RESEARCH ARTICLE

iMeta WILEY

Predicting microbiome compositions from species assemblages through deep learning

Sebastian Michel-Mata[1,2] (ORCID) | Xu-Wen Wang[3] (ORCID) | Yang-Yu Liu[3] (ORCID) |
Marco Tulio Angulo[4]

[1]Center for Applied Physics and Advanced Technology, Universidad Nacional Autónoma de México, Juriquilla, Mexico

[2]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey, USA

[3]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA

[4]CONACyT—Institute of Mathematics, Universidad Nacional Autónoma de México, Juriquilla, Mexico

Correspondence
Yang-Yu Liu, Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA.
Email: yyl@channing.harvard.edu

Marco Tulio Angulo, CONACyT—Institute of Mathematics, Universidad Nacional Autónoma de México, Juriquilla, Mexico.
Email: mangulo@im.unam.mx

Abstract

Microbes can form complex communities that perform critical functions in maintaining the integrity of their environment or their hosts' wellbeing. Rationally managing these microbial communities requires improving our ability to predict how different species assemblages affect the final species composition of the community. However, making such a prediction remains challenging because of our limited knowledge of the diverse physical, biochemical, and ecological processes governing microbial dynamics. To overcome this challenge, we present a deep learning framework that automatically learns the map between species assemblages and community compositions from training data only, without knowing any of the above processes. First, we systematically validate our framework using synthetic data generated by classical population dynamics models. Then, we apply our framework to data from in vitro and in vivo microbial communities, including ocean and soil microbiota, *Drosophila melanogaster* gut microbiota, and human gut and oral microbiota. We find that our framework learns to perform accurate out-of-sample predictions of complex community compositions from a small number of training samples. Our results demonstrate how deep learning can enable us to understand better and potentially manage complex microbial communities.

KEYWORDS

deep learning, microbiome composition, species assemblage

Highlights

- A deep learning framework was developed to predict community compositions from species assemblages.
- The framework does not require knowing any microbial dynamics.
- Validation in various data shows accurate predictions.

INTRODUCTION

Microbes can form complex multispecies communities that perform critical functions in maintaining the integrity of their environment [1,2] or the well-being of their hosts [3–6]. For example, microbial communities play key roles in nutrient cycling in soils [7] and crop growth [8]. In humans, the gut microbiota plays important roles in our nutrition [9], immune system response [10], pathogen resistance [11], and even our central nervous system response [5]. Still, species invasions (e.g., pathogens) and extinctions (e.g., due to antibiotic administration) produce changes in the species assemblages that may shift these communities to undesired compositions [12]. For instance, antibiotic administrations can shift the human gut microbiota to compositions making the host more susceptible to recurrent infections by pathogens [13]. Similarly, intentional changes in the species assemblages, such as by using fecal microbiota transplantations, can shift back these communities to desired "healthier" compositions [14,15]. Therefore, improving our ability to rationally manage these microbial communities requires that we can predict changes in the community composition based on changes in species assemblages [16]. Building these predictions would also reduce managing costs, helping us to predict which changes in the species' assemblages are more likely to yield a desired community composition. Unfortunately, making such a prediction remains challenging because of our limited knowledge of the diverse physical [17], biochemical [18], and ecological [19,20] processes governing the microbial dynamics.

To overcome the above challenge, we present a deep learning framework that automatically learns the map between species assemblages and community compositions from training data only, without knowing the underlying microbial dynamics. We systematically validated our framework using synthetic data generated by classical ecological dynamics models, demonstrating its robustness to changes in the system dynamics and measurement errors. Then, we applied our framework to real data of both in vitro and in vivo communities, including ocean and soil microbial communities [21,22], *Drosophila melanogaster* gut microbiota [23], and human gut [24] and oral microbiota [25]. Across these diverse microbial communities, we find that our framework learns to predict accurate out-of-sample compositions given a few training samples. Our results show how deep learning can be an enabling ingredient for understanding and managing complex microbial communities.

PREDICTING MICROBIOME COMPOSITIONS FROM SPECIES ASSEMBLAGES

Consider the pool $\Omega = \{1, ..., N\}$ of all microbial species (or taxa) that can inhabit an ecological habitat of interest, such as the human gut. A microbiome sample obtained from this habitat can be considered as a local community assembled from Ω with a particular *species assemblage*. The species assemblage of a sample is characterized by a binary vector $z \in \{0, 1\}^N$, where its ith entry z_i satisfies $z_i = 1$ (or $z_i = 0$) if the ith species is present (or absent) in this sample. Each sample is also associated with a *composition* vector $p \in \Delta^N$, where its ith entry p_i is the relative abundance of the ith species, and $\Delta^N = \left\{ p \in \mathbb{R}_{\geq 0}^N \big| \sum_i p_i = 1 \right\}$ is the probability simplex. Therefore, our problem can be formalized as learning the map

$$\varphi : z \in \{0, 1\}^N \longmapsto p \in \Delta^N, \qquad (1)$$

which assigns the composition vector $p = \varphi(z)$ based on the species assemblage z. Note that the above map depends on many physical, biochemical, and ecological processes influencing the dynamics of microbial communities. These processes include the spatial structure of the ecological habitat [17], the chemical gradients of available resources [18], and inter/intraspecies interactions [20], among many others. Therefore, our limited knowledge of all these processes for most microbial communities renders the map of Equation (1) highly uncertain.

Next, we show it is possible to predict the microbial composition from species assemblage without knowing the mechanistic details of all the above processes. Our approach consists in learning the map φ directly from a data set $\mathfrak{D}$ with S microbiome samples. We arrange each of those samples as a pair (z, p) satisfying the map of Equation (1), see Figure 1A. Note that microbiome samples are readily available using standard metagenomic sequencing techniques.

Conditions for predicting compositions from species assemblages

To ensure that the problem of learning φ from $\mathfrak{D}$ is mathematically well-posed, we make the following assumptions. First, we assume that the species pool in the habitat has universal dynamics [26] (i.e., different local communities of this habitat can be described by the same population dynamics model with the same parameters). This assumption is necessary because, otherwise, the
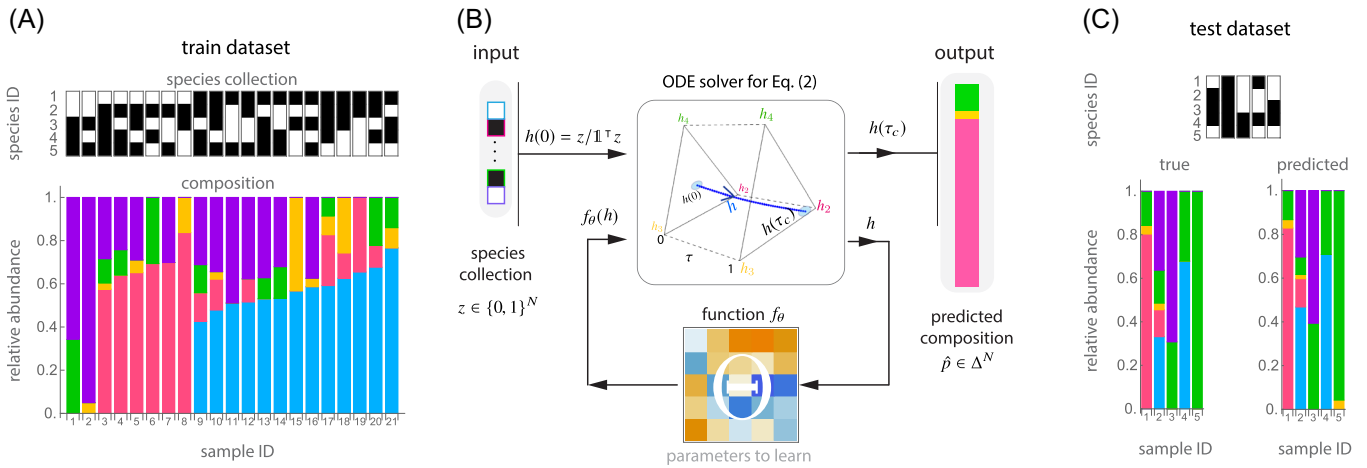
FIGURE 1 A deep learning framework to predict microbiome compositions from species assemblages. We illustrate this framework using experimental data from a pool of $N = 5$ bacterial species in *Drosophila melangaster* gut microbiota [23]: *Lactobacillus plantarum* (blue), *Lactobacillus brevis* (pink), *Acetobacter pasteurianus* (yellow), *Acetobacter tropicalis* (green), and *Acetobacter orientalis* (purple). (A) We randomly split this data set into training ($\mathfrak{D}_1$) and test ($\mathfrak{D}_2$) data sets, which contain 80% and 20% of the samples, respectively. Each data set contains pairs (z, p) with the species assemblage $z \in \{0, 1\}^N$ (top) and its corresponding composition $p \in \Delta^N$ (bottom) from each sample. (B) To predict compositions from species assemblages, our compositional neural ordinary differential equation (cNODE) framework consists of a solver for the ordinary differential equation shown in Equation (2), together with a chosen parametrized function f_θ. During training, the parameters θ are adjusted to learn to predict the composition $\hat{p} \in \Delta^N$ of the species assemblage $z \in \{0, 1\}^N$ in $\mathfrak{D}_1$. (C) After training, the performance is evaluated by predicting the composition of never-seen-before species assemblages in the test data set $\mathfrak{D}_2$. In this experimental microbiota, cNODE learned to perform accurate predictions of the composition in the test data set. For example, in the assemblage of species 3 and 4 (sample 26), cNODE correctly predicts that the composition is strongly dominated by a single species

