Energy landscape analysis of ecological communities elucidates the phase space of community assembly dynamics

Kenta Suzuki[1*], Shinji Nakaoka[2,3], Shinji Fukuda[3-6] and Hiroshi Masuya[1]

* Corresponding author: kenta.suzuki.zk@riken.jp

1. Integrated Bioresource Information Division, Bioresource research center, RIKEN, 3-1-1 Koyadai, Tsukuba, Ibaraki, Japan.

2. Laboratory of Mathematical Biology, Faculty of Advanced Life Science, Hokkaido University, Kita-10 Nishi-8, Kita-ku, Sapporo, Hokkaido, Japan.

3. PRESTO, Japan Science and Technology Agency, Japan.

4. Institute for Advanced Biosciences, Keio University, 246-2 Mizukami, Kakuganji, Tsuruoka, Yamagata 997-0052, Japan.

5. Intestinal Microbiota Project, Kanagawa Institute of Industrial Science and Technology, 3-25-13 Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa, Japan.

6. Transborder Medical Research Center, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan.

Abstract

Studies on historical contingency in community assembly frequently report the presence of alternative stable states as the result of different assembly sequences. If we can observe multiple assembly sequences and resulting community structure, these observations collectively inform the constraints on community assembly dynamics that emerge from various ecological processes. These observations would be a basis for predicting the outcome of assembly processes and understanding the mechanisms that shape species assemblage. However, empirical approaches such as invasion/removal experiments require enormous time and effort and are impossible in some cases. Here, we show that data on multispecies occurrences analyzed using a pairwise maximum entropy model and energy landscape analysis are capable of providing insights into the constraints on community assembly dynamics. This approach is a minimal theoretical framework to systematically and mechanistically study community assembly dynamics. Community assembly has a prominent role in shaping real world ecosystem organization. Our approach provides a new systemic paradigm for developing a predictive theory of community ecology and can have broad impact on ecological studies and its applications including regime shifts and ecosystem management.

Introduction

Community assembly can be defined as the construction and maintenance of local communities through sequential arrival of potential colonists from an external species pool (Drake 1991, Fukami 2010). During assembly processes, the extinction of species from a local community or the migration of species from an external species pool drive transitions from one community state to another. The order and timing of species migration and extinction during community assembly influence structure (Drake 1991, Fukami and Morin 2003) and function (Jiang et al. 2011) of communities, resulting in historical contingency, such as alternative stable states (Fukami 2010, Fukami 2015). Alternative stable states, i.e., stable coexistence of multiple community states under the same environmental conditions, have many empirical examples (see, e.g., review by Schloder et al. 2005). For example, Drake (1991) experimentally showed the effect of various sequences of species invasions on final community composition in aquatic microbial communities. More recently, also in a laboratory experiment with aquatic microbial communities, Pu & Jiang (2015) found that alternative community states were maintained for many generations despite frequent dispersal of individuals among local communities.

Transition among different community states is the outcome of the ecological processes associated with the members of the species pool, and it ultimately constrains community assembly dynamics (Vellend 2010, Weiher et al. 2011, Götzenberger et al. 2012, HilleRisLambers et al. 2012, Wisz et al. 2013). Understanding the global phase space that includes all possible community states of community assembly dynamics (Law and Morton 1993, Capitan et al. 2011) is important for predicting the outcome of assembly processes and understanding the mechanisms that shape local communities. Law and Morton (1993) introduced a graph representation of community assembly dynamics (termed, 'assembly graph') in which nodes represent different community states (species compositions) and directed edges represent transitions from one community state to another. The assembly graph maps out all assembly pathways generated by invasions and extinctions. The same graph also appeared in the toy model of food web assembly introduced by Capitan et al. (2011). While these studies intended to analyze assembly dynamics in mathematical models, Warren et al. (2003) constructed the assembly graph for a pool of six protozoan species by integrating the results of two experimental studies (Weatherby et al. 1998, Law et al. 2000). By analyzing the assembly graph, Warren et al. (2003) showed that the system had a compositional cycle (Morton and Law 1997) with multiple alternative transient trajectories. An assembly graph acknowledges the relationship among community states in the global phase space of a system. These graphs represent the constraints on community assembly and can provide predictions on the outcome of assembly processes (Law and Morton 1993, Capitan et al. 2011). However, repeating experiments as in Weatherby et al. (1998) and Law et al. (2000) requires enormous time and effort, and are unrealistic if the number of species is large or the target community is difficult to manipulate experimentally. Although a theoretical approach would be the alternative, applying the approaches proposed by Law and Morton (1993) or Capitan et al. (2011) to handle observational data is not straightforward. Since these approaches were proposed to analyze the behavior of differential equations, we need to include a statistical framework to fit the parameters of differential equation models. However, it is generally difficult to fit differential equations for multi-

species communities from observational data because of the computational difficulty in fitting model parameters as well as difficulty in obtaining time-series data sufficient to fit these parameters. Therefore, there is a need for developing a novel theoretical framework to systematically and mechanistically study the community assembly dynamics.

In this paper, we propose an approach incorporating a pairwise maximum entropy model (also known as Markov network; Azaele et al. 2010, Araujo et al. 2011, Harris 2016) and the energy landscape analysis (Becker and Karplus 1997, Wales et al. 1998, Watanabe et al. 2014a) to model and analyze community assembly dynamics from observational data (Fig. 1). Our approach starts with community data including observations on the occurrence of species in a set of temporal and/or spatial replicates ('samples') with any accompanying values representing local abiotic environment (explicit abiotic factors). This dataset is then converted to the matrices of presence/absence status and explicit abiotic factors (if available) (Fig. 1B). Second, these matrices are used to fit parameters in a pairwise maximum entropy model (Fig. 1C). A pairwise maximum entropy model has previously been applied to infer species interaction from presence/absence data (Harris 2016). Because it relies on physical association between species it may not represent true ecological interactions (Barner et al. 2018). Previous studies showed that the model could accurately predict the occurrence of species due to its ability to incorporate both biotic and abiotic factors (Azaele et al. 2010, Araujo et al. 2011). Here, we do not focus on the interpretation of model parameters but its 'energy landscape' specified by the fitted pairwise maximum entropy model (Fig. 1D). The energy landscape is a network with nodes representing community states and links representing transitions between community states. The analysis of topological and connection attributes of this weighted network is termed 'energy landscape analysis' (Watanabe et al. 2014a). This analysis was first developed in the studies of molecular dynamics (Wales et al. 1998, Becker and Karplus 1997) and recently applied to the analysis of brain activity (Watanabe et al. 2014a,b, Ezaki et al. 2017, Watanabe & Rees 2017, Ezaki et al. 2018).

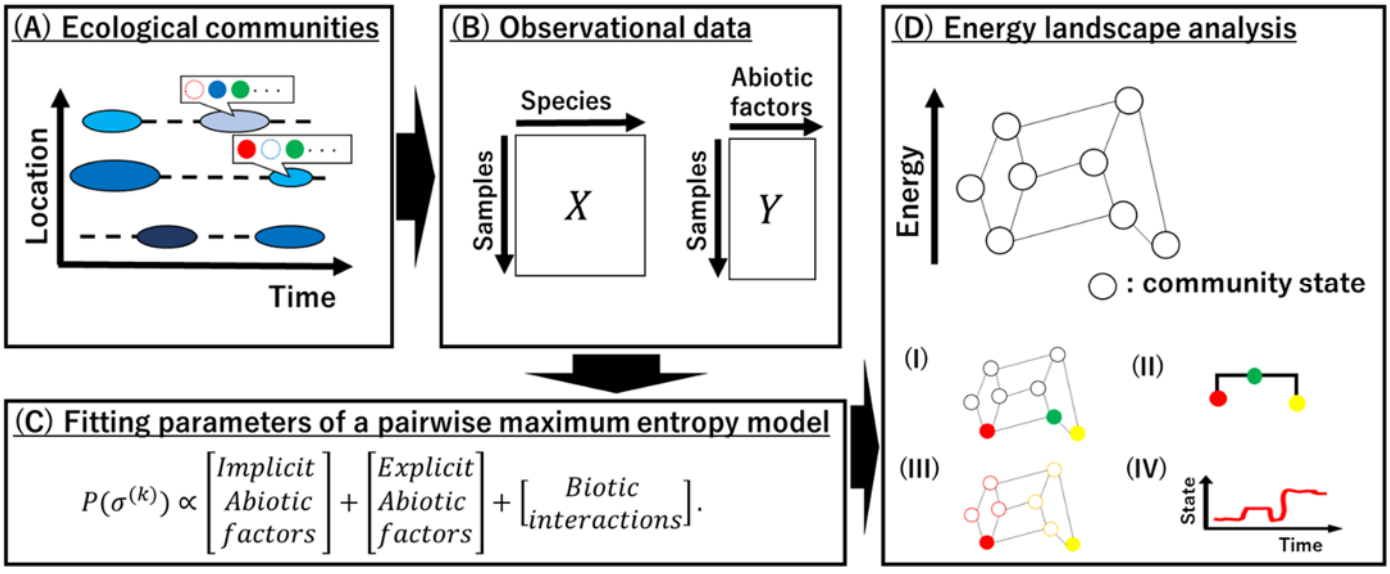

Figure 1. Illustrative explanation of our approach. (A) we assume that the dataset includes occurrence of species in local communities sampled from multiple sites and/or timepoints, with possibly accompanying values representing local abiotic environment (explicit abiotic factors) (in the illustration, circles and filled circles show species that is absent from or present in a local community, and colors and size of ellipses represent differences in local abiotic environment). (B) dataset is converted to matrices of presence/absence status, and explicit abiotic factors (if they are available). (C) these matrices are used to fit parameters in a pairwise maximum entropy model. Here, $p(\vec{\sigma}^{(k)})$ is the probability of a community state $\vec{\sigma}^{(k)}$ (see Materials and Methods for the detail). (D) the fitted pairwise maximum entropy model specifies an energy landscape network with nodes representing community states and links representing transitions between community states. Analysis of the topology and connectivity of this weighted network, i.e., energy landscape analysis, acknowledges (I) the stable states (red and yellow filled-circles) and tipping points (green filled-circle), (II) disconnectivity graph as the summary of hierarchical relationships between the stable states and tipping points, (III) attractive basin of stable states (red and yellow circles indicate attractive basins of the two stable states). (IV) We also able to emulate assembly dynamics constrained on the energy landscape.

The purpose of this paper is to show how the energy landscape analysis allows us to systematically and mechanistically study assembly dynamics. Understanding community assembly in terms of the management of ecological systems has direct relevance to conservation biology, agriculture, and medicine (Fukami et al. 2015). In conservation biology, knowledge of the paths by which communities are assembled helps ecologists to understand the role of history in shaping current communities, and is important for effective community restoration (Drake 1990, Pimm 1991, Lockwood and Pimm 1999, Weiher and Keddy 1999, Lockwood & Samuels 2004, Suding et al. 2004, Wilsey et al. 2015, Young et al. 2005, 2015). In other words, when historical contingency occurs, restoring and maintaining native biodiversity may require specific sequences of exotic species removal and/or native species introduction. This is also relevant to agriculture, e.g., the successful inoculation of agricultural soils with beneficial fungi or other microbes may depend on the timing of inoculation relative to plant growth, as well as the profile of other soil microbes (Verbruggen et al. 2013, Toju et al. 2018). In medicine, the relevance of historical contingency in community assemblies to curing some human diseases is being recognized (Costello et al. 2012, Fierer et al. 2012, Lam & Monack 2014, Devevey et al. 2015). Clinically meaningful evidence for the potential application of modulating the intestinal microbiota for therapeutic gain has created considerable interest and enthusiasm (Smits et al. 2013, Li et al. 2016, Shetty et al. 2017). For example, a disruption to the gut microbiota is associated with several disorders including irritable bowel syndrome, *Clostridium difficile* Infection (CDI), autism, obesity, and cavernous cerebral malformations (Karczewski et al. 2014, Cox et al. 2015, Tang et al. 2017). Driving disrupted microbial communities back to their healthy states could offer novel solutions to prevent and treat complex human diseases (Van Nood et al. 2013).

