esa ECOSPHERE

CONCEPTS & THEORY

A unified concept of dominance applicable at both community and species scales

ZHANSHAN (SAM) MA[1,2,]† AND AARON M. ELLISON[3]

[1] *Computational Biology and Medical Ecology Lab, State Key Laboratory of Genetic Resources and Evolution,*
Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223 China
[2] *Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223 China*
[3] *Harvard University, Harvard Forest, 324 North Main Street, Petersham, Massachusetts 01366 USA*

Citation: Ma, Z., and A. M. Ellison. 2018. A unified concept of dominance applicable at both community and species scales. Ecosphere 9(11):e02477. 10.1002/ecs2.2477

Abstract. Dominance and evenness can be two sides of the same coin or opposite ends of a spectrum. Although evenness and diversity are community-level concepts, dominance can be applied at both community and species scales. Nevertheless, there is not a metric applicable at both these scales that are unified with a single mathematical framework in the existing literature. Here, by extending Lloyd's meaning crowding concept from the population to community scale, we propose a dominance concept and associated metrics that are applicable at both scales. Such metrics can act as proxies for diversity in diversity–stability and diversity–ecosystem service analyses or for population abundance in species interaction network analysis and population stability analysis, with advantages in cross-scale and unified analyses. Our concept of dominance includes three measures (metrics) that link communities and species. The metrics have the same mathematical form, but different interpretations at the community and species scales. Our community-level metric is a function of Simpson's (1949) diversity index, for which we present a rigorous mathematical proof. Our species-level metrics quantify the difference between community dominance and the dominance of a virtual community whose mean population size, per species, equals the population size of the focal species. We demonstrate the use of these metrics using data from a longitudinal study of the human vaginal microbiome and provide new insights relevant for microbiome stability and disease etiology at the community scale. The new metrics also can be used for species dominance network (SDN) analysis at the species scale. Since the species dominance is a function of both species (population) abundance and community dominance, it has an advantage of capturing both community-scale global and species-scale local information, which becomes even more evident when it is used to construct SDNs. These results demonstrate the significance and applications of our unified dominance concept and metrics.

Key words: community dominance; diversity–stability relationship; dominance–stability relationship; mean crowding; species dominance; species dominance networks.

Received 17 July 2018; accepted 24 July 2018; final version received 1 September 2018. Corresponding Editor: Debra P. C. Peters.

† **E-mail:** ma@vandals.uidaho.edu

INTRODUCTION

In community ecology, dominance and evenness (one form of diversity) often are considered as two sides of the same coin. Evenness refers to the similarity of abundances of co-occurring species, whereas dominance of one or a few species is indicated by unevenness in species abundances

(Magurran 2004). Both evenness and dominance as normally measured and used by ecologists are properties of the entire community (Magurran 2004, Hooper et al. 2005, Hillebrand et al. 2008). The relationship between community-level diversity and community stability or ecosystem functions and services is of central importance for both theoretical and applied ecology (Hooper et al. 2005, Thibaut and Connolly 2013, Wang and Loreau 2016), but has been approached largely for organisms we can easily see, monitor, and measure. Sequence-based metagenomics are revealing high levels of microbial diversity, but the stability of microbial communities remains poorly understood (Lozupone et al. 2012, Moya and Ferrer 2016, Oh et al. 2016). Similarly, ecosystem functions that microbes provide are just beginning to be defined, explored, and related to diversity (Sechi et al. 2017, Young 2017).

Communities dominated by a single species (i.e., those with low evenness) tend to have either lower stability or reduced productivity (Rodriguez et al. 2015). However, one could argue that dominance also can be a property of individual species (Crase et al. 2015), with dominant species in a community being assigned large values of a dominance index and subordinate species being assigned smaller (or negative) values. Alternatively, dominance may reflect the contribution of a dominant species to its community. Either way, we suggest that understanding dominance at both the community and species levels will provide new insights into how diversity affects stability and ecosystem functioning.