map φ does not exist, implying that predicting community compositions from species assemblages has to be done in a sample-specific manner, which is a daunting task. The universal dynamics assumption will be satisfied when samples in the data set were collected from similar environments. Indeed, in this case, the environmental factors can be treated as roughly fixed and hence need not be used for composition prediction. For in vitro communities, the universal dynamics assumption is satisfied if samples were collected from the same experiment or multiple experiments but with very similar environmental conditions. For in vivo communities, empirical evidence indicates that the human gut and oral microbiota of healthy adults, as well as certain environment microbiota, display strong universal dynamics [26].

Second, we assume that the compositions of the collected samples represent steady states of the microbial communities. This assumption is natural because the map φ is not well defined for highly fluctuating microbial compositions. We note that observational studies of host-associated microbial communities such as the human gut microbiota indicate that they remain close to stable steady states in the absence of drastic dietary change or antibiotic administrations [24,27,28].

Finally, we assume that for each species assemblage $z \in \{0, 1\}^N$ there is a unique steady-state composition $p \in \Delta^N$. In particular, this assumption requires that true

multistability does not exist for the species pool (or any subset of it) in this habitat. This assumption is required because, otherwise, the map φ is not injective, and the prediction of community compositions becomes mathematically ill-defined.

In practice, we expect that the above three assumptions cannot be strictly satisfied. Therefore, any algorithm that predicts microbial compositions from species assemblages needs to be systematically tested to ensure its robustness against errors due to the violation of such approximations. Note that we can a priori check if a microbiome data set satisfies the universal dynamics assumption using the Dissimilarity-Overlap analysis [26]. Yet, it is mathematically challenging to a priori check if the other two assumptions are satisfied for real data. Nevertheless, the ability to accurately predict microbiome compositions from species assemblage is a posteriori evidence of the validity of the above three assumptions.

Learning to predict species compositions

Consider building a map $\hat{\varphi}_\theta : \{0, 1\}^N \rightarrow \Delta^N$, parametrized by $\theta \in \mathbb{R}^p$, giving the predicted composition $\hat{p} = \hat{\varphi}_\theta(z)$ associated with the species assemblage z. Under the above assumptions, we can in principle learn the map of Equation (1) from the data set $\mathfrak{D}$ by training $\hat{\varphi}_\theta$ (i.e.,

adjusting its parameters θ to ensure that $\hat{\varphi}_\theta$ approximates φ). Existing deep learning network architectures and training methods [29,30], such as ResNet [31] trained with a gradient descent algorithm, are natural candidates to solve this problem (Methods Section). We found that it is possible to train a ResNet architecture for predicting microbiome compositions in simple cases like small in vitro communities (Supporting Information Note S.2.1). But for large in vivo communities like the human gut microbiota, ResNet does not perform very well (Figure S1). The poor performance of ResNet is likely due to a vanishing gradient problem during training [32]. Namely, the ResNet architecture must satisfy two restrictions that are very particular to the map of Equation (1). First, the predicted compositions $\hat{p}$ must be compositional (i.e., $\hat{p} \in \Delta^N$). Second, the predicted relative abundance of any absent species in the assemblage must be identically zero (i.e., $z_i = 0$ should imply that $\hat{p}_i = 0$).

To overcome the limitations of traditional deep learning frameworks based on neural networks (such as ResNet) in predicting microbial compositions from species assemblages, we developed cNODE (compositional Neural Ordinary Differential Equation, see Methods Section and Figure 1B). We design the cNODE framework using the notion of Neural Ordinary Differential Equations, which can be interpreted as a continuous limit of ResNet architecture [33]. Crucially, the architecture and initialization of cNODE ensure that the above two restrictions are satisfied by construction. Furthermore, cNODE's architecture naturally circumvents the typical difficulties of handling zero values associated with compositional data analysis. Zero abundance values often occur in human microbiome datasets because of their highly personalized compositions across hosts (i.e., different individuals tend to have different species assemblages). To evaluate the prediction error of cNODE, one can choose any dissimilarity measure between the predicted and actual compositions related to a given species assemblage. Once this dissimilarity measure is selected, we train cNODE using a meta-learning algorithm for a given number of epochs to minimize the average prediction error in a *training* data set $\mathfrak{D}_1$ (Methods Section). Using this meta-learning algorithm improves the ability of cNODE for predicting the composition of never-seen-before species assemblages. Once trained, we evaluate the performance of cNODE by calculating its average prediction error in a *test* data set $\mathfrak{D}_2$ containing samples not used during the training.

Figure 1 illustrates the application of cNODE in a small experimental community of $N = 5$ bacterial species of *Drosophila melanogaster* microbiota studied by Gould et al. [23]. The data set $\mathfrak{D}$ obtained from this study has $S = 26$ samples (Methods Section). To illustrate the potential of cNODE, we consider a training data set of 21 randomly chosen samples (Figure 1A). As explained before, we arrange each training sample as a pair of "species assemblage" z (top) and "composition" p (bottom). Once trained, the main use of cNODE is to predict the composition of "never-seen-before" species assemblages —namely, "test assemblages" that are not in the training data set. To evaluate the performance of cNODE for predicting such test assemblages, we use as test data set the remaining five experimental samples not included during training. Figure 1C shows that the trained cNODE predicts accurate compositions for the test species assemblages. For example, cNODE predicts that in the assemblage of species 3 with species 4 (which was not used for training), species 3 will become nearly extinct. This prediction agrees well with the actual experimental result (sample 26 in Figure 1C).

RESULTS

In silico validation of cNODE with large species pools

We first evaluated cNODE's performance using in silico microbiome samples generated as steady-state compositions of pools with $N = 100$ species and Generalized Lotka-Volterra (GLV) population dynamics (Methods Section). We characterize the population dynamics of a species pool using two parameters. First, the *connectivity* $C > 0$, characterizing how likely is that two species in the pool interact directly. Second, the *typical interaction strength* $\sigma \geq 0$, characterizing the typical effect of one species over the per-capita growth rate of another species if they interact. Different habitats where the species pool is assembled are thus represented by different parameters (C, σ). Note that, despite its simplicity, the GLV model successfully describes the population dynamics of microbial communities in diverse environments, from the soil [34] and lakes [35] to the human gut [11,36,37].