By making the most of observational data, our approach not only provides information on the constraints on community assembly dynamics (Law and Morton 1993, Warren et al. 2003, Capitan et al. 2011) but describes change across environmental gradients, allows for simulation-based analysis, and provides mechanistic understandings on background ecological processes at work. First, we present a demonstration using a model with predefined (virtual) parameters that describe community assembly dynamics of an eight species metacommunity. Using this model, we explain the basic concepts of our analysis, and then we show how the community assembly dynamics can be emulated as the dynamics constrained on the energy landscape. Furthermore, we introduce an abiotic factor to the model and explain how response of the energy landscape to abiotic factors is captured in our model. Finally, to see how maximum likelihood methods works to infer the energy landscape from observational data, we generated different sized datasets from the models (either with or without abiotic factors) and examined the relationship between the correctness of inferred energy landscape and the dataset size. Subsequently, to demonstrate the application of our approach to a real community, we applied the same analysis to the mouse gut microbiota. We found a major shift in the energy landscape during the transition from young to middle age and discuss the implications on community assembly dynamics across life-stages. We discuss the significance of our approach to community assembly studies, as well as studies on regime shifts and applications to ecosystem management.

Materials and Method

Data structure

We assume that we have a dataset containing the composition of local communities in N samples. The local community state is represented by a string of 0's (absence) and 1's (presence) whose length S is the total number of species included in the whole dataset. Thus, there is a total of 2^S community states. We denote a community state of kth sample as a random variable $\vec{\sigma}^{(k)} = (\sigma_1^{(k)}, \sigma_2^{(k)}, \dots, \sigma_S^{(k)})$, where $\sigma_i^{(k)} \in \{0,1\}$ for $i = 1,2,\dots,S$ and $k = 1,2,\dots,N$. We denote the community states of N samples as an $S \times N$ matrix $X = \{\vec{\sigma}^{(1)}, \vec{\sigma}^{(2)}, \dots, \vec{\sigma}^{(N)}\}$. If available, the abiotic environment of the samples is represented by a vector of length M that may include, for example, resource availability, pH, altitude, or age of host organism, which is referred to as explicit abiotic factors. We denote the explicit abiotic factors of N samples as $M \times N$ matrix $Y = \{\vec{\epsilon}^{(1)}, \vec{\epsilon}^{(2)}, \dots, \vec{\epsilon}^{(N)}\}$, where $\vec{\epsilon}^{(k)} = \{\epsilon_1^{(k)}, \epsilon_2^{(k)}, \dots, \epsilon_M^{(k)}\}$ represents the abiotic factors of the kth sample.

Pairwise maximum entropy model

To calculate the probability of community states $\vec{\sigma}^{(k)}$, i.e., $p(\vec{\sigma}^{(k)})$, we use a pairwise maximum entropy model derived from the principle of maximum entropy (Jaynes 1982, Mead and Papanicolau 1984, Jaynes 2003, Azaele et al. 2011). The principle of maximum entropy is a powerful tool to explain statistical patterns in natural world (Azaele et al. 2010). In its simplest form, i.e., without including explicit abiotic factors, the principle requires the realized distribution of community states to maximize the following pairwise maximum entropy model (see Azaele et al. 2010 for the detail):

$$p(\vec{\sigma}^{(k)}) = e^{-E(\vec{\sigma}^{(k)})}/Z, \quad (1)$$

$$E(\vec{\sigma}^{(k)}) = -\sum_{i=1}^{S} h_i \sigma_i^{(k)} - \sum_{i=1}^{S} \sum_{j=1, j\neq i}^{S} J_{ij} \sigma_i^{(k)} \sigma_j^{(k)} /2. \quad (2)$$

Here, $E(\vec{\sigma}^{(k)})$ is the energy of community state $\vec{\sigma}^{(k)}$, and

$$Z = \sum_{i=1}^{2^S} e^{-E(\vec{\sigma}^{(i)})}. \quad (3)$$

Parameters in this model are h_is and J_{ij}s which are elements of a vector $h = \{h_1, h_2, .., h_S\}$ and a matrix $J = \{J_{ij}\}_{i=1,2,...,S; j=1,2,...,S}$, respectively. As a model for ecological community, h_i may be interpreted as a parameter representing the net effect of the implicit abiotic factors (gloss effect from the abiotic environment), which may favor the presence ($h_i > 0$) or absence ($h_i < 0$) of species i. Moreover, each species is coupled to all others in a pairwise manner through J_{ij}s, and correspondingly, $J_{ij} > 0$ favors the co-occurrence of species i and j and $J_{ij} < 0$ disfavors the co-occurrence of species i and j. Here, the term energy is used as the diversion from statistical physics since this model was first proposed in this field (Brush 1967). In ecology, it is nothing but a multiplier of the Napier's constant as in eq.(1), i.e., the logarithm of the probability of a community state is inversely proportional to $E(\vec{\sigma}^{(k)})$ where $\log p(\vec{\sigma}^{(k)}) = -E(\vec{\sigma}^{(k)}) - \log Z = -E(\vec{\sigma}^{(k)}) + \sum_{i=1}^{2^S} E(\vec{\sigma}^{(i)})$. Hence, this equation implies that a community state having lower energy than others is more frequently observed.

If observations on some abiotic factors are available, equation (2) can be extended to:

$$E(\vec{\sigma}^{(k)}) = -\sum_{i=1}^{S} h_i \sigma_i^{(k)} - \sum_{i=1}^{S} \sum_{j=1}^{M} g_{ij} \epsilon_j^{(k)} \sigma_i^{(k)} - \sum_{i=1}^{S} \sum_{j=1, j\neq i}^{S} J_{ij} \sigma_i^{(k)} \sigma_j^{(k)} /2. \quad (4)$$

Here, g_{ij} represents the effect of ith abiotic factor on the occurrence of jth species, and it is an element of a $M \times S$ matrix $g = \{g_{ij}\}_{i=1,2,...,M; j=1,2,...,S}$. Now, the first term of eq. (4) should be interpreted as the net effect of implicit (unobserved) abiotic factors, while the second term represents the effect of explicit (observed) abiotic factors.

Both eq. (2) and (4) are based on the pairwise interaction. Thus, it implicitly assumes that all higher-order occurrence patterns are well-captured by the information encapsulated in the first two moments. This is relaxed by including the high order terms in these equations. However, such extension would increase the dataset size required to obtain enough predictive performance (Nguyen et al. 2017).

As we explained in Appendix A, the maximum likelihood estimates of h, J and g for observational data can be obtained by the mean occurrences (number of occurrences) for each species, the mean co-occurrences (number of co-occurrences) between each species and the cross-product between the species and environment matrices. The relevance of the pairwise maximum entropy model is grounded on the assumption that a statistical population of local community states satisfies the principle of maximum entropy and the dataset is an unbiased (random) sample from such a statistical population. We adopt this assumption as a working hypothesis for our analysis, although relevance of this assumption requires further investigations. However, it is worth noting that the principle of maximum

entropy has a long history of theoretical development (Jaynes, 1957a, b), and played an important role for developing predictive models in a vast array of scientific fields (Schneidman et al., 2006; Lezon and et al., 2006; reviewed by Nguyen et al. 2017), including ecology (Shipley et al. 2006, Dewar and Porte´ 2008, Harte et al. 2008).

Ecological processes in pairwise maximum entropy model

We describe how the essential ecological processes that can shape local assemblages are involved in our model. We present the following argument for conceptual clarity; the main purpose of this paper is not to provide detailed analysis for discussing the relative strength of different processes.

In terms of temporal scale, our method infers the constraints of community assembly dynamics attained by the set of species and environmental factors at an interval in which observations took place. In other words, we do not account for the processes that can change the composition of species pool, e.g., migration of species from the outside of the set of communities, and species extinction from or evolution within the set of communities. These processes might be handled by extending our model, though we do not consider this to avoid complications.

Even without considering these processes, local species assemblages are affected by numerous deterministic and stochastic processes, typically including the response of species to abiotic environment (environmental filtering), biotic interactions (biotic filtering), dispersal limitation, and stochastic processes (e.g., disturbance, demographic stochasticity, and ecological drift) (Vellend 2010, Weiher et al. 2011, Götzenberger et al. 2012, HilleRisLambers et al. 2012, Gravel 2013, Wisz et al. 2013). As in eq. (2) and (4), our model explicitly accounts for the effect of abiotic factors and biotic interactions. On the other hand, stochasticity is implemented by the absolute value of parameters. To explain this, let us introduce a control parameter $\tau \in (0, \infty)$ and denote h/τ, J/τ and g/τ. We can characterize the relative strength of stochasticity in community assembly processes in terms of τ. When $\tau \to \infty$, the differences among species disappear. This means that all community states will be equally observed, i.e., $\mathrm{p}(\vec{\sigma}^{(k)}) = 1/2^{\mathrm{S}}$ for $\mathrm{k} = 1, \ldots, \mathrm{N}$. On the other hand, τ augments the difference in deterministic effects on species' occurrence when $\tau \to 0$. In this limit, $\mathrm{p}(\vec{\sigma}^{(k)})$ of a community state having the smallest energy approaches 1 while that of the others approach 0. Community state is fixed to a composition since any stochastic effects have disappeared. This consideration highlights that the absolute value of the parameters determines the balance between deterministic and stochastic community assembly processes. Dispersal (and thus dispersal limitation) is not explicitly implemented in our model. A species included in the model occurs with the same probability throughout different samples. However, effect of dispersal limitation would be considered when we include spatial attributes as the element of abiotic factors. In such a model, occurrence of a species will vary as a function of the spatial attributes.

Our approach differs from species distribution models (Elith & Leathwick 2009) and joint species distribution models (Warton et al. 2015) because it includes biotic interactions. However, this does not mean that the same ecological interaction would always result in the same consequence in species' presence/absence status. For example, let us consider a simple example of a predator-prey metapopulation (Hanski and Gilpin1991); even when species A consumes B, and A exhausts B in a site, B may coexist with A in the metacommunity level if B is better at dispersing relative to A. In this case, we may see a checkerboard distribution where A and B typically occupy different sites. However, if A does not exhaust B and coexist within a site, then we see co-occurrence of A and B because presence of B facilitates presence of A. There may be an opposite consequence for the predator-prey relationship. Therefore, signal of the biotic interaction (J_{ij}) can be identified by its effect on the presence/absence status. Any biotic interactions (and other deterministic ecological processes) are ultimately identified by their effect on presence/absence status rather than the type of distinct ecological processes.

Energy landscape analysis

The energy landscape of the ecological dataset is defined on a network with nodes representing community states and links representing their transition. Two nodes are defined to be adjacent by a link only if they take the opposite status (i.e., 0/1) for just one species. In other words, the hamming distance between two adjacent nodes is always 1. To each node, $E(\vec{\sigma}^{(k)})$, i.e., the energy of the corresponding community states $\vec{\sigma}^{(k)}$, is assigned. The distribution of energy over the network characterizes the energy landscape. Energy landscape analysis is a methodology to analyze the topology and connectivity of this high dimensional phase space represented as a weighted network (Becker & Karplus 1997, Wales et al. 1998, Wales, 2010, Watanabe et al. 2014a,b, Ezaki et al. 2017, Watanabe & Rees 2017, Ezaki et al. 2018).

Energy minima - A local minimum is a node with energy less than all S neighboring nodes. We exhaustively examined whether each of the 2^S nodes were local minima. Since the local minima have the lowest energy compared to all neighboring nodes, these corresponding community states constitute end-points when assembly processes are completely deterministic (i.e., when community states must always go down the energy landscape). We identified energy minima as stable states of the community assembly dynamics. Presence of alternative stable states can be identified as multiple energy minima within an energy landscape.

Basin of attraction - The attractive basin of an energy minimum was computed as follows. First, we selected a node i in the energy landscape. If the selected node was not a local minimum, we moved to the node with the lowest energy value among the nodes adjacent to the current node. We repeated moving downhill in this manner until a local minimum was reached. The initial node i belongs to the basin of the finally reached local minimum. We ran this procedure for each of 2^S nodes. The basin size of a local minimum is the fraction of nodes that belong to the basin of the local minimum. When assembly processes are completely deterministic, all the community states belonging to an attractive

basin eventually reach one distinct stable state. We denote the basin of an energy minima A as B_A and union of the basin of A and B as B_{AB}.