Here, we introduce and develop a new framework for quantifying dominance that includes three dominance metrics for describing communities of organisms and the species that make up the communities. Our metrics are based on Lloyd's index of mean crowding (Lloyd 1967), but we extend it to include both species and communities in a single framework. Our approach is applicable to both communities and populations of individual species, which is a property that, to the best of our knowledge, is not a characteristic of other dominance or diversity metrics used by ecologists. This approach is applicable to any ecological community, but we illustrate it through an exploration of the diversity–stability relationship in the human vaginal microbial community (HVMC). Our exploration also generates new insights into the etiology of bacterial vaginosis (BV).

Background and Motivation

Lloyd (1967) developed his index of mean crowding (m^*) for measuring aggregation in populations of animals. Although Lloyd (1967) was inspired by neighborhood- or distance-based approaches for investigating tree competition and spatial point patterns in forests, he recognized that it was infeasible to measure distances between individuals of animals that move. Instead, Lloyd defined m^* based on the mean population density per quadrat and its corresponding variance:

$$m^* = m + \frac{\sigma^2}{m} - 1 \tag{1}$$

where m is mean population density (abundance) across n spatial point samples and σ^2 is its corresponding variance.

This measure of mean crowding not only included an estimate of neighborhood information, but also overcame a critical issue associated with abundance data: its non-normality. Iwao (1968) extended Lloyd (1967) by discovering a linear relationship between m^* and mean population density (m) and applied the m^*–m model ($m^* = \alpha + \beta m$) to assess insect population aggregation and their patterns of spatial distribution. The work of Iwao and his colleagues (Iwao 1968, Kuno 1991) and Taylor's variance–mean (V–m) power relationship ($V = am^b$) identified by Taylor (1961, 1984, 1986) remain the two primary approaches for describing spatial patterns of many animal populations (Ma 2013, 2015, Cohen and Xu 2015).

Discussions about the relative merits of the m^*–m and V–m models have focused on two issues: which model has more general applicability to field data; and which model's parameters (α, β or a, b) have better statistical properties and more reasonable biological interpretations (Iwao 1968, Taylor 1984, 1986, Kuno 1991). There are two points of agreement in these discussions. First, both models are better than an index of aggregation (dispersion) (e.g., V/m) or fitting frequency distributions (e.g., negative binomial) to abundance data. Second, the validity of either model depends on a third parameter, m_0, the population aggregation critical density (PACD), and the two models are nearly equivalent in terms of the PACD (Ma 1989, 1991, 2015).

We also note that there are obvious analogies between measures of aggregation of individuals within a population and measures of diversity (unevenness) of a community, and between fitting frequency distributions to individual abundances within a population and the species abundance distribution used by community ecologists. Ma (2012*a*, 2015) extended Taylor's *V–m* power law model to the community level. Zhang et al. (2014) and Oh et al. (2016) applied this extended power law to assess spatial heterogeneity in the human gut microbiome and the temporal stability of the human skin microbiome. In this paper, we describe how to extend *meaning crowding* to the community level for analyzing dominance. This extension has the additional advantage, unavailable with the *V–m* model, of being applicable to both community dominance *and* species dominance. Furthermore, it has a unified mathematical framework.

APPROACH AND ORGANIZATION OF THE PAPER

We extend the mean crowding concept from the scale of populations of individuals of a single species to encompass assemblages of species (communities) by defining three new dominance metrics: *community dominance* (D_c), *species dominance* (D_s), and *species dominance distance* (D_{sd}). (We note that in the community ecology literature, index, measures, and metrics have been used interchangeably. Here, we use metric in a general sense [as a "type"], as opposed to index, which is an "instance" of a metric.) With respect to a microbiome, we think that dominance is more meaningful than other measures of diversity for two reasons. First, D_c is a simple linear function of Simpson's diversity index (see mathematical proof in Appendix S1), but it is easier to examine dominance and its relationship to stability in microbial assemblages like the HVMC.