Figure 2A shows the performance of cNODE during training. The training and test datasets have $S = N$ samples for this panel. Note that the training prediction error decreases with the number of training epochs, especially for low values of σ. Interestingly, the test prediction error reaches a plateau after sufficient training epochs, regardless of the value of σ. This plateau implies that cNODE was adequately trained with low overfitting. Note that the plateau's value increases with σ (i.e., the test prediction error increases). This result remains valid for different training data set sizes and different values for the parameters (C, σ). In all these cases, the test prediction error reaches a plateau whose value increases both by increasing C (Figure 2B) or σ (Figure 2C). But, crucially, such an increase can be compensated by increasing the number of samples in the training data set.
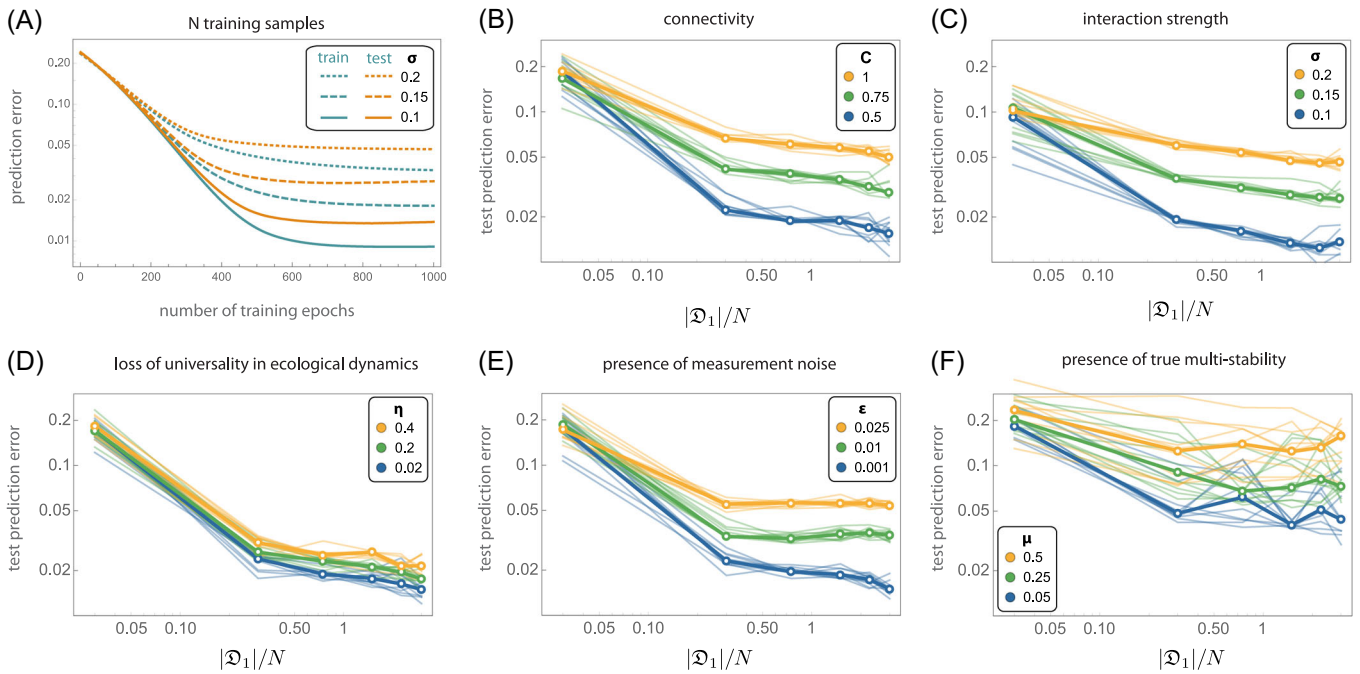
FIGURE 2 In silico validation of compositional neural ordinary differential equation (cNODE). Results are for pools of $N = 100$ species with Generalized Lotka-Volterra population dynamics (A–E) or population dynamics model with nonlinear functional responses that admits true multistability (F). The population dynamics is characterized by two parameters: the connectivity $C > 0$ and the typical interaction strength $\sigma \geq 0$. In panels B–F, thin lines represent the prediction errors for 10 validations of training cNODE with a different data set. Mean errors are shown with thick lines. (A) Training cNODE using $S = N$ samples with connectivity $C = 0.5$ and different typical interaction strengths σ. (B) Performance of cNODE for in-silico data sets with $\sigma = 0.1$ and different connectivity C. (C) Performance of cNODE for in-silico datasets with $C = 0.5$ and different interaction strengths σ. (D) Performance of cNODE for in silico data sets with nonuniversal dynamics. The lack of universal dynamics is quantified by the value of η. For all datasets, $\sigma = 0.1$ and $C = 0.5$. (E) Performance of cNODE for in-silico data sets with measurement errors quantified by ε. For all data sets, $\sigma = 0.1$ and $C = 0.5$. (F) Performance of cNODE for in-silico data sets with multiple interior equilibria, quantified by the probability $\mu \in [0, 1]$ of finding multiple equilibria. For all data sets, $C = 0.5$, $\sigma = 0.1$

This result implies that, in general, cNODE requires a larger number of training samples in species pools with higher connectivity or higher typical interaction strength between species. Overall, these results suggest that using $S = 2N$ or more training samples is enough to adequately train cNODE, regardless of the habitat type. In this case, we also observe a high correlation between the true and predicted compositions in the test data set, as expected from a low test prediction error (Figure S2).

To systematically evaluate the robustness of cNODE against violation of its three key assumptions, we performed three types of validations. In the first validation, we generated datasets that violate the assumption of universal dynamics (Methods Section). In this case, if two species interact, the effect of one species over the per-capita growth rate of the other species changes on average by $\eta \geq 0$ among samples in the data set. Therefore, the value $\eta = 0$ corresponds to universal dynamics, and larger values of η correspond to more significant losses of universal dynamics. We find that cNODE is robust against universality loss as its asymptotic

prediction error changes continuously and maintains a reasonably low test prediction error up to $\eta = 0.4$ (Figure 2D). cNODE is also robust to losses of universal dynamics that occur when species interact with different species in a sample-specific manner (Figure S3).

In the second validation, we evaluated the robustness of cNODE against measurement noises in the relative abundance of species (Methods Section). We characterize the noise intensity by a constant $\varepsilon \geq 0$. The measurement noise may cause some absent species to be measured as present and vice-versa. We find that cNODE performs adequately up to $\varepsilon = 0.025$ (Figure 2E).

In the final validation, we generated datasets with true multistability by simulating a population dynamics model with nonlinear functional responses (Methods Section). For each species assemblage, these functional responses generate two interior equilibria in different "regimes": one regime with low biomass and the other with high biomass. Therefore, each species assemblage can have two associated compositions. We built training datasets by choosing a fraction $(1 - \mu)$ of samples from the first regime and the

rest from the second regime. We find that cNODE is robust enough to provide reasonable predictions up to $\mu = 0.2$ (Figure 2F).

cNODE predicts microbiome compositions in real microbial communities

We evaluated cNODE using six microbiome datasets of different habitats (Supporting Information Note S4). The first data set consists of $S = 275$ samples [38] of the ocean microbiome at the phylum taxonomic level, resulting in $N = 73$ different taxa. The second data set consists of $S = 26$ in vivo samples of *Drosophila melanogaster* gut microbiota of $N = 5$ species [23], as described in Figure 1. The third data set has $S = 93$ samples of in vitro communities of $N = 8$ soil bacterial species [21]. The fourth data set contains $S = 113$ samples of the Central Park soil microbiome [22] at the phylum level ($N = 36$ phyla). The fifth data set contains $S = 150$ samples of the human oral microbiome [25] at the genus level ($N = 73$ genera). The final data set has $S = 106$ samples of the human gut microbiome from the Human Microbiome Project [24] at the genus level ($N = 58$ genera). Note that for each data set, to ensure cNODE has enough training samples, we chose to work at a specific taxonomic level so that the number of samples $S \geq 2N$, where N is the total number of taxa at the specific taxonomic level. Note that, based on the Dissimilarity-Overlap analysis, all the six microbiome datasets display the signature of universal microbial dynamics to some extent (Supporting Information Note S4.5 and Figure S4).