Disconnectivity graph, energy barriers and tipping points - A disconnectivity graph summarizes the structure of the energy landscape. It can be obtained as follows, after obtaining the local minima. First, we set an energy threshold value, denoted by E_{th}, to the energy value of the community state that attained the second highest energy value among the 2^S community states. Second, we removed the nodes corresponding to the community states whose energy exceeded E_{th}. When E_{th}, is the second highest energy, the node with the highest energy was removed. We also removed the links incident to the removed nodes. Third, we checked whether each pair of local minima was connected in the reduced network. Forth, we lowered E_{th} to the next highest energy value realized by a community state. Then, we repeated the third to fifth steps, that is, removal of the nodes and links, checking for connectivity between local minima, and lowering of E_{th}, until all the local minima were isolated. During this process, for each pair of local minima, we recorded the lowest E_{th} value below which the two local minima were disconnected. This value is equal to the energy barrier that the assembly dynamics must overcome to reach from one local minimum to another. We referred the community state having the threshold energy value as a tipping point. Finally, we constructed a hierarchical tree whose terminal leaves represented the local minima. The vertical positions of these leaves represent their energy values. Those of the branches represent the height of the energy barrier that separates the local minima belonging to the two branches. Thus, a disconnectivity graph represents the hierarchical relationship among alternative stable states.

Basin boundary - We regard a pair of adjacent community states belonging to different attractive basins as a boundary pair. Boundary pairs differ from the tipping point because they acknowledge the entire structure of the boundary between two basins. To summarize the energy of basin boundary, for each boundary pair, we assigned $\mathrm{Max}(E(\vec{\sigma}^{(A)}), E(\vec{\sigma}^{(B)}))$ as the energy of the pair, where $\vec{\sigma}^{(A)}$ and $\vec{\sigma}^{(B)}$ are community sates included in the boundary pair.

Path between two energy minima - If the hamming distance between two energy minima A and B is L, then there are $L!$ paths connecting the two energy minima with L steps. Following the previous study (Ezaki et al. 2018), we defined the path energy (PE) as the sum of energy required for a path connecting two energy minima:

$$\mathrm{PE}(P_i) = \sum_{t=2}^{X} \theta\big(E(\sigma_{it}) - E(\sigma_{it-1})\big).$$

Here, $P_i = \{\sigma_{i1}, \sigma_{it}, \ldots, \sigma_{iL}\}$ represents the ith path connecting A and B, and σ_{ij} is the jth community state on the path. $\theta(x)$ is a function that returns zero when $x \leq 0$ and x when $x > 0$. We refer to this value as an index of the ease of transition between community states.

Numerical simulations - We carried out numerical simulations to emulate community assembly dynamics constrained on an energy landscape. We employed the heat-bath (also known as Gibbs sampling) method (Gilks et al. 1996) as follows. First, we selected an initial community state. Then, in each time step, a transition from the current community state $\vec{\sigma}^{(k)}$ to one of its S adjacent community state $\vec{\sigma}^{(k')}$, selected with probability 1/S, was attempted ($\vec{\sigma}^{(k)}$ and $\vec{\sigma}^{(k')}$ differs only with respect to the presence/absence status of one of S species). The transition to the selected community state took place with probability $e^{-E\left(\vec{\sigma}^{(k')}\right)}/(e^{-E(\vec{\sigma}^{(k)})} + e^{-E(\vec{\sigma}^{(k')})})$. This procedure provides a sequence of transition of community states constrained on an energy landscape, and we refer to it as the emulated community assembly dynamics.

Sample Data for demonstration

Eight-species metacommunity model – To explain the basic concepts and work flow of the analysis, we use an eight species metacommunity model with predefined (virtual) parameters. Therefore, parameters for the model were not inferred from real data but given *a priori*. The parameter values were selected so that the system had three alternative stable states. We first evaluated the model without no explicit environmental factors, i.e., we set parameters for the biotic interactions (J) and the responses to implicit abiotic factors (h). We then extended the same model by including an explicit abiotic factor (ϵ), and set the additional parameters (g) for the responses to the explicit abiotic factor. Since we assumed only one abiotic factor, ϵ and g were defined as vectors with N elements instead of matrices. Values of these parameters are presented as supplementary information.

Mouse gut microbiota - We applied our approach to the data of gut-microbiota taken from the feces of six male C57BL/6J mice, which is in the DDBJ database (http://trace.ddbj.nig.ac.jp/DRASearch/) under accession number DRA004786 (Nakanishi et al. submitted). Feces were sampled once every 4 weeks between 4 to 72 weeks of age, thus 18 data points were obtained per mouse. Hence, 96 data points are available. We transformed the relative abundance data into presence/absence data by setting a cutoff level as 1%, and we picked up OTUs that found between 20% to 80% samples. As the result we obtained the presence/absence status of 8 OTUs specified at the genus level. We also used age of mouse (4-72 weeks) as an explicit environmental parameter. In the analysis, we scaled 4-72 weeks to a value within a 0-1 range. We assumed that 4 weeks interval was sufficiently longer than the transient dynamics of the gut microbiota (Gerber 2014), and treated microbiota composition of the same mouse at different ages as independent data. Since we included only age as the abiotic factor in the analysis, ϵ and g were defined as vectors with N elements instead of matrices.

Results

Analysis of metacommunity models

We first explain the basic concepts and work flow of our approach, using an eight species metacommunity model with predefined (virtual) parameters. As explained in Materials and Methods, we regard the energy minima as stable states henceforth.

Basic Concepts - The eight-species metacommunity model is defined so that it has three alternative stable states (Fig. 2a). We regard them as stable states A, B and C. The energy of A, B and C were -10.56, -10.49 and -9.70, respectively, and these values corresponded to their probability (Fig. 2b). The relative basin size of A, B, and C is shown in Fig. 2c. The height of the energy barrier for the transition between stable states is shown in Fig. 2d and was estimated as the difference of energy between the departed stable state to the tipping point (indicated as b1 and b2 in Fig. 2a). The energy landscape acknowledges the hierarchical relationship among the three alternative stable states in this model. There was a tipping point b1 at -9.97 that connects stable state A and B (Fig. 2a). This means that b1 connects the attractive basin of stable state A (B_A) and that of stable state B (B_B) at this energy level. A second tipping point b2 was found at -8.49 (Fig. 2a) that connects b1 and C. This means that the union of the two basins B_A and B_B (B_{AB}) was connected to the attractive basin of stable state C (B_C) via b2 at this energy level.

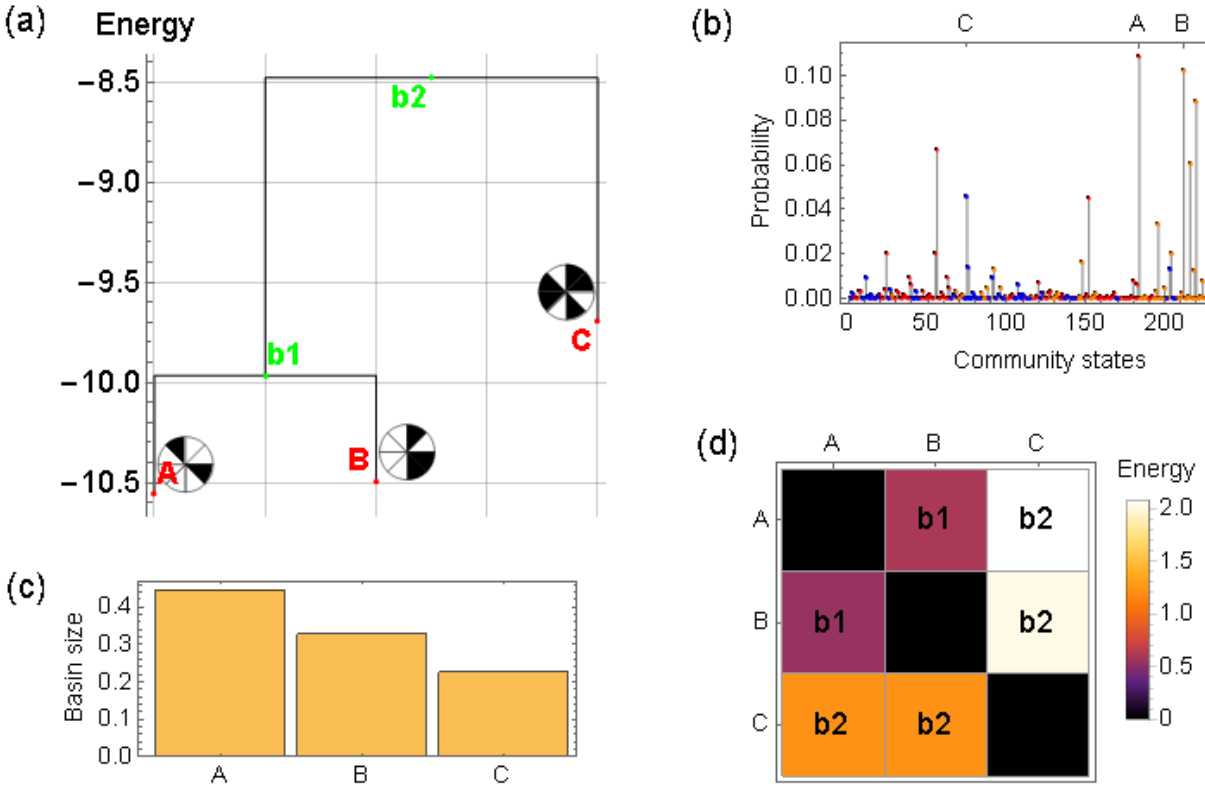

Figure 2. Summary of the energy landscape of the eight species metacommunity model. (a) disconnectivity graph showing the energy minima; stable states (red points), tipping points (green points), and community composition of stable states. Pie chart indicates presence (white) or absence (black) of species $i = 1, \ldots, 8$ in clockwise order. (b) probability of community states calculated from eq.

(2). Stable state A, B and C and their basins are indicated by Red, Orange and Blue points, respectively. (c) relative basin size of energy minima. (d) height of energy barrier (difference in energy between the departed stable state to the tipping point) for the transition from stable states shown in columns to rows.

Emulating community assembly dynamics - To consider the effect of the energy landscape on the community assembly dynamics, we considered the transition dynamics between stable states (steady state dynamics) rather than the dynamics initiated by absence of species (transient dynamics) because these provide general information on energy landscape dynamics, and have implications on the ecosystem management issues that deals with transition between established communities.

Transitions between A and B occurred more frequently than between A or B to C (Fig. 3a) because of the lower energy barrier between them (Fig. 2d). We considered two strategies to investigate the transition dynamics between the stable states: the first one focused on tipping points and basin boundary and the second one focused on the paths connecting two stable states. Here, we considered transition between the stable states A and C.

Since energy of the tipping point between A and B (b1) was lower than that of A and C (b2), we needed to consider the boundary between B_{AB} and B_C rather than B_A and B_C for transition between A and C. There were 186 boundary pairs that constituted the basin boundary between B_{AB} and B_C. We performed a numerical simulation (10^5 steps heat-bath algorithm) and calculated the frequency that each boundary pair was visited (Fig. 3b). There was a negative correlation between the frequency and the energy of boundary pairs (Pearson's correlation coefficient, -0.53, $P<0.001$). The top 10 low energy boundary pairs (5.3% of all boundary pairs) covered 56.7% of visited boundary pairs during transition (Fig. 3c), and these pairs included tipping point b2 (Table 1). This result shows that we can characterize the transition between stable states using a small number of critical community states. The set of these community states can be understood as the channel that mediates transition between alternative stable states.

