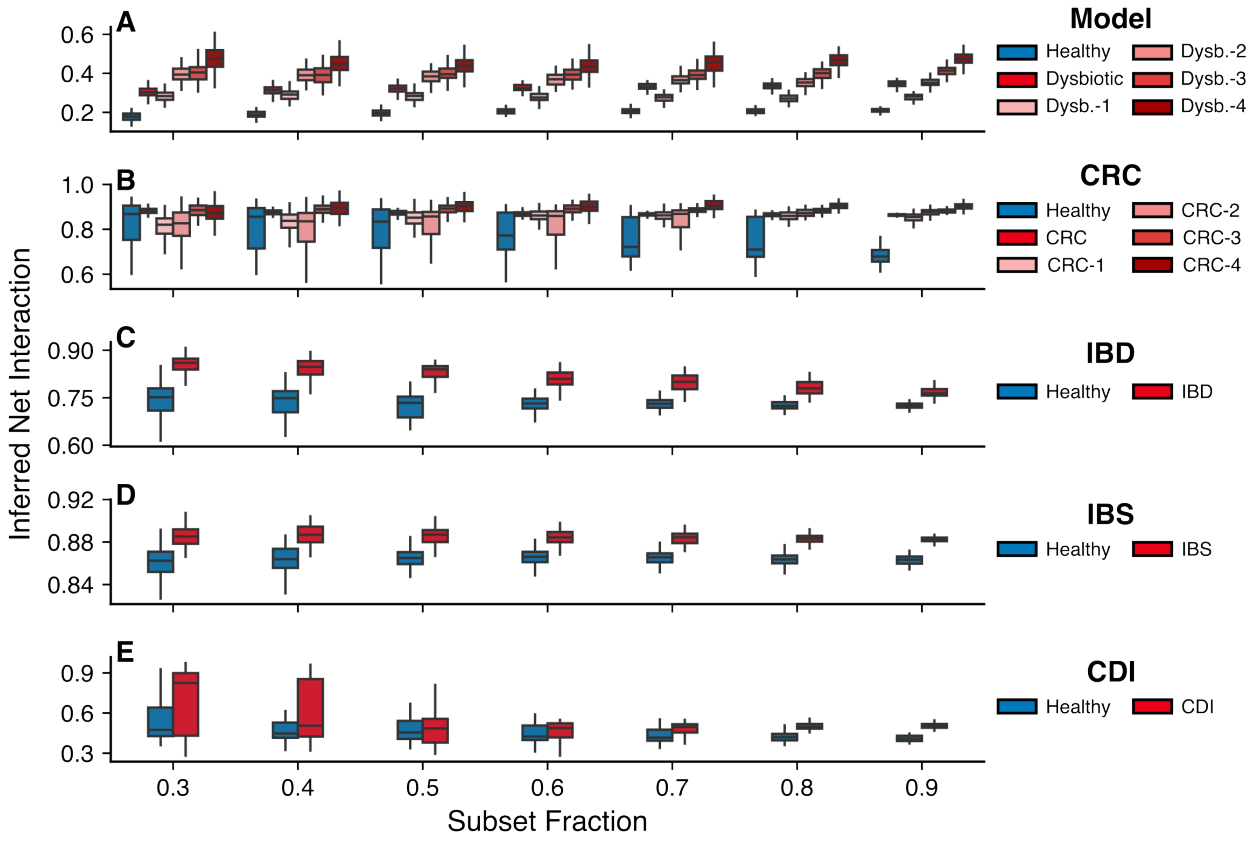

Figure S9: **The inferred net interaction is robust under bootstrapping.** Inferred net interaction calculated for different subsampling fractions of the original dataset for the model (**A**), CRC (**B**), IBD (**C**), IBS (**D**), and CDI (**E**). The inferred net interaction reliably captures the qualitative trends of the true underlying net interaction, ρ, in the model across all subsampled datasets. Healthy (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. All results are highly statistically significant (500 bootstraping-generated samples)