To evaluate cNODE, we performed the leave-one-out cross-validation on each data set (Methods Section). The median test prediction errors were 0.06, 0.066, 0.079, 0.107, 0.211, and 0.242 for the six datasets, respectively (Figure 3A). These errors are consistent with the strength of universality observed in each data set. To understand the meaning of these errors, for each data set we inspected five pairs $(p, \hat{p})$ corresponding to the observed and out-of-sample predicted composition of five samples. We chose the five samples based on their test prediction error. Specifically, we selected those samples with the minimal error, close to the first quartile, closer to the median, closer to the third quartile, and with the maximal error (columns in Figure 3B–G, from left to right). We found that samples with errors below the third quartile provide acceptable predictions (left three columns in Figure 3B–G), while samples with errors close to the third quartile or with the maximal error do demonstrate salient differences between the observed and predicted compositions (right two columns in Figure 3B–G). Note that in the

sample with largest error of the human gut data set (Figure 3G, rightmost column), the observed composition is dominated by *Prevotella* (pink) while the predicted sample is dominated by *Bacteroides* (blue). This drastic difference is likely due to different dietary patterns [39]. These results also confirm that $2N$ or more training samples are enough to adequately train cNODE, regardless of the habitat type. Note that using other taxonomic levels in these experimental datasets may change the performance of cNODE because it will effectively change the sample size.

DISCUSSION

cNODE is a deep learning framework to predict microbial compositions from species assemblages only. We validated its performance using in silico, in vitro, and in vivo microbial communities, finding that cNODE learns to perform accurate out-of-sample predictions using a few training samples. Classic methods for predicting species abundances in microbial communities use inference based on population dynamics models [21,36,40,41]. However, these methods typically require high-quality time-series data of species absolute abundances, which can be difficult and expensive to obtain in vivo microbial communities. cNODE circumvents needing absolute abundances or time-series data. However, compared to the classic methods, the cost to pay is that cNODE cannot be mechanistically interpreted because of the lack of identifiability inherent to compositional data [42,43]. We also found that cNODE can outperform existing deep-learning architectures like ResNet, specially when predicting the composition of large in-vivo microbiomes. Recently, Maynard et al. [44] proposed a statistical method to predict the steady-state abundance in ecological communities [44]. This method requires absolute abundance data of species, which are not available in most microbiome datasets. cNODE can outperform this statistical method despite using only relative abundances (Supporting Information Note S6). See also Supporting Information Note S5 and Figure S5 for a discussion of how our framework compares to other related works.

Deep learning techniques are actively applied in microbiome research [45–53], such as for classifying samples that shifted to a diseased state [54], predicting infection complications in immunocompromised patients [55], or predicting the temporal or spatial evolution of certain species collection [56,57]. However, to the best of our knowledge, the potential of deep learning for predicting the effect of changing species assemblages was not explored nor validated before. Our framework, based
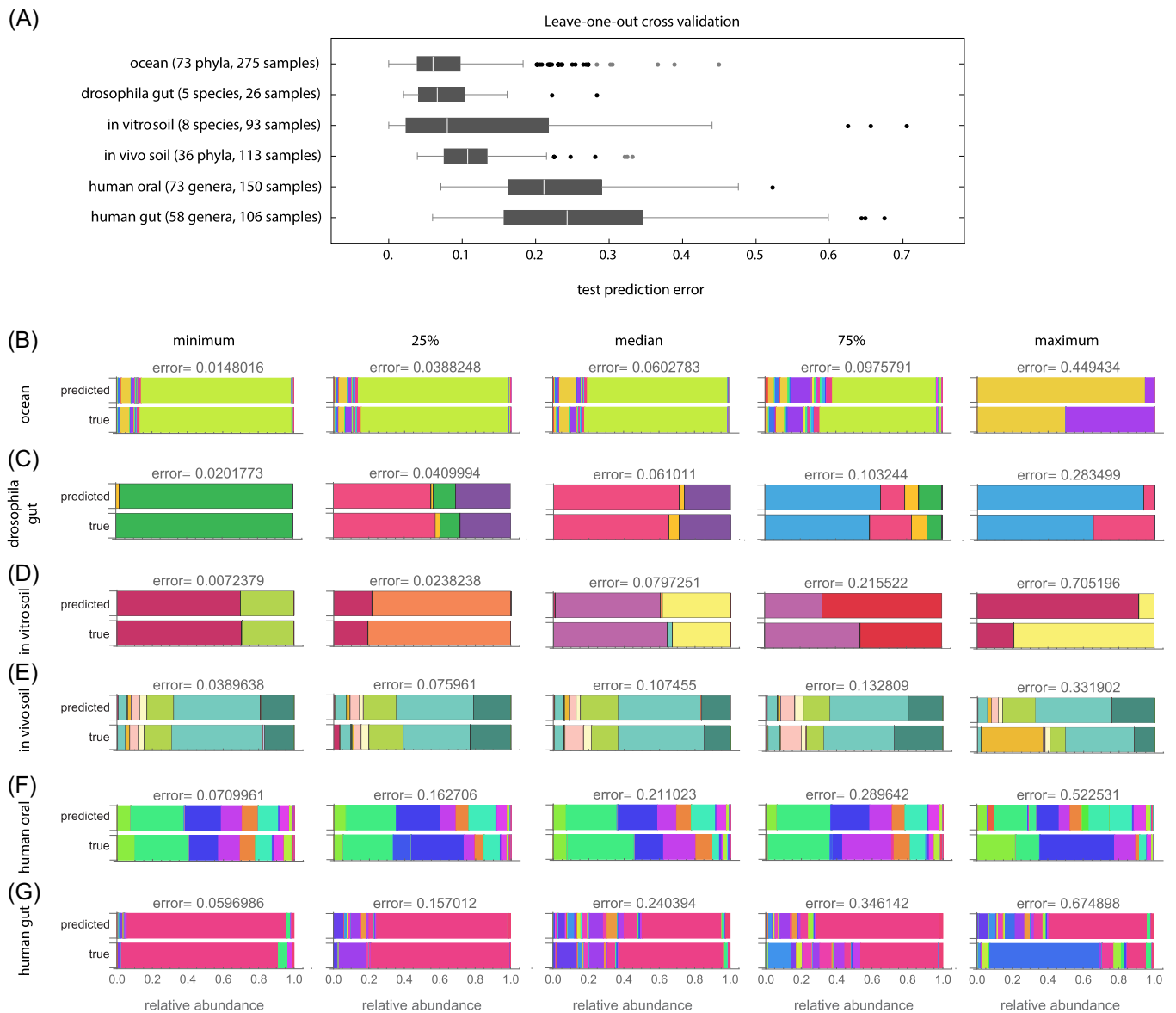
FIGURE 3 Predicting the composition of real microbiomes from species assemblages. Results of the compositional neural ordinary differential equation applied to six experimental microbial communities using a leave-one-out crossvalidation. (A) Prediction error obtained from a leave-one-out crossvalidation of each data set. (B–G) For each data set, we show the true and predicted compositions corresponding to the minimal prediction error, closer to the first quartile, median, closer to the third quartile, maximum prediction error (including outliers). All compositions shown in (B–G) are out-of-sample predictions

on the notion of neural ODE [33], is a baseline that could be improved by incorporating additional information. For example, incorporating available environmental information such as pH, temperature, age, BMI, body-site, and host's diet could enhance the prediction accuracy. This additional information would help us predict the species present in different environments. Adding "hidden variables" such as the unmeasured total biomass or unmeasured resources to our ODE will enhance the expressivity of the cNODE [58,59], but this may result in more challenging training. Finally, if available, knowledge of the genetic similarity between species can be leveraged into the loss function by using the phylogenetic Wasserstein distance [60] that provides a well-defined gradient [61].

We anticipate that a useful application of our framework is to predict if by adding some species collection to a local community we can bring the abundance of target species below the practical extinction threshold. Thus, given a local community containing the target (and potentially pathogenic) species, we could use a greedy optimization algorithm to identify a minimal collection of species to add such that our architecture predicts that they will decolonize the target species.