Since hamming distance between A and C was 7, there were 5040 (= 7!) possible paths connecting A and C with the shortest path length. However, there may be a small number of effective paths since the paths that eventually connect two states will have lower $PE(P_i)$ than the others. Figure 3d shows the negative correlation between PE and the frequency of paths (Pearson's correlation coefficient was -0.43, $P<0.001$). Here, the top 20 low PE paths (0.3% of all paths) covered 19.2% of transitions between A and C (Fig. 3e), again indicating that there was a relatively small fraction of paths that eventually connected stable states A and C. Since coverage of these paths will decrease when considering paths with different length (i.e., $L = 8,9,\ldots$), we need to consider many more paths as eventual transition paths. However, the presented result has implications for effective sequencing of species introduction

or removal to control community states. As is shown by the emulated community assembly dynamics here, some low PE paths will be more efficient than others that would be selected by random trials.

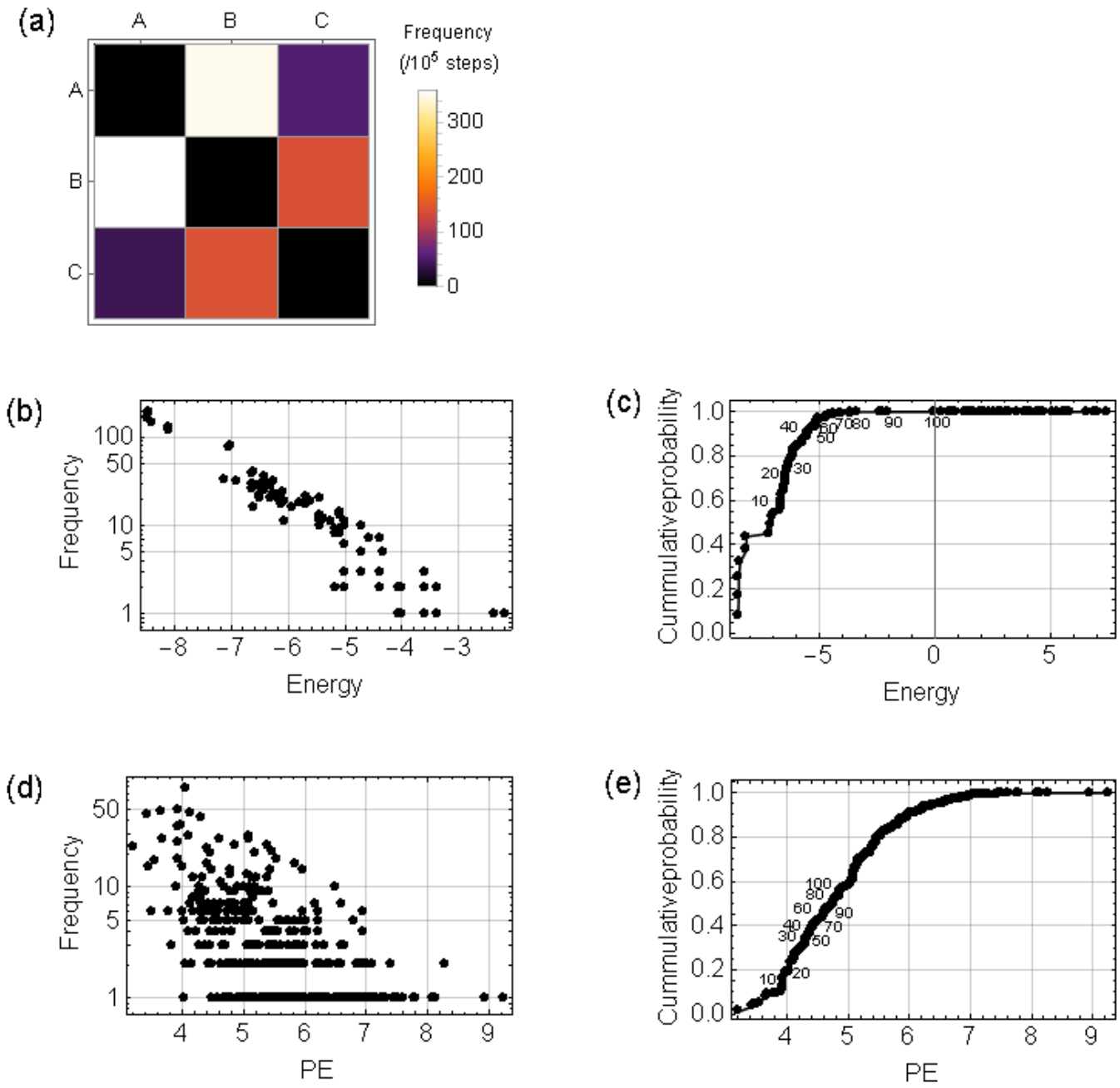

Figure 3. Transition dynamics between community state A and C. (a) frequency of transition between stable states that has a negative correlation to the energy barrier represented in Fig. 2d (Pearson's correlation coefficient between the transition frequency and energy barrier was -0.79, P=0.06). (b) relationship between the energy of boundary pairs and the frequency that they are visited. (c) cumulative probability distribution of (b). Here, points were indicated by the rank of boundary pairs with ascending order of their energy. (d) relationship between PE and the frequency of paths. (e) cumulative probability distribution of (d). Points were indicated by the rank of paths with ascending order of their PE. (a-c) were calculated from a numerical simulation with 10^5 steps heat-bath

algorithm. For (d-e), we repeated 10^5 steps heat-bath simulations until we obtain 2,000 paths whose length was 7. To obtain (a), we picked up the realized paths that start from one of the stable states and arrive at one of the others. To obtain (d-e), we first picked up the realized paths that start from A (C) and arrive at C (A) without visiting B, and pruned loops (intervals that come back to the same community state) and removed redundancy (states that was not moved to other state), and then picked up the path whose length was 7.

Rank	Energy	Community state in BAB	Community state in BC	Cumulative probability
1	-8.48	{1, 1, 0, 0, 1, 0, 1, 1}	{0, 1, 0, 0, 1, 0, 1, 1}*	0.08
2	-8.45	{1, 1, 0, 0, 1, 0, 1, 1}	{1, 1, 0, 0, 1, 0, 1, 0}	0.163
3	-8.45	{0, 1, 0, 1, 1, 0, 1, 1}	{0, 1, 0, 0, 1, 0, 1, 1}*	0.233
4	-8.39	{1, 1, 0, 1, 1, 0, 1, 0}	{1, 1, 0, 0, 1, 0, 1, 0}	0.309
5	-8.12	{0, 1, 0, 1, 1, 0, 1, 1}	{0, 1, 0, 1, 1, 0, 1, 0}	0.37
6	-8.12	{1, 1, 0, 1, 1, 0, 1, 0}	{0, 1, 0, 1, 1, 0, 1, 0}	0.425
7	-7.12	{0, 0, 0, 1, 1, 1, 1, 0}	{0, 0, 0, 1, 1, 0, 1, 0}	0.45
8	-7.05	{1, 1, 0, 0, 0, 0, 1, 0}	{1, 1, 0, 0, 1, 0, 1, 0}	0.479
9	-7.01	{0, 1, 0, 0, 0, 0, 1, 1}	{0, 1, 0, 0, 1, 0, 1, 1}*	0.507
10	-6.9	{0, 0, 1, 0, 1, 1, 1, 0}	{0, 0, 1, 0, 1, 0, 1, 0}	0.522

Table 1. Profile of top 10 low energy boundary pairs. * indicates tipping point between B_{AB} and B_C.

Energy landscape across environmental gradient - If the occurrence of species depends on some abiotic factors, change in the constraints on community assembly dynamics across the environmental gradient can be captured by the change in the energy landscape. We introduced an abiotic factor (ϵ) to the previous model and studied the change in the energy landscape over $\epsilon \in [0,1]$. Figure 4a-f shows the snapshot of the energy landscape for different ϵ values, and figure 4g is a stable state diagram that shows the energy of stable states (solid lines) and tipping points (dashed lines). This result shows that C was no longer a stable state when $\epsilon > 0.62$, and A and B became more stable with increasing ϵ. Further, alteration of community composition of A, B and b1 to A', B' and b1', respectively, occurred with increasing ϵ.

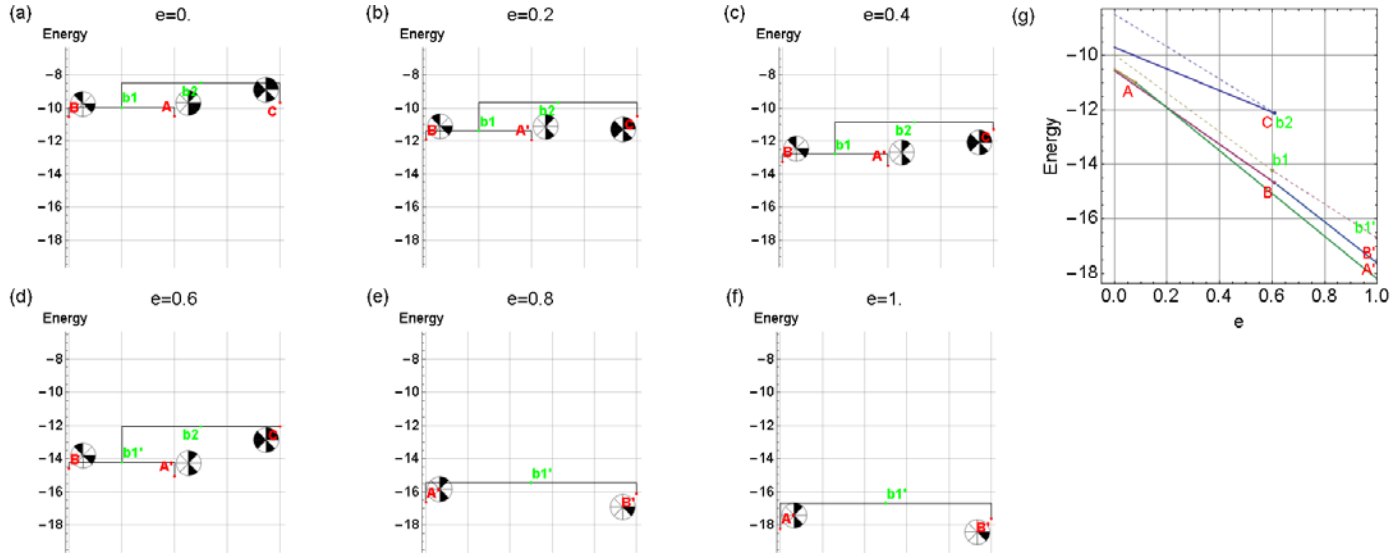

Figure 4. Energy landscape of an eight species metacommunity subject to an abiotic factor (ϵ). (a-f) disconnectivity graph for different ϵ values (stable states (red) and tipping points (green)). Pie chart indicates presence (white) or absence (black) of species $i = 1, \ldots, 8$ in clockwise order. (g) stable state diagram showing the energy of stable states and tipping points. Here, energy of stable states (solid lines, community state is indicated by red letters) and tipping points (dashed lines, community state is indicated by green letters) are shown. Each line segments (labeled by letters identifying stable states or tipping points) represent the range of age with stable states (or tipping points).

Inferring energy landscape from observational data - In reality, the energy landscape analysis will be applied after inferring the energy landscape from observational data. We examined the relationship between the accuracy of the inferred energy landscape and the size of dataset used for the inference by generating the dataset from the presented models. We evaluated the accuracy using Spearman's rank correlation (ρ) and relative mean squared error (RMSE) between the energy of all community states specified by the pairwise maximum entropy model with actual and inferred parameter values. In the eight species metacommunity model either without (Fig. 5a, b) or with (Fig. 5c-n) an abiotic factor, both ρ and RMSE were mostly saturated when dataset size was larger than 200 (78% of the total number of possible community states, i.e., $2^8 = 256$), and the result was still reliable even if the data set size was 100 (39% of the total number of possible community states).