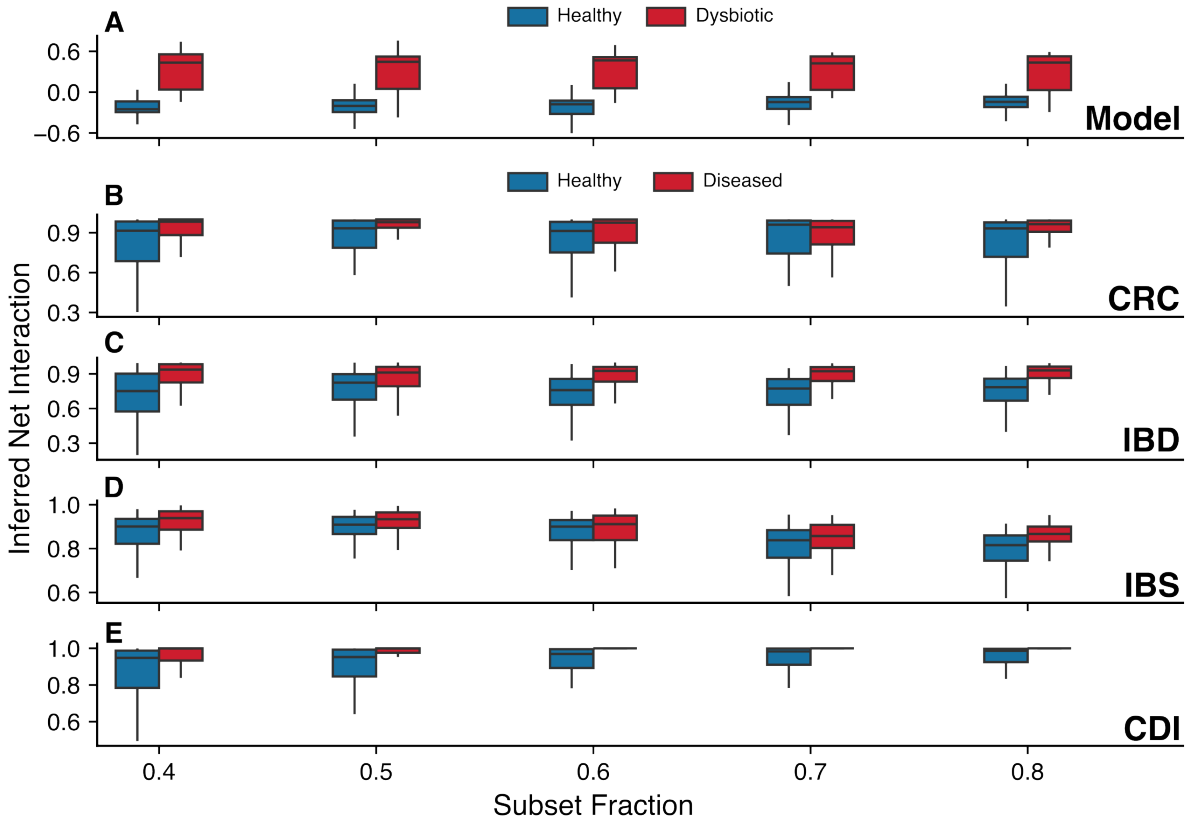

Figure S10: **Inferred net interaction using *BEEM-Static*.** Inferred net interaction estimated with the *BEEM-Static* algorithm (*105*) (see (*86*)) for different subsampling fractions of the original dataset for the model (**A**), CRC (**B**), IBD (**C**), IBS (**D**), and CDI (**E**). Across all 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 (500 bootstraping-generated samples).