Second and more generally, the concept of dominance can be applied to both populations of individual species and multi-species communities. For example, we commonly refer to communities with high species diversity. However, it often matters greatly which species dominates an assemblage (Ellison et al. 2005, Ma et al. 2011, Valls et al. 2015), and its identity refers to an individual species or population. Except in population genetics, when we are discussing genetic diversity within a population (Romiguier et al. 2014), we rarely refer to high-diversity species. To the best of our knowledge, there is not an existing index of dominance or diversity that can be applied simultaneously to species and assemblages.

In the following sections, we first sketch how to extend Lloyd's (1967) index of mean crowding to both species and community scales (complete technical details and mathematical proofs are provided in Appendix S1). We then use a phenomenological modeling approach to dominance–stability relationships, focusing on five linear and non-linear diversity–stability models (linear, logistic, sine-logistic, linear–quadratic, and quadratic–quadratic). This approach reveals three fundamental components of the diversity–stability relationship: dominance-dependent stability (DDS), dominance-inversely-dependent stability (DID), and dominance-independent stability (DIS). Our modeling approach also distinguishes between stability and resilience; the latter is defined as the derivative of the former and characterizes the rate of changes in stability, analogously to the relationship between acceleration and velocity. Last, the models illustrate the (in)stability of the community equilibrium in HVMC.

THE DOMINANCE CONCEPT AND ITS METRICS

Community and species dominance based on mean crowding

Lloyd (1967) defined aggregation or dispersion of an animal population, m^*, as "the mean number per individual of other individuals in the same quadrat" (Lloyd 1967) or "the average number of other individuals per quadrat, per individual" (ISI 1986). As mentioned in *Background and Motivation*, m^* includes some neighborhood information—that is, density of individuals in a given area—but does so without measuring distances between individuals in the neighborhood. Such measurements generally are infeasible for free-living animals in a relatively continuous habitat (Lloyd 1967). The measurement of mean crowding also should be applicable to free-moving bacterial species in a largely continuous habitat such as the HVMC.

Another motivation for using mean crowding (or the *V–m* power law model) was to avoid estimating an arithmetic mean from the commonly encountered, highly skewed, long-tailed distribution of abundances of organisms that are either aggregated or patchily distributed (Ma 2015). Measures of aggregation, dispersion, patchiness, heterogeneity, skewness, evenness, and dominance often are used to characterize the abundance distribution of biological species; their temporal variability often is associated with the (in)stability of populations. Because aggregation essentially is inversely related to evenness, extending mean crowding to dominance (i.e., unevenness) should be straightforward.

Mean crowding of a community and community dominance.—We start by defining the *mean crowding of a community* using the same mathematical structure as that for mean crowding of a population (Eq. 1):

$$m_c^* = m_c + \frac{\sigma_c^2}{m_c} - 1 \tag{2}$$

where m_c = the mean abundance per species of all species in the community and σ_c^2 is its variance. Note that m_c is computed across species, not individuals of a single species, as in Eq. 1. By analogy with m^*, m_c^* is a measure of community unevenness or dominance, and we interpret m_c^* as the average abundance of other species per individual species. Essentially, we treat each species as a virtual quadrat (sensu Lloyd 1967); the quadrat is the sampling unit for m^* (Eq. 1), and the species is the sampling unit for m_c^* (Eq. 2).

Since there are many species in the community, it makes sense to divide m_c^* by the mean abundance of all the species in the community, m_c. This standardized value of m_c^* we call *community dominance*, which we abbreviate as D_c:

$$D_c = \frac{m_c^*}{m_c} = 1 + \frac{\sigma_c^2}{m_c^2} - \frac{1}{m_c} \tag{3}$$

and which is the direct counterpart of population-level patchiness or heterogeneity (aggregation) of population distribution (Lloyd 1967, Iwao 1968). We interpret D_c as unevenness in a community and express it in terms of the deviation from the average species. Analogously with aggregation of individuals in a population, D_c also can be interpreted as the "center of gravity" of a community, measuring how crowded the individuals of the average species are crowded by the individuals of its neighbor species.