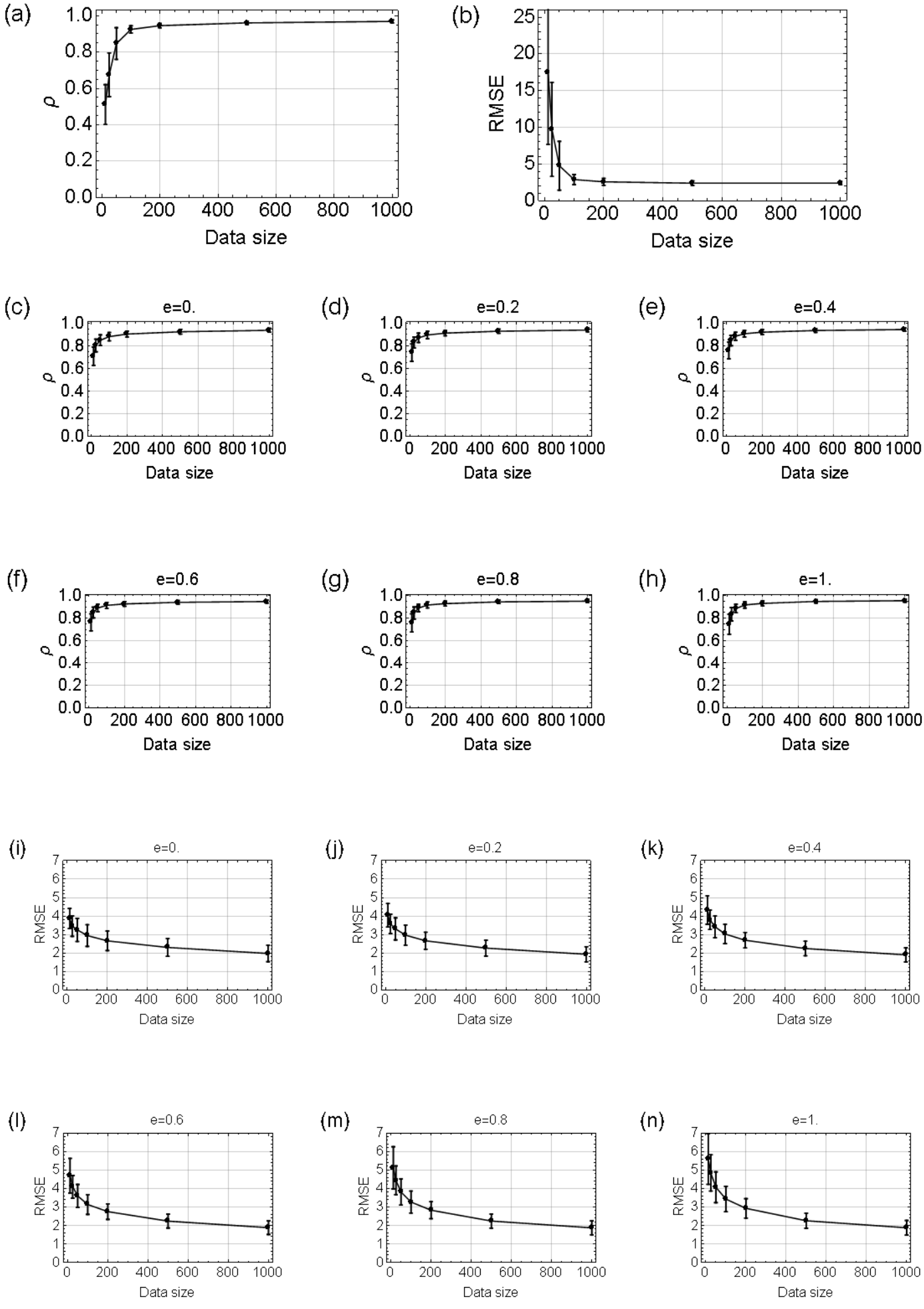

Figure. 5. Inferring energy landscapes using datasets of varying sizes. (a,b) spearman rank correlation and RMSE (root mean squared error) of inferred and actual energy of all community states calculated for 25, 50, 100, 200, 500, 1000 dataset size with 100 independent trials. In each trial, we repeated 25-1000 independent runs of 10^4 steps heat-bath simulations, then we picked up the final community states as a dataset to infer the parameters of the pairwise maximum entropy model using gradient descent algorithm (Appendix A). (c-h) spearman rank correlation of inferred and actual energy of all community states calculated for 25, 50, 100, 200, 500, 1000 dataset size with 100 independent trials. (i-n) RMSE of inferred and actual energy of all community state calculated for 25, 50, 100, 200, 500, 1000 dataset size with 100 independent trials. To obtain (c-n), in each trial we repeat 25-1000 independent runs of 10^4 steps heat-bath simulations where ϵ values are randomly assigned from the uniform distribution [0,1], then we picked up the final community states as a dataset to infer the parameters of the pairwise maximum entropy model using stochastic approximation algorithm (Appendix A).

Application to real data

These results show how energy landscapes can be used to study community assembly dynamics and indicate 100 samples are sufficient to infer the energy landscape of systems with eight species. We applied this approach to the gut microbiome data of six mice with 96 samples and concluded that this dataset was sufficient to apply our approach. Community states were represented by eight genus level OTUs that can be identified as species in the above analysis, and we used age of mice as the abiotic factor.

At the community level, all stable states found at initial age (C_{40}, C_{44}, C_{138} in Fig. 6a; see Table 2 for the membership of community states) increased energy as age increased. C_{138} was no longer a stable state after 20 weeks because its energy exceeded that of the tipping point (C_{140}) between C_{12} and C_{138} (Fig. 6a). C_{44} changed to C_{12} at 4 weeks age and then showed reduced energy with increased age. The lowest energy minimum was altered from C_{138} to C_{12} at 12 weeks age. C_{40} showed reduced energy with increased age when it changed to C_8 at 12 weeks age.

These community level responses can be understood by analyzing the estimated parameters. In figure 6b, the community level response is shown by the bars marked as 'Total' in addition to the genus level responses. The net effect of biotic interactions ($\Sigma_j J_{ij}\sigma_i^{(k)}\sigma_j^{(k)}$), the bacterial responses to age ($g_i\epsilon$), and the implicit effect of abiotic factors (h_i; here, these values represent the sum of the effects other than the bacterial response to age or the biotic interaction between bacteria) are indicated by blue, orange and green, respectively. The correlation between age and the occurrence of genus became positive if g (genus level response to age; Table 2) was positive; this correlation became negative when g was negative. The sum of g across community members determined the community level response to age. *Bifidobacterium* (in C_{138}), *Turicibacter* (in C_{40} and C_{44}) and unclassified RF39 (in all) had negative g

values (Table 2). The transition from C_{40} to C_8 and C_{44} to C_{12} occurred as the alteration of stable states between those containing *Turicibacter* (C_{40} or C_{44}) and those without *Turicibacter* (C_8 or C_{12}) since its absence altered the sign of the community level response to age (Table2; in Fig. 6b, the community level response to age is positive for C_8 or C_{12} while the same value was negative for C_{40} and C_{44}). *Turicibacter* disappeared later in C_{40} than in C_{44} (Fig. 6a) because it had a stronger positive relationship with *Oscillatospira* (in C_{40}) than unclassified Ruminococcaceae (in C_{44}) (Fig 6c). The net effect of biotic interactions to *Turicibacter* was negative in both cases since other members, i.e., *Suttellela* and RF39, had a negative relationship with *Turicibacter* (Fig. 6b,c; in figure 6b, the net effect of biotic interactions was negative for *Turicibacter* both in C_{40} and C_{44}). RF39 also had a negative g value, though it remained present in C_8 and C_{12} because its occurrence largely depended on h (Table 2, Fig 6b). The difference between C_8 and C_{12} was presence of *Oscillatospira* (C_8) or Ruminococcaceae (C_{12}) (Table 2). The two genera had a negative relationship with each other (Fig. 6c). Thus, they could be mutually exclusive. Interestingly, the tipping point between C_8 and C_{12}, i.e., C_{24} or C_{52} (Fig. 6a, Table2), contained *Lachnospiraceae* that was not included in any stable states.

Similarly, C_{138} lost its stability due to the negative response of *Bifidobacterium* to age (Table 2, Fig. 6b). The energy of C_{138} increased faster than that of C_{140}, i.e., tipping point between C_{12} and C_{138} (Fig. 6a), because presence of *Suterella* (positively affected by age) in C_{140} reduced the inclination of its community level response to age. Different from C_{138}, C_{12} continued to be a stable state over age because it included *Sutterella* instead of *Bifidobacterium* (Table 2), resulted in the alteration of the lowest energy minimum from C_{138} to C_{12} at 12 weeks age.

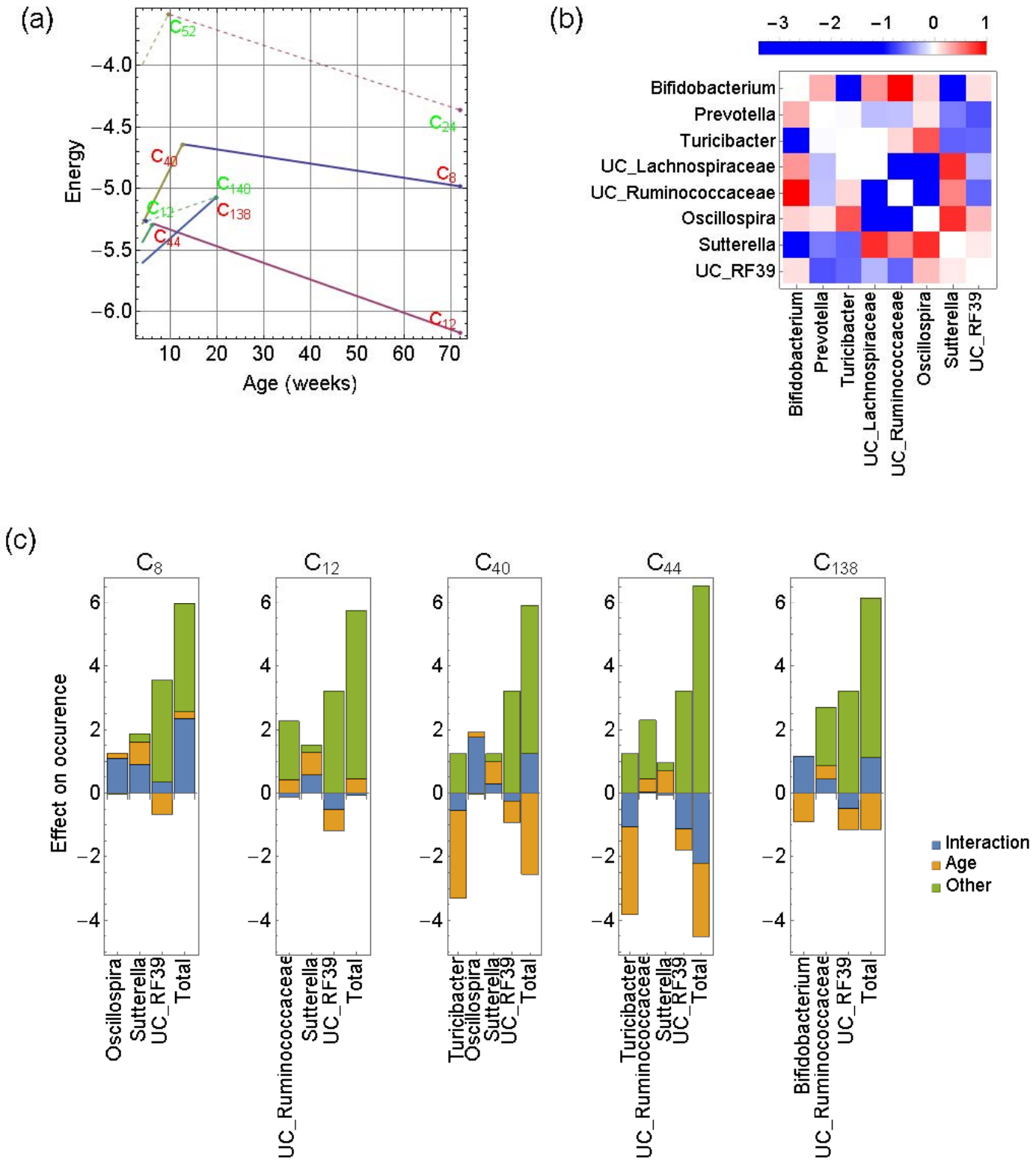

Figure 6. Energy landscape analysis of a mouse gut microbiota. (a) stable state diagram showing the energy of stable states and tipping points. Here, energy of stable states (solid lines, community state is indicated by red letters) and tipping points (dashed lines, community state is indicated by green letters) are shown. Each line segment (labeled by community state C_x at their right end) represents

the range of age with stable states (or tipping points). Subscripts in C_X identifies a community type where binary vectors are converted to a decimal number, e.g., $(0,0,0,0,0,0,0,0) = 1$, $(0,0,0,0,0,0,0,1) = 2$, etc.. (b) Strength of the net effect of biotic interactions ($\Sigma_j J_{ij}\sigma_i^{(k)}\sigma_j^{(k)}$), explicit abiotic factor (response to age) ($g_i\epsilon$) and implicit abiotic factors (h_i) for each genus i (genus level effects) and their sum over community members (community level effects), $\Sigma_i\Sigma_j J_{ij}\sigma_i^{(k)}\sigma_j^{(k)}$, $\Sigma_i g_i\epsilon\sigma_i^{(k)}$ and $\Sigma_i h_i\sigma_i^{(k)}$ (shown as 'total'). $\sigma_i^{(k)}$ represents membership of each communities (Table 2). For comparison, we set $\epsilon = 0.5$. (c) elements of biotic interactions (J_{ij}). The value represents the strength of association between two genera (shown in columns and rows). There is a positive association between two genera if the value is positive whereas there is negative association if it is negative.