Our framework has a few limitations. For example, cNODE cannot accurately predict the abundance of taxa that have never been observed in the training data set. An additional limitation of our current architecture is that it assumes that true multistability does not exist—namely, a community with a given species assemblage permits only one stable steady-state, where each species in the collection has a positive abundance. For complex microbial communities such as the human gut microbiota, the highly personalized species collections make it very difficult to decide if true multistability exists or not. We could extend our framework to handle multistability by predicting a probability density function for the abundance of each species. True multistability would correspond to predicting a multimodal density function in such a case. Datasets with insufficient sequencing depth or coverage can produce samples with "fake" multistability, leading to prediction errors that our framework cannot resolve. Indeed, the in-silico validation of cNODE in Figure 2 indicates that measurement errors can significantly degrade the performance of cNODE.

In conclusion, the many species and the complex, uncertain dynamics that microbial communities exhibit, have been fundamental obstacles in our ability to learn how they respond to alterations, such as removing or adding species. Moving this field forward may require losing some ability to interpret the mechanism behind their response. In this sense, deep learning methods could enable us to rationally manage and predict complex microbial communities' dynamics.

METHODS

A ResNet architecture for predicting microbiome compositions from species assemblages

As a top-rated tool in image processing, ResNet is a cascade of $L \geq 1$ hidden layers where the state $h_\ell \in \mathbb{R}^N$ of the ℓth hidden layer satisfies $h_\ell = h_{\ell-1} + f_\theta(h_{\ell-1})$, $\ell = 1, ..., L$, for some parametrized function f_θ with parameters θ. These hidden layers are plugged to the input $h_0 = g_{in}(z)$ and the output $\hat{p} = g_{out}(h_L)$ layers, where g_{in} and g_{out} are some functions. Crucially, for our problem, any architecture must satisfy two restrictions: (1) vector $\hat{p}$ must be compositional (i.e., $\hat{p} \in \Delta^N$); and (2) the predicted relative abundance of any absent species must be identically zero (i.e., $z_i = 0$ should imply that $\hat{p}_i = 0$). Simultaneously satisfying both restrictions requires that the output layer is a normalization of the form $\hat{p}_i = z_i h_{L,i} / \sum_j z_j h_{L,j}$, and that f_θ is a non-negative function (because $h_L \geq 0$ is required to

ensure the normalization is correct). This result is likely due to the normalization of the output layer, which challenges the training of neural networks because of vanishing gradients [30]. The vanishing gradient problem is often solved by using other normalization layers such as the `softmax` or `sparsemax` layers [62]. However, we cannot use these layers because they do not satisfy the second restriction. We also note that ResNet becomes a universal approximation only in the limit $L \to \infty$, which again complicates the training [32].

The cNODE architecture

In cNODE, an input species assemblage $z \in \{0, 1\}^N$ is first transformed into the initial condition $h(0) = z/\mathbb{1}^\intercal z \in \Delta^N$, where $\mathbb{1} = (1, ..., 1)^\intercal \in \mathbb{R}^N$ (left in Figure 1B). This initial condition is used to solve the set of nonlinear ODEs

$$\frac{dh(\tau)}{d\tau} = h(\tau) \odot [f_\theta(h(\tau)) - \mathbb{1} h(\tau)^\intercal f_\theta(h(\tau))]. \quad (2)$$

Here, the independent variable $\tau \geq 0$ represents a virtual "time". The expression $h \odot v$ is the entry-wise multiplication of the vectors $h, v \in \mathbb{R}^N$. The function $f_\theta : \Delta^N \to \mathbb{R}^N$ can be any continuous function parametrized by θ. For example, it can be the linear function $f_\theta(h) = \Theta h$ with parameter matrix $\Theta \in \mathbb{R}^{N \times N}$ (bottom in Figure 1B), or a more complicated function represented by a feedforward deep neural network. Note that Equation (2) is a general form of the replicator equation—a canonical model in evolutionary game theory [63]—with f_θ representing the fitness function. By choosing a final integration "time" $\tau_c > 0$, Equation (2) is numerically integrated to obtain the prediction $\hat{p} = h(\tau_c)$ that is the output of cNODE (right in Figure 1B). We choose $\tau_c = 1$ without loss of generality, as τ in Equation (2) can be rescaled by multiplying f_θ by a constant. The cNODE thus implements the map

$$\hat{\varphi}_\theta : z \in \{0, 1\}^N \longmapsto \hat{p} \in \Delta^N, \quad (3)$$

taking an input species assemblage z to the predicted composition $\hat{p}$ (see Supporting Information Note S.1 for implementation details). Note that Equation (2) is key to cNODE because its architecture guarantees that the two restrictions imposed before are naturally satisfied. Namely, $\hat{p} \in \Delta^N$ because the conditions $h(0) \in \Delta^N$ and $\mathbb{1}^\intercal dh/d\tau = 0$ imply that $h(\tau) \in \Delta^N$ for all $\tau \geq 0$. Additionally, $z_i = 0$ implies $\hat{p}_i = 0$ because $h(0)$ and z have the same zero pattern, and the right-hand side of Equation (2) is entry-wise multiplied by h.

Training cNODE

We train cNODE by adjusting the parameters θ to approximate φ with $\hat{\varphi}_{\theta}$. To do this, we first choose a distance or dissimilarity measure $d(p, q)$ to quantify how dissimilar are two compositions $p, q \in \Delta^N$. We choose the Bray-Curtis [64] dissimilarity to present our results, however, the performance of cNODE is quite robust to the specific distance or dissimilarity measure used (Figure S6). Specifically, for a data set $\mathfrak{D}_i \subseteq \mathfrak{D}$, we use as loss function the *prediction error*

$$E(\mathfrak{D}_i) = \frac{1}{|\mathfrak{D}_i|} \sum_{(z,p) \in \mathfrak{D}_i} d(p, \hat{\varphi}_{\theta}(z)). \tag{4}$$

Second, we randomly split the data set $\mathfrak{D}$ into training $\mathfrak{D}_1$ and test $\mathfrak{D}_2$ datasets. Next, we choose an adequate functional form for f_{θ}. In our experiments, we found that the linear function $f_{\theta}(h) = \Theta h$, $\Theta \in \mathbb{R}^{N \times N}$, provides accurate predictions for the composition of in silico, in vitro, and in vivo communities. Importantly, despite f_{θ} is linear, the map $\hat{\varphi}_{\theta}$ is nonlinear because it is the solution of the nonlinear ODE of Equation (2). Finally, we adjust the parameters θ by minimizing Equation (4) on $\mathfrak{D}_1$ using a gradient-based meta-learning algorithm [65]. This learning algorithm enhances the generalizability of cNODE (Supporting Information Note S1.2 and Figure S1). Training cNODE with a data set of 100 species, 100 training samples, and 100 epochs takes about 30 min on a Linux machine with six Intel Xeon CPUs (E7-4870 v2) @ 2.30 GHz.

Once trained, we calculate cNODE's test prediction error $E(\mathfrak{D}_2)$ that quantifies cNODE's performance in predicting the compositions of never-seen-before species assemblages. Test prediction errors could be due to a poor adjustment of the parameters (i.e., inaccurate prediction of the training set), low ability to generalize (i.e., inaccurate predictions of the test data set), or violations of our three assumptions (universal dynamics, steady-state samples, no true multistability).