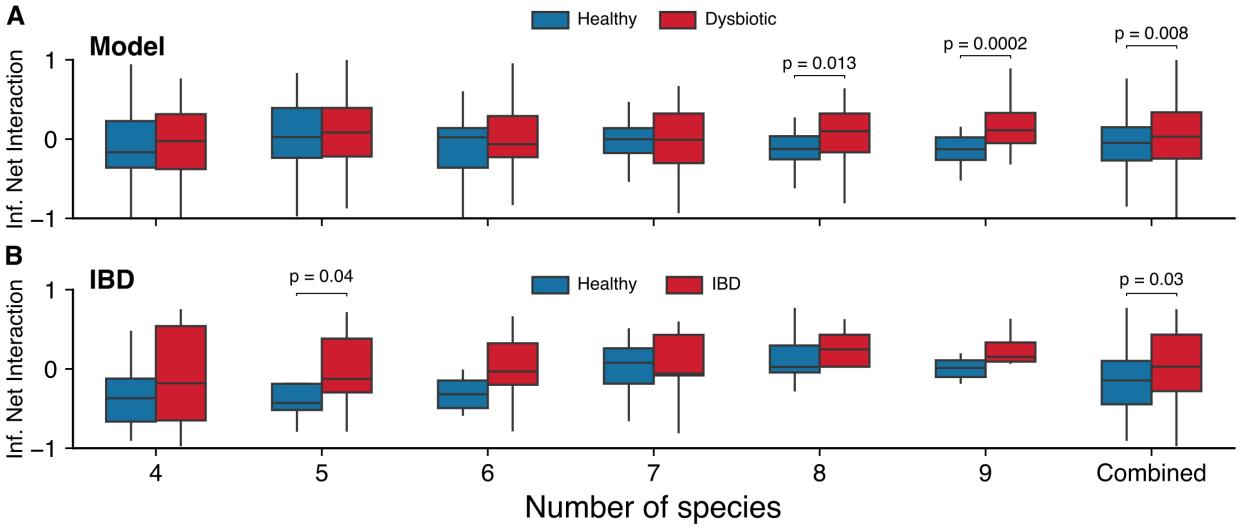

Figure S11: **Inferred net interaction using *LIMITS*.** Inferred net interaction computed with the *LIMITS* algorithm (*106*) (see (*86*)) for different numbers of species in the model (**A**) and for the IBD dataset (**B**). In both cases, dysbiotic states in the model and diseased states in IBD patients show a higher inferred net interaction compared to their respective healthy counterparts, indicating a shift towards a more cooperative microbial network.

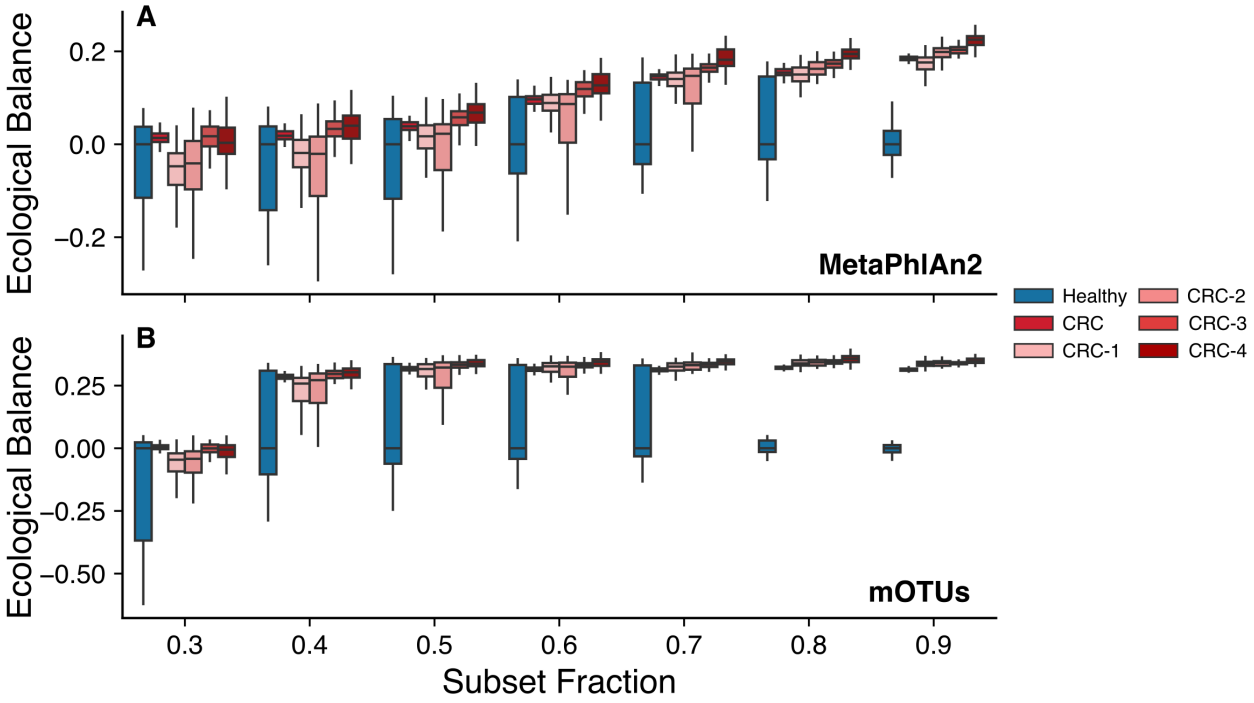

Figure S12: **The ecological network balance index (ENBI) is robust across different taxonomic profiling methodologies.** ENBI calculated for different subsampling fractions of the original CRC datasets using two distinct taxonomic profiling methodologies: *MetaPhlAn2* (*122*) (**A**) and *mOTUs* (*123*) (**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).