We also observe that D_c is precisely linearly related to Simpson's (1949) diversity index D:

$$D_c = \frac{nD - n}{\sum_{i=1}^{n} m_{s_i}} \tag{4}$$

where n is the number of species in the community, and m_{s_i} is the mean abundance (size) of the ith species (see mathematical proof in Appendix S1, where we also examine the relationship between D_c and two other familiar measures of community dominance or diversity: Shannon–Weiner's H' and the Berger–Parker index of dominance, Berger and Parker 1970). We fit a simple linear model between D_c and the existing dominance indexes (Fig. 1; Appendix S1: Table S1).

The dominance of individual species.—We use D_c to define a dominance index for each species in the community. We define the *species dominance distance* of an individual species s as follows:

$$D_{sd} = \frac{m_c^*}{m_s} = \frac{m_c}{m_s} + \frac{\sigma_c^2}{m_c m_s} - \frac{1}{m_s} \tag{5}$$

which defines a dominance index that "distributes" the mean community crowdedness (m_c^*) over a specific species (focal species) (m_s) rather than over an average species (m_c). D_{sd} ranges from 0 to +∞. Although the behavior of D_{sd} seems counterintuitive—dominant species may have large values of m_s and hence small values of D_{sd}—its interpretation is intuitive if we imagine community as a sphere with a center of gravity = D_c. The dominant species (in terms of abundances) are arrayed closer to the center of this metaphorical sphere than are the rare ("satellite") species, which should be farther from the center and have larger value of the species dominance distance.

We define *species dominance* (range = −∞ to D_c) as the difference between community dominance (D_c) and species dominance distance (D_{sd}):

$$\begin{aligned} D_s = D_c - D_{sd} &= \frac{m_c^*}{m_c} - \frac{m_c^*}{m_s} \\ &= 1 - \frac{m_c}{m_s} + \frac{\sigma_c^2}{m_c^2} - \frac{\sigma_c^2}{m_c m_s} + \frac{m_c - m_s}{m_c m_s}. \end{aligned} \tag{6}$$

Dominant species have larger values of D_s than other species, which accords with our intuition. Two considerations led us to this definition of

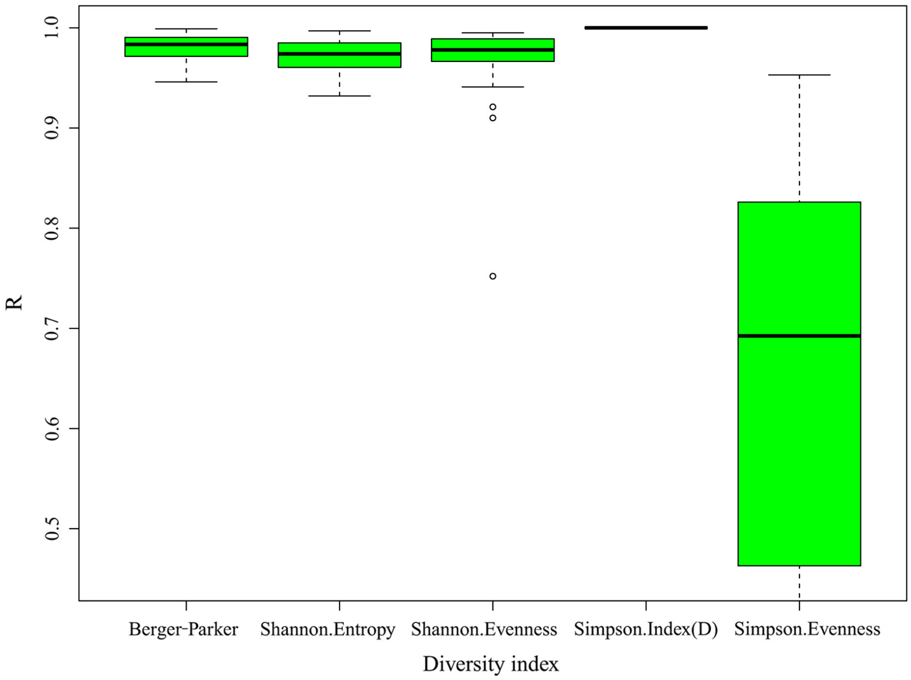

Fig. 1. Associations between our community dominance metric (D_c) and other existing diversity indexes. See Appendix S1 for fitted parameters and additional discussion.