Generating in-silico data for validating cNODE

We generated in silico data for validating cNODE as steady-state compositions of pools with N species and generalized Lotka-Volterra (GLV) population dynamics. The GLV model reads [66]:

$$\frac{dx_i(t)}{dt} = x_i(t)\left[r_i + \sum_{j=1}^{N} a_{ij} x_j(t)\right], i = 1, ..., N. \tag{5}$$

Above, $x_i(t)$ denotes the abundance of the ith species at time $t \geq 0$. The GLV model has as parameters the interaction matrix $A = (a_{ij}) \in \mathbb{R}^{N \times N}$, and the intrinsic growth-rate vector $r = (r_i) \in \mathbb{R}^N$. The parameter a_{ij} denotes the inter- (if $j \neq i$) or intra- (if $j = i$) species *interaction strength* of species j to the per-capita growth rate of species i. The parameter r_i is the intrinsic growth rate of species i. The interaction matrix A determines the ecological network $\mathcal{G}(A)$ underlying the species pool. Namely, this network has one node per species and edges $(j \rightarrow i) \in \mathcal{G}(A)$ if $a_{ij} \neq 0$. The *connectivity* $C \in [0, 1]$ of this network is the proportion of edges it has compared to the N^2 edges in a complete network.

To validate cNODE, we generated synthetic microbiome samples as steady-state compositions of GLV models with random parameters by choosing $a_{ij} \sim \text{Bernoulli}(C)\text{Normal}(0, \sigma)$ if $i \neq j$, $a_{ii} = -1$, and $r_i \sim \text{Uniform}[0, 1]$, for different values of connectivity C and characteristic inter-species interaction strength $\sigma > 0$ (Supporting Information Note S3).

Generating in silico data to test the robustness of cNODE

For this, given a "base" GLV model with parameters (A, r), we consider two forms of universality loss (Supporting Information Note S3). First, samples are generated using a GLV with the same ecological network but with those non-zero interaction strengths a_{ij} replaced by $a_{ij} + \text{Normal}(0, \eta)$, where $\eta > 0$ characterizes the changes in the typical interaction strength. Second, samples are generated using a GLV with slightly different ecological networks obtained by randomly rewiring a proportion $\rho \in [0, 1]$ of their edges.

In the second validation, we evaluated the robustness of cNODE against measurement noises in the relative abundance of species. For this, for each sample p, we first change the relative abundance of the ith species from p_i to $\max\{0, p_i + \text{Normal}(0, \varepsilon)\}$, where $\varepsilon \geq 0$ characterizes the measurement noise intensity. Then, we normalize the vector p to ensure it is still compositional, that is, $p \in \Delta^N$. Due to the measurement noise, some species that were absent may be measured as present and vice-versa.

In the third validation, we generated datasets with true multistability by simulating a population dynamics model with nonlinear functional responses (Supporting Information Note S3). For each species collection, these functional responses generate two interior equilibria in different "regimes": one regime with low biomass, and the other regime with high biomass. We then train cNODE with datasets obtained by choosing a fraction

$(1 - \mu)$ of samples from the first regime, and the rest from the second regime.

Validating cNODE using real microbiome data sets

To validate cNODE, we performed a leave-one-out cross-validation over real microbiome data sets (see descriptions on Supporting Information Note S4). For each data set, we measured the prediction error of cNODE using each sample as a test set and the rest of the samples as a training set. We repeated this procedure for different learning rates and mini-batch sizes and selected the hyperparameters that minimized the average prediction error over the samples (see Table S1).

CONFLICT OF INTERESTS
The authors declare that there are no conflict of interests.

AUTHOR CONTRIBUTIONS
Marco Tulio Angulo and Yang-Yu Liu conceived and designed the project. Sebastian Michel-Mata did the numerical analysis. Sebastian Michel-Mata and Xu-Wen Wang performed the real data analysis. All authors analyzed the results. Marco T. Angulo and Yang-Yu Liu wrote the manuscript. Sebastian Michel-Mata and Xu-Wen Wang edited the manuscript.

DATA AVAILABILITY STATEMENT
The data and code used in this study are available at https://github.com/michel-mata/cNODE.jl.

ORCID
Sebastian Michel-Mata https://orcid.org/0000-0003-3424-6669
Xu-Wen Wang http://orcid.org/0000-0001-7670-3544
Yang-Yu Liu https://orcid.org/0000-0003-2728-4907

REFERENCES
1. East, Roger. 2013. "Soil Science Comes to Life." *Nature* 501: S18. https://doi.org/10.1038/501S18a
2. Busby, Posy E., Chinmay Soman, Maggie R. Wagner, Maren L. Friesen, James Kremer, Alison Bennett, Mustafa Morsy, Jonathan A. Eisen, Jan E. Leach, and Jeffery L. Dangl. 2017. "Research Priorities for Harnessing Plant Microbiomes in Sustainable Agriculture." *PLOS Biology* 15: e2001793. https://doi.org/10.1371/journal.pbio.2001793
3. Thursby, Elizabeth, and Nathalie Juge. 2017. "Introduction to the Human Gut Microbiota." *Biochemical Journal* 474: 1823–36. https://doi.org/10.1042/BCJ20160510
4. Chittim, Carina L., Stephania M. Irwin, and Emily P. Balskus. 2018. "Deciphering Human Gut Microbiota-Nutrient Interactions: A Role for Biochemistry." *Biochemistry* 57: 2567–77. https://doi.org/10.1021/acs.biochem.7b01277
5. Cryan, John F., and Timothy G. Dinan. 2012. "Mind-altering Microorganisms: The Impact of the Gut Microbiota on Brain and Behaviour." *Nature Reviews Neuroscience* 13: 701. https://doi.org/10.1038/nrn3346
6. Arrieta, Marie-Claire, Leah T. Stiemsma, Nelly Amenyogbe, Eric M. Brown, and Brett Finlay. 2014. "The Intestinal Microbiome in Early Life: Health and Disease." *Frontiers in Immunology* 5: 427. https://doi.org/10.3389/fimmu.2014.00427
7. Fierer, Noah. 2017. "Embracing the Unknown: Disentangling the Complexities of the Soil Microbiome." *Nature Reviews Microbiology* 15: 579–90. https://doi.org/10.1038/nrmicro.2017.87
8. Mueller, Ulrich G., and Joel L. Sachs. 2015. "Engineering Microbiomes to Improve Plant and Animal Health." *Trends in Microbiology* 23: 606–17. https://doi.org/10.1016/j.tim.2015.07.009
9. Gill, Steven R., Mihai Pop, Robert T. DeBoy, Paul B. Eckburg, Peter J. Turnbaugh, Buck S. Samuel, Jeffrey I. Gordon, David A. Relman, Claire M. Fraser-Liggett, and Karen E. Nelson. 2006. "Metagenomic analysis of the human distal gut microbiome." *Science* 312: 1355–9. https://doi.org/10.1126/science.1124234
10. Hooper, Lora V., Dan R. Littman, and Andrew J. Macpherson. 2012. "Interactions Between the Microbiota and the Immune System." *Science* 336: 1268–73. https://doi.org/10.1126/science.1223490
11. Buffie, Charlie G., Vanni Bucci, Richard R. Stein, Peter T. McKenney, Lilan Ling, Asia Gobourne, Daniel No, Hui Liu, Melissa Kinnebrew, Agnes Viale, et al. 2015. "Precision Microbiome Reconstitution Restores Bile Acid Mediated Resistance to Clostridium Difficile." *Nature* 517: 205–8. https://doi.org/10.1038/nature13828
12. Costello, Elizabeth K., Keaton Stagaman, Les Dethlefsen, Brendan J. M Bohannan, and David A. Relman. 2012. "The Application of Ecological Theory Toward an Understanding of the Human Microbiome." *Science* 336: 1255–62. https://doi.org/10.1126/science.1224203
13. Lemon, Katherine P., Gary C. Armitage, David A. Relman, and Michael A. Fischbach. 2012. "Microbiota-Targeted Therapies: An Ecological Perspective." *Science Translational Medicine* 4: 137rv5. https://doi.org/10.1126/scitranslmed.3004183
14. Borody, Thomas J., Sudarshan Paramsothy, and Gaurav Agrawal. 2013. "Fecal Microbiota Transplantation: Indications, Methods, Evidence, and Future Directions." *Current Gastroenterology Reports* 15: 1–7. https://doi.org/10.1007/s11894-013-0337-1

15. Xiao, Yandong, Marco Tulio Angulo, Songyang Lao, Scott T. Weiss, and Yang-Yu Liu. 2020. "An Ecological Framework to Understand the Efficacy of Fecal Microbiota Transplantation." *Nature Communications* 11: 1–17. https://doi.org/10.1038/s41467-020-17180-x