	h	G	C_8	C_{12}	C_{40}	C_{44}	C_{138}	C_{12}*	C_{24}*	C_{52}*	C_{140}*
Bifidobacterium	0.003	-1.814	0	0	0	0	1	0	0	0	1
Prevotella	-0.509	-0.337	0	0	0	0	0	0	0	0	0
Turicibacter	1.246	-5.512	0	0	1	1	0	0	0	1	0
UC_Lachnospiraceae	-0.114	0.458	0	0	0	0	0	0	1	1	0
UC_Ruminococcaceae	1.84	0.844	0	1	0	1	1	1	0	0	1
Oscillospira	-0.038	0.318	1	0	1	0	0	0	1	0	0
Sutterella	0.249	1.395	1	1	1	1	0	1	1	1	1
UC_RF39	3.203	-1.321	1	1	1	1	1	1	1	1	1

Table 2. Profile of eight genus included in the analysis. h (implicit effect from abiotic factors) and g (genus level response to age) are the parameters inferred by stochastic approximation (Appendix A) and 0/1 values indicate membership of each genus in stable states and tipping points (indicated by *). C_{12} appeared both as a stable state and tipping point since it was a tipping point at 4 weeks of age (Fig. 6a).

Discussion

We developed a framework to study community assembly dynamics by incorporating a pairwise maximum entropy model (Azaele et al. 2010, Araujo et al. 2011, Harris 2016) and an energy landscape analysis (Becker and Karplus 1997). Using a pairwise maximum entropy model, a set of community states can be seen as an energy landscape, which is a network with nodes representing community states and links representing their transitions. The nodes are weighted by their energy and characterize the energy landscape. Energy landscape analysis incorporated the number, composition and basin size of stable states (energy minima) as well as tipping points between them and overall basin boundary structure (represented by boundary pairs). Disconnectivity graphs (Fig. 2a, 4a) showed the hierarchical relationships among alternative stable states and tipping points. Therefore, our

approach provides systematic understanding on the constraints of community assembly dynamics by embedding the relationship of alternative stable states and other community states as the structure of energy landscape. We also introduced a heat-bath (Gibbs sampling) algorithm (Gilks et al. 1996) to emulate community assembly dynamics constrained on the energy landscape. We used it to evaluate how transition paths between two alternative stable states were constrained by the basin boundary structure. In our simulation, only 5.3% of basin boundary structure mediated 56.7% of realized transitions (Fig. 3c). This suggests that a small fraction of community states may be a channel for transition between alternative stable states, which would be a basis for developing an early warning signal (Scheffer et al. 2012) for community level transitions. Furthermore, we showed that the sum of energy that must go up over a path connecting two stable states ('path energy', PE) can inform how easy the transition will occur with these sequences. It will help us to find reliable assembly sequences by which we can easily change one community state to another. The presented approach can include one or more abiotic factors that affect species occurrence, and thus addresses change in energy landscape structure along these environmental axes. We introduced a stable state diagram to summarize the number and energy of stable states and tipping points (Fig. 4g, 6a), which indicated change in constraints on community assembly dynamics along with the environmental gradient. We also confirmed that energy landscapes can be reconstructed from observational data if dataset size was sufficient. In short, our approach provided information on the constraints on community assembly dynamics (Law and Morton 1993, Warren et al. 2003, Capitan et al. 2011), described change across environmental gradients, and allowed for simulation-based analysis.

In the mouse gut microbiota, our approach showed a major shift in energy landscape structure that occurred during early to middle age, which accompanied the disappearance of *Bifidobacterium* and *Turicibacter* from the stable states. This suggested an age-related regime shift in mouse gut microbiota as was previously reported for human gut microbiota (Lahti et al. 2014). Lahti et al. (2014) showed evidence of 'tipping elements' whose presence/absence status control community structure in the human gut microbiota across age. Our result suggested that, at least for its compositional shift in early to middle age, *Bifidobacterium* and *Sutterella* may be the potential 'tipping elements' in the mouse gut microbiota since their presence/absence status in stable states were altered across life stage. In human gut microbiota, *Bifidobacterium* is introduced as a beneficial bacterium. However, it usually disappears within a few days and the 'probiotic' effects do not last beyond a few days without its continuous supplementation (Kim et al. 2013). Our results suggest that in mouse gut microbiota, the same thing may be observed due to the lack of stable states including *Bifidobacterium*. Attempts to establish a *Bifidobacterium* population will not succeed without finding environmental conditions that support at least one stable state including *Bifidobacterium*. Our result also suggested that *Lachnospiraceae* may be a catalytic taxon (Warren et al. 2003) which appeared in the tipping point but was absent from any stable states. In this analysis, we also revealed how the genus level differences in response to age, the effect of implicit abiotic factors, and interactions with other genera are responsible for community level responses to age and result in age dependence of community assembly processes. The response of each genus to age naturally had a prominent role in the community level response,

though interaction with other factors was also identified. The approach outlined here provided mechanistic understanding of the processes that drive community assembly dynamics. Overall, this approach allowed us to systematically and mechanistically analyze community assembly dynamics across environmental gradient.

Comparison to other approaches

Community structures across spatial and temporal scale have been studied by species distribution models (SDMs) or joint species distribution models (JSDMs) (Norberg et al. 2019, Wilkinson et al. 2019). However, these models do not directly include biotic interactions and thus it is controversial whether they are reliable to study mechanisms that shape species assemblage (Barner et al. 2018, Freilich et al. 2018). More importantly, these models do not directly describe dynamic relationships between different community states (Baselga and Araújo 2009, Elith and Leathwick 2009, Norberg et al. 2019). Therefore, they cannot be perceived as models of community assembly dynamics. Dynamical models (such as differential equations) are able to describe shifts in community structure based on the change in population abundance of a species (Gravel et al. 2011). However, it is generally difficult to develop fully mechanistic models for multi-species communities from time-series abundance data. On the other hand, our model considers only the consequence of species extinctions or migrations in local communities. Hence, it ignores the time scale of transient population dynamics and approximates the continuous state space as a network of community states. By discarding detailed descriptions of transient dynamics, our approach offers a practical way to study the community assembly dynamics from observational data.

Ecological implications

Community assembly dynamics play a prominent role in real world ecosystem organization, and therefore the methodological advancement we presented here will influence conceptual development in other important topics in ecology. For example, regime shifts in ecology have mainly addressed ecological states over one environmental axis (Scheffer et al. 2001, Beisner et al. 2003, Walker et al. 2004). The concept can be extended to a multi-species system (Shoroder et al. 2005) but it is typically assumed that the system is dominated by a few ecological variables and control parameters, as in the relationship of phosphorus concentration with abundance of phytoplankton and plant species in lake systems (Scheffer and Jeppsen 2007). Attempts to analyze regime shifts in real world multi-species systems are limited by development of full mechanistic models based on observational data. Our approach offers a practical way to understand how complex stability landscape (Walker et al. 2004) of a multi-species system could change across environmental axes and trigger regime shifts. This ability will have important implications for ecosystem management. Further, our approach can provide both systematic and mechanistic understanding of the processes that structure and maintain spatiotemporal heterogeneity in the composition of ecosystems, and thus will make a significant contribution to conservation of ecosystem function and services.

Limitations and challenges

Although our approach has great potential to advance studies of community assembly and related fields, it still has some flaws that need to be addressed. First, further verification is required to assess the effect of replacing species' abundance with presence/absence status. Does the pairwise maximum model always provide good approximation to the global phase space of ecological dynamics? For example, since our approach relies on a gradient system, it cannot account for attractor dynamics that often appear in ecological systems. Heteroclinic cycles that cause cyclic alteration of community states (Morton and Law 1997, Fukami et al. 2015) are one examples. These dynamics may still be identified as a set of energy minima separated by low energy barriers, though the overall consequence of approximating dynamics in a continuous phase space into a coarse-grained phase space (where nodes of the weighted network represent each sub-system of the original phase space) is unknown. Second, it is not clear if the principle of maximum entropy actually fits the distribution of community states in natural species assemblages. For example, due to disturbance and patch dynamics, community states may be in different developmental stages at different sites at any given point in time. Stochastic consequences of biotic interactions would also be responsible for such non-equilibrium community dynamics (e.g., see Warren et al. 2003). Third, increase of the number of species (S) included in the analysis will cause discrepancy between the available dataset size and the number of possible community states (2^S) and this might reduce the accuracy of the pairwise maximum entropy model for species rich systems. If community dynamics occurred only within a fraction of phase space, the discrepancy between dataset size and possible community states will not significantly reduce the performance of the pairwise maximum entropy model. More information is required on the relationship between the number of species and the proportion of phase space needed to explain community dynamics in ecological systems. Last, causal relationships between species' presence/absence status and transition of one community state to another are not well represented using our approach. Incorporating causal analysis (e.g., Sugihara et al. 2012, Runge et al. 2017) will strengthen our approach especially when considering its application to control community states.

Conclusion

There is an urgent need to move ecology from empirical and conceptual work to application and management issues (Mouquet et al. 2015). Although some further verification and improvement is required, we believe that the methodological advancement presented here will be a new systemic paradigm for developing a predictive theory of community ecology (Long and Karel 2002, Chase 2003, Fukami 2010).

Acknowledgements

This work was supported by the Management Expenses Grant for RIKEN BioResource Research Center, MEXT, and in part by the Center of Innovation Program from Japan Science and Technology Agency (JST) (to K.S. and S.N.), JST PRESTO Grant Number JPMJPR16E9 (to S.N.), the Japan Society for the Promotion of Science (JSPS) (S) JP15H05707 (to S.N.), JSPS KAKENHI (18H04805 to S.F.), JST PRESTO (JPMJPR1537 to S.F.), AMED-CREST (JP19gm1010009 to S.F.), the Takeda

Science Foundation (to S.F.) and the Food Science Institute Foundation (to S.F.). We declare that we have no conflict of interest.