species dominance. First, we used the difference between community dominance and species dominance distance rather than the inverse of species dominance distance, to avoid a non-linear transformation. Second, D_s turns out to be more suitable than D_{sd} for analyzing dominance by individual species with an approach to analyzing species dominance networks (SDNs) (Z. Ma and A. M. Ellison, *unpublished manuscript*).

Phenomenological modeling of community dominance (diversity)-stability relationships

We start by defining *community dominance stability* (referred to hereafter as *community stability*), $S_c(t)$, as follows:

$$S_c(t) = \frac{D_c(t+1) - D_c(t)}{D_c(t)}. \tag{7}$$

$S_c(t)$ measures the rate of change of community dominance (for many other definitions of ecological stability, see Grimm and Wissel 1997).

Similarly, we define *population dominance stability* (hereafter, *population stability*), $S_s(t)$, as the rate of change of species dominance over time,

$$S_s(t) = \frac{D_s(t+1) - D_s(t)}{D_s(t)} \tag{8}$$

noting that there would be a separate measure of stability for each species in the community. Finally, we assume that community dynamics can be modeled by a general (set of) differential equation(s) such as:

$$S(t) = \frac{dD(t)}{D(t)dt} = f[D(t), Z] \tag{9}$$

where $D(t)$ is one of the dominance metrics at time t, $S(t)$ is the parallel stability metric at time t, and Z is an optional vector of covariates.

Since we do not know the categorical form of function f in Eq. 8, we adopt a data-driven, phenomenological modeling approach. Through trial-and-error and exploratory curve-fittings, we found that five alternative models can describe basic community dominance dynamics:

1. a two-parameter linear model

$$S_c(t) = a + bD_c(t);$$

2. a five-parameter linear–quadratic (L–Q) model

$$S_c(t) = a + bD_c(t) + cD_c^2(t) + [D_c(t) - d] \times \mathrm{Sign}(D_c(t) - d)[c(D_c(t) + d) + e)];$$

3. a six-parameter quadratic–quadratic (Q–Q) model

$$S_c = a + bD_c + cD_c^2 + (D_c - d) \times \mathrm{Sign}(D_c - d)[e(D_c + d) + f)];$$

4. a three-parameter logistic

$$S_c(t) = \frac{K}{1 + a^* \exp[-rD_c(t)]};$$

5. and a three-parameter periodic logistic-sine model

$$S_c(t) = \frac{K}{1 + a^* \exp[-rD_c(t)]} \sin\left(\frac{D_c(t)}{\pi}\right).$$

A detailed discussion of these five models and the derivations of their parameters are given in Appendix S1.

For each of the stability models listed above, there is a corresponding dominance dynamics model. For community dominance, this would be:

$$D_c(t+1) = D_c(t)[1 + S_c(t)] \tag{10}$$

and so, for example, the corresponding model for dominance dynamics with the three-parameter logistic stability model would be as follows:

$$D_c(t+1) = D_c(t)\left\{1 + \frac{K}{[1 + a\exp(-rD_c(t))]}\right\}. \tag{11}$$

Like density-dependence models for population regulation (Kot 2001, Berryman et al. 2002, Pastor 2008), the generalized stability model (Eqs. 7–9) may display three types of local behavior: (1) DDS, in which stability increases with dominance; (2) dominance-inversely-dependent stability (DID), in which stability decreases with dominance; and (3) DIS, in which stability does not change with dominance level. That is, $S(t) \propto kD(t)$, with $k < 0$, $k > 0$, $k = 0$, corresponding, respectively, to DDS, DID, and DIS.