16. Widder, Stefanie, Rosalind J. Allen, Thomas Pfeiffer, Thomas P. Curtis, Carsten Wiuf, William T. Sloan, Otto X. Cordero, Sam P. Brown, Babak Momeni, Wenying Shou, et al. 2016. "Challenges in Microbial Ecology: Building Predictive Understanding of Community Function and Dynamics." *The ISME Journal* 10: 2557. https://doi.org/10.1038/ismej.2016.45

17. Tropini, Carolina, Kristen A. Earle, Kerwyn Casey Huang, and Justin L. Sonnenburg. 2017. "The Gut Microbiome: Connecting Spatial Organization to Function." *Cell Host & Microbe* 21: 433–42. https://doi.org/10.1016/j.chom.2017.03.010

18. Quinn, Robert A., William Comstock, Tianyu Zhang, James T. Morton, Ricardo da Silva, Alda Tran, Alexander Aksenov, Louis-Felix Nothias, Daniel Wangpraseurt, Alexey V. Melnik, et al. 2018. "Niche Partitioning of a Pathogenic Microbiome Driven by Chemical Gradients." *Science Advances* 4: eaau1908. https://doi.org/10.1126/sciadv.aau1908

19. Kuramitsu, Howard K., Xuesong He, Renate Lux, Maxwell H. Anderson, and Wenyuan Shi. 2007. "Interspecies Interactions within Oral Microbial Communities." *Microbiology Molecular Biology Reviews* 71: 653–70. https://doi.org/10.1128/MMBR.00024-07

20. Coyte, Katharine Z., Jonas Schluter, and Kevin R. Foster. 2015. "The Ecology of the Microbiome: Networks, Competition, and Stability." *Science* 350: 663–6. https://doi.org/10.1126/science.aad2602

21. Friedman, Jonathan, Logan M. Higgins, and Jeff Gore. 2017. "Community Structure Follows Simple Assembly Rules in Microbial Microcosms." *Nature Ecology & Evolution* 1: 109. https://doi.org/10.1038/s41559-017-0109

22. Ramirez, Kelly S., Jonathan W. Leff, Albert Barberán, Scott Thomas Bates, Jason Betley, Thomas W. Crowther, Eugene F. Kelly, Emily E. Oldfield, E. Ashley Shaw, Christopher Steenbock, et al. 2014. "Biogeographic Patterns in Below-Ground Diversity in New York City's Central Park are Similar to Those Observed Globally." *Proceedings of the Royal Society B: Biological Sciences* 281: 20141988. https://doi.org/10.1098/rspb.2014.1988

23. Gould, Alison L., Vivian Zhang, Lisa Lamberti, Eric W. Jones, Benjamin Obadia, Nikolaos Korasidis, Alex Gavryushkin, Jean M. Carlson, Niko Beerenwinkel, and William B. Ludington. 2018. "Microbiome Interactions Shape Host Fitness." *Proceedings of the National Academy of Sciences* 115: E11951–60. https://doi.org/10.1073/pnas.1809349115

24. Huttenhower, Curtis, Dirk Gevers, Rob Knight, Sahar Abubucker, Jonathan H. Badger, Asif T. Chinwalla, Heather H. Creasy, Ashlee M. Earl, Michael G. FitzGerald, Robert S. Fulton, et al. 2012. "Structure, Function and Diversity of the Healthy Human Microbiome." *Nature* 486: 207. https://doi.org/10.1038/nature11234

25. Vogtmann, Emily, Jun Chen, Muhammad G. Kibriya, Amnon Amir, Jianxin Shi, Yu Chen, Tariqul Islam, Mahbubul Eunes, Alauddin Ahmed, Jabun Naher, et al. 2019. "Comparison of Oral Collection Methods for Studies of Microbiota." *Cancer Epidemiology and Prevention Biomarkers* 28: 137–43. https://doi.org/10.1158/1055-9965.EPI-18-0312

26. Bashan, Amir, Travis E. Gibson, Jonathan Friedman, Vincent J. Carey, Scott T. Weiss, Elizabeth L. Hohmann, and Yang-Yu Liu. 2016. "Universality of Human Microbial Dynamics." *Nature* 534: 259. https://doi.org/10.1038/nature18301

27. Lozupone, Catherine A., Jesse I. Stombaugh, Jeffrey I. Gordon, Janet K. Jansson, and Rob Knight. 2012. "Diversity, Stability and Resilience of the Human Gut Microbiota." *Nature* 489: 220. https://doi.org/10.1038/nature11550

28. Gibbons, Sean M., Sean M. Kearney, Chris S. Smillie, and Eric J. Alm. 2017. "Two Dynamic Regimes in the Human Gut Microbiome." *PLOS Computational Biology* 13: 1–20. https://doi.org/10.1371/journal.pcbi.1005364

29. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. "Deep Learning." *Nature* 521: 436. https://doi.org/10.1038/nature14539

30. Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. Cambridge, MA, USA: MIT Press. https://doi.org/10.1007/s10710-017-9314-z

31. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. "Deep Residual Learning for Image Recognition." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–8. Las Vegas, NV, USA. https://doi.org/10.1109/cvpr.2016.90

32. Lin, Hongzhou, and Stefanie Jegelka. 2018. "ResNet with One-Neuron Hidden Layers is a Universal Approximator." In *Advances in Neural Information Processing Systems,* 6169–78. Montreal, Canada: Curran Associates Inc. https://hdl.handle.net/1721.1/129326

33. Chen, Tian Qi, Yulia Rubanova, Jesse Bettencourt, and David K. Duvenaud. 2018. "Neural Ordinary Differential Equations." In *Advances in Neural Information Processing Systems*, 6571–83. Montreal, Canada: Curran Associates Inc. https://dl.acm.org/doi/abs/10.5555/3327757.3327764

34. Moore, John C., Peter C. de Ruiter, H. William Hunt, David C. Coleman, and Diana W. Freckman. 1996. "Microcosms and Soil Ecology: Critical Linkages Between Fields Studies and Modelling Food Webs." *Ecology* 77: 694–705. https://doi.org/10.2307/2265494

35. Dam, Phuongan, Luis L. Fonseca, Konstantinos T. Konstantinidis, and Eberhard O. Voit. 2016. "Dynamic Models of the Complex Microbial Metapopulation of Lake Mendota." *NPJ Systems Biology and Applications* 2: 1–7. https://doi.org/10.1038/npjsba.2016.7

36. Stein, Richard R., Vanni Bucci, Nora C. Toussaint, Charlie G. Buffie, Gunnar Rätsch, Eric G. Pamer, Chris Sander, and Joao B. Xavier. 2013. "Ecological Modeling from Time-series Inference: Insight into Dynamics and Stability of Intestinal Microbiota." *PLOS Computational Biology* 9: e1003388. https://doi.org/10.1371/journal.pcbi.1003388

37. Rao, Chitong, Katharine Z. Coyte, Wayne Bainter, Raif S. Geha, Camilia R. Martin, and Seth Rakoff-Nahoum. 2021. "Multikingdom Ecological Drivers of Microbiota Assembly in Preterm Infants." *Nature* 591: 633–8. https://doi.org/10.1038/s41586-021-03241-8

38. Laverock, Bonnie, Cindy J. Smith, Karen Tait, A. Mark Osborn, Steve Widdicombe, and Jack A. Gilbert. 2010. "Bioturbating

Shrimp Alter the Structure and Diversity of Bacterial Communities in Coastal Marine Sediments." *The ISME Journal* 4: 1531–44. https://doi.org/10.1038/ismej.2010.86

39. Wu, Gary D., Jun Chen, Christian Hoffmann, Kyle Bittinger, Ying-Yu Chen, Sue A. Keilbaugh, Meenakshi Bewtra, Dan Knights, William A. Walters, Rob Knight, et al. 2011. "Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes." *Science* 334: 105–108. https://doi.org/10.1126/science.1208344