References

1. Drake, J. A. (1991). Community-assembly mechanics and the structure of an experimental species ensemble. *The American Naturalist*, 137(1), 1-26.
2. Fukami, T. 2010. Community assembly dynamics in space. In Community ecology: Processes, models, and applications. Edited by Herman A. Verhoef and Peter J. Morin, 45–54. Oxford: Oxford Univ. Press.
3. Fukami, T., Morin, P. J. (2003). Productivity–biodiversity relationships depend on the history of community assembly. *Nature*, 424(6947), 423.
4. Jiang, L., Joshi, H., Flakes, S. K., Jung, Y. (2011). Alternative community compositional and dynamical states: the dual consequences of assembly history. *Journal of Animal Ecology*, 80(3), 577-585.
5. Fukami, T. (2015). Historical contingency in community assembly: integrating niches, species pools, and priority effects. *Annual Review of Ecology, Evolution, and Systematics*, 46, 1-23.
6. Fukami, T., Dickie, I. A., Paula Wilkie, J., Paulus, B. C., Park, D., Roberts, A., et al. (2010). Assembly history dictates ecosystem functioning: evidence from wood decomposer communities. *Ecology letters*, 13(6), 675-684.
7. Schröder, A., Persson, L., De Roos, A. M. (2005). Direct experimental evidence for alternative stable states: a review. *Oikos*, 110(1), 3-19.
8. Pu, Z., Jiang, L. (2015). Dispersal among local communities does not reduce historical contingencies during metacommunity assembly. *Oikos*, 124(10), 1327-1336.
9. Vellend, M. (2010). Conceptual synthesis in community ecology. *The Quarterly review of biology*, 85(2), 183-206.
10. Weiher, E., Freund, D., Bunton, T., Stefanski, A., Lee, T., Bentivenga, S. (2011). Advances, challenges and a developing synthesis of ecological community assembly theory. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 366, 2403–2413.
11. Gotzenberger, L., de Bello, F., Bråthen, K.A., Davison, J., Dubuis, A., Guisan, A. et al. (2012). Ecological assembly rules in plant communities- approaches, patterns and prospects. *Biol. Rev.*, 87, 111–127.
12. HilleRisLambers, J., Adler, P.B., Harpole, W.S., Levine, J.M., Mayfield, M.M. (2012). Rethinking community assembly through the lens of coexistence theory. *Annu. Rev. Ecol. Evol. Syst.*, 43, 227–248.
13. Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F. et al. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biol. Rev.*, 88, 15–30.
14. Law, R., Morton, R. D. (1993). Alternative permanent states of ecological communities. *Ecology*, 74(5), 1347-1361.

15. Capitán, J. A., Cuesta, J. A., Bascompte, J. (2011). Statistical mechanics of ecosystem assembly. *Physical Review Letters*, 103(16), 168101.
16. Warren, P. H., Law, R., Weatherby, A. J. (2003). Mapping the assembly of protist communities in microcosms. *Ecology*, 84(4), 1001-1011.
17. Weatherby, A. J., P. H. Warren, R. Law. 1998. Coexistence and collapse: an experimental investigation of the persistent communities of a protist species pool. *Journal of Animal Ecology* 67:554–566
18. Law, R., A. J. Weatherby, P. H. Warren. 2000. On the invasibility of persistent protist communities. *Oikos* 88: 319–326.
19. Morton RD, Law R. 1997. Regional species pools and the assembly of local ecological communities. *J. Theor. Biol.* 187:321–31
20. Azaele, S., Muneepeerakul, R., Rinaldo, A., Rodriguez-Iturbe, I. (2010). Inferring plant ecosystem organization from species occurrences. *Journal of theoretical biology*, 262(2), 323-329.
21. Araujo MB, Luoto M. The importance of biotic interactions for modelling species distributions under climate change. *Glob. Ecol. Biogeogr.* 2007; 16:743–753.
22. Harris, D. J. (2016). Inferring species interactions from co - occurrence data with Markov networks. *Ecology*, 97(12), 3308-3314.
23. Becker, O. M., Karplus, M. (1997). The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics. *The Journal of chemical physics*, 106(4), 1495-1517.
24. Wales, D. J., Miller, M. A., Walsh, T. R. (1998). Archetypal energy landscapes. *Nature*, 394(6695), 758.
25. Watanabe, T., Hirose, S., Wada, H., Imai, Y., Machida, T., Shirouzu, I., Konishi, S., Miyashita, Y., Masuda, N. (2014a). Energy landscapes of resting-state brain networks. *Frontiers in Neuroinformatics*, 8, 12.
26. Barner, A. K., Coblentz, K. E., Hacker, S. D., Menge, B. A. (2018). Fundamental contradictions among observational and experimental estimates of non - trophic species interactions. *Ecology*, 99(3), 557-566.
27. Watanabe, T., Masuda, N., Megumi, F., Kanai, R., Rees, G. (2014b). Energy landscape and dynamics of brain activity during human bista- ble perception. *Nature Communications*, 5, 4765.
28. Ezaki, T., Watanabe, T., Ohzeki, M., Masuda, N. (2017). Energy landscape analysis of neuroimaging data. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 375(2096), 20160287.
29. Watanabe, T., Rees, G. (2017). Brain network dynamics in high- functioning individuals with autism. *Nature Communications*, 8, 16048.
30. Ezaki, T., Sakaki, M., Watanabe, T., Masuda, N. (2018). Age–related changes in the ease of dynamical transitions in human brain activity. *Human brain mapping*, 39(6), 2673-2688.
31. Drake, J. A. (1990). Communities as assembled structures: do rules govern pattern? *Trends in Ecology and Evolution* 5:159–164.
32. Pimm, S. L. (1991). The balance of nature. University of Chicago Press, Chicago, Illinois, USA.

33. Lockwood, J. L., Pimm, S. L. (1999). When does restoration succeed. Ecological assembly rules: perspectives, advances, retreats, 363-392.
34. Weiher, E., Keddy, P. (Eds.). (2001). Ecological assembly rules: perspectives, advances, retreats. Cambridge University Press.
35. Lockwood JL, Samuels CL. (2004). Assembly models and the practice of restoration. In Assembly Rules and Restoration Ecology: Bridging the Gap Between Theory and Practice, ed. VM Temperton, RJ Hobbs, T Nuttle, S Halle, pp. 55–70. Washington, DC: Island Press
36. Suding KN, Gross KL, Houseman GR. (2004). Alternative states and positive feedbacks in restoration ecology. *Trends Ecol. Evol.* 19:46–53
37. Wilsey BJ, Barber K, Martin LM. (2015). Exotic grassland species have stronger priority effects than natives regardless of whether they are cultivated or wild genotypes. *New Phytol.* 205:928–37
38. Young TP, Petersen DA, Clary JJ. (2005). The ecology of restoration: historical links, emerging issues and unexplored realms. *Ecol. Lett.* 8:662–73
39. Young, TP, Zefferman EP, Vaughn KJ, Fick S. (2015). Initial success of native grasses is contingent on multiple interactions among exotic grass competition, temporal priority, rainfall, and site effects. *AoB PLANTS* 7:081
40. Verbruggen E, van der Heijden MG, Rillig MC, Kiers ET. (2013). Mycorrhizal fungal establishment in agricultural soils: factors determining inoculation success. *New Phytol.* 197:1104–9
41. Toju, H., Peay, K. G., Yamamichi, M., Narisawa, K., Hiruma, K., Naito, K., et al. (2018). Core microbiomes for sustainable agroecosystems. *Nat. Plants*, 4(5), 247-257.
42. Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J., Relman, D. A. (2012). The application of ecological theory toward an understanding of the human microbiome. *Science*, 336(6086), 1255-1262.
43. Fierer, N., Ferrenberg, S., Flores, G. E., González, A., Kueneman, J., Legg, T., et al. (2012). From animalcules to an ecosystem: application of ecological concepts to the human microbiome. *Annual Review of Ecology, Evolution, and Systematics*, 43, 137-155.
44. Lam, L. H., Monack, D. M. (2014). Intraspecies competition for niches in the distal gut dictate transmission during persistent Salmonella infection. *PLoS pathogens*, 10(12), e1004527.
45. Devevey, G., Dang, T., Graves, C. J., Murray, S., Brisson, D. (2015). First arrived takes all: inhibitory priority effects dominate competition between co-infecting Borrelia burgdorferi strains. *BMC microbiology*, 15(1), 61.
46. Smits LP, Bouter KE, de Vos WM et al. (2013). Therapeutic potential of fecal microbiota transplantation. *Gastroenterology* 145:946–53.
47. Li SS, Zhu A, Benes V et al. (2016). Durable coexistence of donor and recipient strains after fecal microbiota transplantation. *Science* 352:586–9.
48. Shetty, S. A., Hugenholtz, F., Lahti, L., Smidt, H., de Vos, W. M. (2017). Intestinal microbiome landscaping: insight in community assemblage and implications for microbial modulation strategies. *FEMS microbiology reviews*, 41(2), 182-199.

49. Karczewski, J., Poniedziałek, B., Adamski, Z., Rzymski, P. (2014). The effects of the microbiota on the host immune system. *Autoimmunity*, 47(8), 494-504.
50. Cox, L. M., Blaser, M. J. (2015). Antibiotics in early life and obesity. *Nature Reviews Endocrinology*, 11(3), 182.
51. Tang, A. T., Choi, J. P., Kotzin, J. J., Yang, Y., Hong, C. C., Hobson, N., et al. (2017). Endothelial TLR4 and the microbiome drive cerebral cavernous malformations. *Nature*, 545(7654), 305.
52. Van Nood, E., Vrieze, A., Nieuwdorp, M., Fuentes, S., Zoetendal, E. G., de Vos, W. M., et al. (2013). Duodenal infusion of donor feces for recurrent Clostridium difficile. *New England Journal of Medicine*, 368(5), 407-415.
53. Jaynes, E.T., (1982). On the rationale of maximum-entropy methods. Proc. IEEE 70, 939–952.
54. Mead, L.R., Papanicolau, N., (1984). Maximum entropy in the problem of moments. *J. Math. Phys.* 25, 2404–2417.
55. Jaynes, E.T., (2003). Probability Theory. Cambridge University Press, Cambridge.
56. Brush, S. G. (1967). History of the Lenz-Ising model. *Reviews of modern physics*, 39(4), 883.
57. Nguyen, H. C., Zecchina, R., Berg, J. (2017). Inverse statistical problems: from the inverse Ising problem to data science. *Advances in Physics*, 66(3), 197-261.
58. Jaynes, E.T., (1957a). Information theory and statistical mechanics. *Phys. Rev.* 106, 620–630.
59. Jaynes, E.T., (1957b). Information theory and statistical mechanics. II. *Phys. Rev.* 108, 171–190.
60. Schneidman, E., Berry II, M.J., Segev, R., Bialek, W., (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440, 1007–1012.
61. Lezon, T., et al., (2006). Using the principle of entropy maximization to infer genetic interaction networks from gene expression patterns. *Proc. Natl. Acad. Sci. USA* 103, 19033–19038.
62. Shipley, B., Vile, D., Garnier, E´., (2006). From plant traits to plant communities: a statistical mechanistic approach to biodiversity. *Science* 314, 812–814.
63. Dewar, R.C., Porte´, A., (2008). Statistical mechanics unifies different ecological patterns. *J. Theor. Biol.* 251, 389–403.
64. Harte, J., Zillio, T., Conslik, E., Smith, A.B., (2008). Maximum entropy and the state- variable approach to macroecology. *Ecology* 89 (10), 2700–2711.
65. Elith, J., Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677–697.
66. Warton, D.I., Blanchet, F.G., O'Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C., Hui, F.K.C. (2015) So many variables: joint modeling in community ecology. *Trends in Ecology & Evolution*, 30, 766–779.
67. Hanski, I., Gilpin, M. (1991). Metapopulation dynamics: brief history and conceptual domain. *Biological journal of the Linnean Society*, 42(1-2), 3-16.
68. Gilks, W. R., Richardson, S., Spiegelhalter, D. J. (1996). Introducing Markov chain Monte Carlo. Markov chain Monte Carlo in practice, 1, 19.
69. Nakanishi et al. (submitted)
70. Gerber, G. K. (2014). The dynamic microbiome. *FEBS letters*, 588(22), 4131-4139.

71. Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., et al. (2012). Anticipating critical transitions. *Science*, 338(6105), 344-348.
72. Lahti, L., Salojärvi, J., Salonen, A., Scheffer, M., De Vos, W. M. (2014). Tipping elements in the human intestinal ecosystem. *Nature communications*, 5, 4344.
73. Kim, S.-W. et al. (2013). Robustness of gut microbiota of healthy adults in response to probiotic intervention revealed by high-throughput pyrosequencing. *DNA Res*. 20, 241–253.
74. Norberg, A., Abrego, N., Blanchet, F. G., Adler, F. R., Anderson, B. J., Anttila, J., et al. (2019). A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological Monographs*, e01370.
75. Wilkinson, D. P., Golding, N., Guillera - Arroita, G., Tingley, R., McCarthy, M. A. (2019). A comparison of joint species distribution models for presence–absence data. *Methods in Ecology and Evolution*, 10(2), 198-211.
76. Freilich, M. A., Wieters, E., Broitman, B. R., Marquet, P. A., Navarrete, S. A. (2018). Species co - occurrence networks: Can they reveal trophic and non - trophic interactions in ecological communities?. *Ecology*, 99(3), 690-699.
77. Baselga, A., and M. B. Araújo. (20099. Individualistic vs. community modelling of species distributions under climate change. *Ecography* 32: 55-65.
78. Gravel, D., Guichard, F., Hochberg, M. E. (2011). Species coexistence in a variable world. *Ecology Letters*, 14(8), 828-839.
79. Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, 413(6856), 591.
80. Beisner, B. E., Haydon, D. T. Cuddington, K. Alternative stable states in ecology. *Front. Ecol. Env.* 1, 376–382 (2003).
81. Walker, B., Holling, C. S., Carpenter, S., Kinzig, A. (2004). Resilience, adaptability and transformability in social–ecological systems. *Ecology and society*, 9(2).
82. Scheffer, M., Jeppesen, E. (2007). Regime shifts in shallow lakes. *Ecosystems*, 10(1), 1-3.
83. Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, 338(6106), 496-500.
84. Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., Sejdinovic, D. (2017). Detecting causal associations in large nonlinear time series datasets. arXiv preprint arXiv:1702.07007.
85. Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D., et al. (2015). Predictive ecology in a changing world. *Journal of Applied Ecology*, 52(5), 1293-1310.
86. Long, Z.T. and Karel, I. (2002) Resource specialization determines whether history influences community structure. *Oikos*, 96, 62–9.
87. Chase, J.M and Leibold, M.A. (2003) Ecological Niches: Linking Classical and Contemporary Approaches. University of Chicago Press, Chicago, IL

Appendix A

The maximum likelihood estimate for the model parameters can be obtained by minimizing the discrepancy between the values of the data's sufficient statistics and the corresponding sufficient statistics within the model (Bickel and Doksum 1977, Azaele et al. 2010, Murphy 2012, Harris 2015, Lee and Hastie 2015).

Gradient descent algorithm

For a dataset that does not include explicit abiotic factors and energy can be calculated by eq. (2), a gradient descent algorithm can be applied (Watanabe et al. 2014a,b, Harris 2015, Harris 2016). For a model with parameter h^* and J^*, let the expected probability of species i be $\langle\sigma_i\rangle^* = 1/2^S \sum_{k=1}^{2^S} \sigma_i^{(k)} p(\vec{\sigma}^{(k)})$ and the co-occurrence be $\langle\sigma_i\sigma_j\rangle^* = 1/2^S \sum_{k=1}^{2^S} \sigma_i^{(k)} \sigma_j^{(k)} p(\vec{\sigma}^{(k)})$. The parameter h and J can be fitted to the data by iteratively adjusting $\langle v_i\rangle^*$ and $\langle v_i v_j\rangle^*$ toward $\langle v_i\rangle$ and $\langle v_i v_j\rangle$, respectively, by updating the parameters as

$$h_i^{new} \leftarrow h_i^{old} + \alpha \log\langle\sigma_i\rangle/\langle\sigma_i\rangle^*$$
$$J_{ij}^{new} \leftarrow J_{ij}^{old} + \log\langle\sigma_i\sigma_j\rangle/\langle\sigma_i\sigma_j\rangle^*$$

at each step. Here, $\langle\sigma_i\rangle$ and $\langle\sigma_i\sigma_j\rangle$ is mean occurrence and co-occurrence calculated from the observational data. We set $\alpha = 0.25$ and the number of iteration as 5,000.

Stochastic approximation

The likelihood function of the pairwise maximum entropy model becomes computationally intractable when we need to include explicit abiotic factors (as in eq. (4)), because it requires repeating above computations independently for every sample. Therefore, it calls for a different model-fitting algorithm that is computationally more efficient. Here, following Harris (2015), we introduce a 'stochastic approximation' (Robbins and Monro 1951, Salakhutdinov and Hinton 2012) for this purpose. This algorithm replaces the intractable computations with tractable Monte Carlo estimates of the same quantities. Despite the sampling error introduced by this substitution, stochastic approximation provides strong guarantees for eventual convergence to the maximum likelihood estimate (Younes 1999, Salakhutdinov and Hinton 2012).

Stochastic approximation (Robbins and Monro 1951, Salakhutdinov and Hinton 2012) estimates the expected values of the sufficient statistics by averaging over a more manageable number of simulated assemblages during each model-fitting iteration, while still retaining maximum likelihood convergence guarantees. The advantage of this algorithm is Z (eq. (3)) does not have to be calculated at each step, which significantly improves computational efficiency. This is due to the use of a heat-bath algorithm that only requires calculating energy of two adjacent community states. The procedure iterates through the following three steps as many times as needed (Here, we set $T = 50000$ for these analyses).

Stochastic approximation

1. Set t=0, and initial parameter values for h, j and g.

2. Calculate learning rate l as:

$$\mathrm{l} = \mathrm{l}_0 \frac{1000}{999 + t},$$

momentum m as:

$$m = 0.9\left(1 - \frac{1}{0.1\mathrm{t} + 2}\right),$$

and logistic priors as $\mathrm{p_h} = -\tanh(\mathrm{h}/2/2)/2$, $\mathrm{p_g} = -\tanh(\mathrm{g}/2/2)/2$ and $\mathrm{p_J} = -\tanh(\mathrm{J}/0.5/2)/2$. Here, l_0 is the initial learning rate and we set it as 1.

3. For $\mathrm{k} = 1$ to M, run one step heat-bath algorithm based on current parameters (h, J and g): transition from the current community state $\vec{\sigma}^{(k)}$ to one of its S adjacent community state $\vec{\sigma}^{(k')}$, selected with probability $1/\mathrm{S}$, was attempted ($\vec{\sigma}^{(k)}$ and $\vec{\sigma}^{(k')}$ differs only with respect to the presence/absence status of one of S species). The transition to the selected state took place with probability $\mathrm{e}^{-\mathrm{E}\left(\vec{\sigma}^{(k')}\right)}/(\mathrm{e}^{-\mathrm{E}(\vec{\sigma}^{(k)})} + \mathrm{e}^{-\mathrm{E}(\vec{\sigma}^{(k')})})$ ($\mathrm{e}^{-\mathrm{E}(\vec{\sigma}^{(k)})}$ and $\mathrm{e}^{-\mathrm{E}(\vec{\sigma}^{(k')})}$ are given by eq. (4)).

4. Subtract the simulated sufficient statistics from the observed ones to calculate the approximate likelihood gradient. Sufficient statistics are calculated as, $SS1 = \mathrm{XX^t}$, and $\mathrm{SS2} = \mathrm{Y} \times \mathrm{X}$. Here, $\mathrm{X} = \left\{\vec{\sigma}^{(k)}\right\}_{\mathrm{k}=1,2,\ldots,\mathrm{N}}$ is the matrix of presence/absence status and $\mathrm{Y} = \left\{\epsilon^{(k)}\right\}_{\mathrm{k}=1,2,\ldots,\mathrm{N}}$ is the matrix of abiotic factors. Then, we obtain the difference of sufficient statistics as:

$$\Delta SS1 = \mathrm{SS1}^* - \mathrm{SS1},$$

and

$$\Delta SS2 = \mathrm{SS2}^* - \mathrm{SS2}.$$

Here, $\mathrm{SS1}^*$ and $\mathrm{SS2}^*$ is the sufficient statistics calculated for actual data.

5. Adjust the model parameters to climb the approximate gradient, using a schedule of step sizes as:

$$\mathrm{h_{new}} \leftarrow \mathrm{h_{old}} + \Delta\mathrm{h_{new}},$$
$$\mathrm{J_{new}} \leftarrow \mathrm{J_{old}} + \Delta\mathrm{J_{new}},$$
$$\mathrm{g_{new}} \leftarrow \mathrm{g_{old}} + \Delta\mathrm{g_{new}}.$$

Here,

$$\Delta\mathrm{h_{new}} = \mathrm{l}G_h + m\Delta\mathrm{h_{old}},$$
$$\Delta\mathrm{J_{new}} = \mathrm{l}\mathrm{G_J} + m\Delta\mathrm{J_{old}},$$
$$\Delta\mathrm{g_{new}} = \mathrm{l}\mathrm{G_g} + m\Delta\mathrm{g_{old}},$$

and,

$$G_h = \frac{\mathrm{diag}(\Delta\mathrm{SS1}) + \mathrm{p_h}}{\mathrm{M}},$$

$$G_J = \frac{\Delta SS1 + p_J}{M} |I(S) - 1|,$$

$$G_g = \frac{\Delta SS2 + p_g}{M},$$

are the approximated likelihood gradients.

6. Set h_{new}, J_{new}, g_{new}, Δh_{new}, ΔJ_{new} and Δg_{new} as h_{old}, J_{old}, g_{old}, Δh_{old}, ΔJ_{old} and Δg_{old}, respectively. If $t < T$, increment t by 1 and back to 2, else terminate the loop.

Here, the simulations in Step 1 use one step heat-bath algorithm (Gibbs sampling) to generate a community state distribution based on the model's current parameter estimates. While the subsequent community state distributions produced by Gibbs sampling are autocorrelated, this does not prevent convergence to the maximum likelihood estimates (Younes 1999, Salakhutdinov and Hinton 2012). The approximate likelihood gradients in Step 5 match those of gradient descent, except that they are averaged over a set of Monte Carlo samples rather than over all possible community states. These gradients were augmented with a momentum term (Hinton 2012) and by regularizers based on a logistic prior with location 0 and scale 2.0 (for environmental responses) or 0.5 (for pairwise interactions).

References

1. Bickel, P., K. Doksum. (1977). Mathematical Statistics: Basic Ideas and Selected Topics. San Francisco: Holden-Day.
2. Azaele, S., R. Muneepeerakul, A. Rinaldo, I. Rodriguez-Iturbe. (2010). Inferring plant ecosystem organization from species occurrences. *Journal of theoretical biology*, 262:323–329.
3. Lee, J. D., Hastie, T. J. (2015). Learning the structure of mixed graphical models. *Journal of Computational and Graphical Statistics*, 24(1), 230-253.
4. Murphy, K. P. (2012). Machine Learning: A Probabilistic Perspective. The MIT Press.
5. Harris, D. J. (2015). Multi-Process Statistical Modeling of Species' Joint Distributions. University of California, Davis.
6. Watanabe, T., Hirose, S., Wada, H., Imai, Y., Machida, T., Shirouzu, I., Konishi, S., Miyashita, Y., Masuda, N. (2014a). Energy landscapes of resting-state brain networks. *Frontiers in Neuroinformatics*, 8, 12.
7. Watanabe, T., Masuda, N., Megumi, F., Kanai, R., Rees, G. (2014b). Energy landscape and dynamics of brain activity during human bistable perception. *Nature Communications*, 5, 4765.
8. Harris, D. J. (2016). Inferring species interactions from co - occurrence data with Markov networks. *Ecology*, 97(12), 3308-3314.
9. Robbins, H., S. Monro. 1951. A Stochastic Approximation Method. *The Annals of Mathematical Statistics* 22:400–407.
10. Salakhutdinov, R., and G. Hinton. (2012). An efficient learning procedure for deep Boltzmann machines. *Neural Computation* 24:1967–2006.

11. Younes, L. 1999. On the convergence of Markovian stochastic algorithms with rapidly de- creasing ergodicity rates. *Stochastics: An International Journal of Probability and Stochastic Processes* 65:177–228.
12. Hinton, G. E. 2012. A practical guide to training restricted boltzmann machines. Pages 599–619 in Neural Networks: Tricks of the Trade. Springer. Hong,