40. Mounier, Jérôme, Christophe Monnet, Tatiana Vallaeys, Roger Arditi, Anne-Sophie Sarthou, Arnaud Hélias, and Françoise Irlinger. 2008. "Microbial Interactions Within a Cheese Microbial Community." *Applied and Environmental Microbiology* 74: 172–81. https://doi.org/10.1128/AEM.01338-07

41. Venturelli, Ophelia S., Alex V. Carr, Garth Fisher, Ryan H. Hsu, Rebecca Lau, Benjamin P. Bowen, Susan Hromada, Trent Northen, and Adam P. Arkin. 2018. "Deciphering Microbial Interactions in Synthetic Human Gut Microbiome Communities." *Molecular Systems Biology* 14: e8157. https://doi.org/10.15252/msb.20178157

42. Aitchison, John. 1994. "Principles of Compositional Data Analysis." *Lecture Notes-Monograph Series* 73–81. https://doi.org/10.1214/lnms/1215463786

43. Cao, Hong-Tai, Travis E. Gibson, Amir Bashan, and Yang-Yu Liu. 2017. "Inferring Human Microbial Dynamics from Temporal Metagenomics Data: Pitfalls and Lessons." *BioEssays* 39: 1600188. https://doi.org/10.1002/bies.201600188

44. Maynard, Daniel S., Zachary R. Miller, and Stefano Allesina. 2020. "Predicting Coexistence in Experimental Ecological Communities." *Nature Ecology & Evolution* 4: 91–100. https://doi.org/10.1038/s41559-019-1059-z

45. Reiman, Derek, Ahmed Metwally, and Yang Dai. 2017. "Using Convolutional Neural Networks to Explore the Microbiome in Engineering in Medicine and Biology Society (EMBC)." In *2017 39th Annual International Conference of the IEEE*, 4269–72. Jeju Island, South Korea: Curran Associates Inc. https://doi.org/10.1109/EMBC.2017.8037799

46. García-Jiménez, Beatriz, Jorge Muñoz, Sara Cabello, Joaquín Medina, and Mark D. Wilkinson. 2021. "Predicting Microbiomes Through a Deep Latent Space." *Bioinformatics* 37(10): 1444–51. https://doi.org/10.1093/bioinformatics/btaa971

47. Metwally, Ahmed A., Philip S. Yu, Derek Reiman, Yang Dai, Patricia W. Finn, and David L. Perkins. 2019. "Utilizing Longitudinal Microbiome Taxonomic Profiles to Predict Food Allergy via Long Short-Term Memory Networks." *PLOS Computational Biology* 15: e1006693. https://doi.org/10.1371/journal.pcbi.1006693

48. Sharma, Divya, Andrew D. Paterson, and Wei Xu. 2020. "TaxoNN: Ensemble of Neural Networks on Stratified Microbiome Data for Disease Prediction." *Bioinformatics* 36: 4544–50. https://doi.org/10.1093/bioinformatics/btaa542

49. Galkin, Fedor, Polina Mamoshina, Alex Aliper, Evgeny Putin, Vladimir Moskalev, Vadim N. Gladyshev, and Alex Zhavoronkov. 2020. "Human Gut Microbiome Aging Clock Based on Taxonomic Profiling and Deep Learning." *iScience* 23: 101199. https://doi.org/10.1016/j.isci.2020.101199

50. Asgari, Ehsaneddin, Kiavash Garakani, Alice C. McHardy, and Mohammad R. K. Mofrad. 2018. "MicroPheno: Predicting Environments and Host Phenotypes from 16S rRNA Gene Sequencing Using a K-mer Based Representation of Shallow Sub-Samples." *Bioinformatics* 34: i32–42. https://doi.org/10.1093/bioinformatics/bty652

51. Zhu, Qiang, Xingpeng Jiang, Qing Zhu, Min Pan, and Tingting He. 2019. "Graph Embedding Deep Learning Guides Microbial Biomarkers' Identification." *Frontiers in Genetics* 10: 1182. https://doi.org/10.3389/fgene.2019.01182

52. Le, Vuong, Thomas P. Quinn, Truyen Tran, and Svetha Venkatesh. 2020. "Deep in the Bowel: Highly Interpretable Neural Encoder-Decoder Networks Predict Gut Metabolites from Gut Microbiome." *BMC Genomics* 21: 1–15. https://doi.org/10.1186/s12864-020-6652-7

53. LaPierre, Nathan, Chelsea J.-T. Ju, Guangyu Zhou, and Wei Wang. 2019. "MetaPheno: A Critical Evaluation of Deep Learning and Machine Learning in Metagenome-Based Disease Prediction." *Methods* 166: 74–82. https://doi.org/10.1016/j.ymeth.2019.03.003

54. Yazdani, Mehrdad, Bryn C. Taylor, Justine W. Debelius, Weizhong Li, Rob Knight, and Larry Smarr. 2016. "Using Machine Learning to Identify Major Shifts in Human Gut Microbiome Protein Family Abundance in Disease." In *IEEE International Conference on BigData* 1272–80. https://doi.org/10.1109/BigData.2016.7840731

55. Espinoza, J. Luis. 2018. "Machine Learning for Tackling Microbiota Data and Infection Complications in Immunocompromised Patients with Cancer." *Journal of Internal Medicine* 284: 189–92. https://doi.org/10.1111/joim.12746

56. Larsen, Peter E., Dawn Field, and Jack A. Gilbert. 2012. "Predicting Bacterial Community Assemblages Using an Artificial Neural Network Approach." *Nature Methods* 9: 621. https://doi.org/10.1038/nmeth.1975

57. Zhou, Guangyu, Jyun-Yu Jiang, Chelsea J.-T. Ju, and Wei Wang. 2019. "Prediction of Microbial Communities for Urban Metagenomics Using Neural Network Approach." *Human Genomics* 13: 47. https://doi.org/10.1186/s40246-019-0224-4

58. Zhang, Han, Xi Gao, Jacob Unterman, and Tom Arodz. 2019. "Approximation Capabilities of Neural Ordinary Differential Equations." arXiv preprint arXiv:1907.12998. https://arxiv.org/abs/1907.12998

59. Dupont, Emilien, Arnaud Doucet, and Yee Whye Teh. 2019. "Augmented Neural ODEs." arXiv preprint arXiv:1904.01681. https://arxiv.org/abs/1904.01681

60. Evans, Steven N., and Frederick A. Matsen. 2012. "The Phylogenetic Kantorovich-Rubinstein Metric for Environmental Sequence Samples." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 74: 569–92. https://doi.org/10.1111/j.1467-9868.2011.01018.x

61. Peyré, Gabriel, and Marco Cuturi. 2019. "Computational Optimal Transport." *Foundations and Trendső in Machine Learning* 11: 355–607. https://doi.org/10.1561/2200000073

62. Martins, Andre, and Ramon Astudillo. 2016. "From Softmax to Sparsemax: A Sparse Model of Attention and Multi-Label Classification." In *International Conference on Machine Learning*, 1614–23. New York, New York, USA: PMLR. https://dl.acm.org/doi/10.5555/3045390.3045561

63. Hofbauer, Josef, and Karl Sigmund. 1988. *The Theory of Evolution and Dynamical Systems: Mathematical Aspects of Selection* (7, London Mathematical Society Students Text). https://searchworks.stanford.edu/view/1317720

64. Legendre, Pierre and Louis Legendre. eds. 2012. "Numerical Ecology." In *Developments in Environmental Modelling* (3rd English edition, 24). Amsterdam: Elsevier. https://www.elsevier.com/books/numerical-ecology/legendre/978-0-444-53868-0

65. Nichol, Alex, Joshua Achiam, and John Schulman. 2018. "On First-Order Meta-Learning Algorithms." arXiv preprint arXiv:1803.02999. https://arxiv.org/abs/1803.02999

66. Case, T. J. 2000. *An Illustrated Guide to Theoretical Ecology.* Oxford: Oxford University Press. https://doi.org/10.1644/1545-1542(2001)082<0247:>2.0.CO;2

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.