In practice, except for the simple two-parameter linear model (where the parameter b is equivalent to k in the generalized stability model), we may not be able to determine precisely the value of k. However, the non-linear models we considered are simple enough that we can evaluate piecewise relationships between dominance and stability. We also note that the three-parameter logistic and two-parameter linear model may capture only one of the three dominance-dependence behaviors in a specific model, whereas the other three models are more flexible and may capture all three behaviors in a single model. The three-parameter logistic-sine model also can capture periodic fluctuation of the three types of dependence relationships.

Since our dominance metrics also are related to common measures of diversity, this modeling approach offers an equally powerful method for modeling classic diversity–stability relationships. For the simple linear model, the slope b is the derivative of the linear stability function, that is, the rate of change of stability with respect to dominance. It is a measure of resilience: the speed at which a community returns to local equilibrium after perturbation.

DEMONSTRATION OF THE METRICS

The HVMC dataset

We compare our dominance metrics to three standard diversity indices (Appendix S1) and illustrate the phenomenological model selection approach to dominance–stability modeling using a "32-healthy cohort dataset." This dataset is from a longitudinal study of 32 healthy women done in 2006–2007 at the Johns Hopkins University School of Medicine, Baltimore, Maryland, USA (Gajer et al. 2012). The participants self-collected mid-vaginal swabs and vaginal smears twice weekly for 16 weeks. The extraction of genomic DNA from frozen vaginal swabs, PCR amplification, sequencing of the V1–V2 region of bacterial 16S rRNA genes, and the archive of sequence data are described by Gajer et al. (2012). QIIME (Caporaso et al. 2010), UCLUST (Edgar 2010), UCHIME (Edgar et al. 2011), RDP Naïve Bayesian Classifier (Wang et al. 2007), and speciateIT (https://speciateIT.sourceforge.net) were used to obtain the OTU table and included sequence read counts and relative abundances of the taxonomic assignments at the bacterial species level of 97% similarity (Gajer et al. 2012).

Comparisons between our dominance metrics and existing diversity indices

Fig. 1 compares D_c with three other diversity indexes. As D_c is a linear transformation of Simpson's D, those two are perfectly correlated ($r = 1$). Correlations between D_c and both the Shannon–Weiner and the Berger–Parker diversity indices exceeded 0.95 among all but two of the HVMC samples (r for those two outliers equaled 0.94 and 0.75) and were statistically significant in all cases ($P < 0.001$). The overall correlation between D_c and Simpson's measure of evenness (D/S) averaged 0.61 (median $r = 0.69$; Fig. 1) and again was statistically significant ($P < 0.001$). We conclude that D_c is comparable quantitatively to existing measures of diversity.

Why introduce another diversity (dominance) metric? We assert that the utility of D_c is its straightforward extension to, and interpretation at, the species level, with the same mathematical form (but with different interpretations). This extension (D_s) allows us to identify quantitatively which species dominates a community and to what extent it dominates the community. An example is shown in Tables 1, 2 for seven microbial species from one of the subjects in the 32-healthy cohort HVMC dataset (the results for all species of this subject [number 400] are given in the Appendix S1 and computed with code in Data S1).

Table 1. The *species dominance distances* (D_{sd}) of seven species selected from the 29 longitudinal samples of Subject number (#) 400.

Sample ID	OTU#1	OTU#8	OTU#28	OTU#11	OTU#115	OTU#57	OTU#2
400_0101	0.774	13.259	2.175	1325.9	∞	∞	∞
400_0105	0.884	858.50	9.090	∞	∞	∞	∞
400_0108	0.953	159.27	∞	2269.6	∞	∞	2269.6
400_0112	0.941	304.78	∞	∞	∞	∞	∞
400_0119	0.931	35.271	∞	∞	∞	∞	2160.4
400_0122	0.989	444.94	∞	∞	∞	∞	∞
400_0126	0.906	909.75	61.102	∞	∞	∞	∞
400_0129	42.40	186.56	4.056	∞	3.97	∞	∞
400_0202	0.945	132.94	∞	∞	∞	∞	2226.8
400_0205	0.973	86.030	2365.8	∞	∞	∞	∞
400_0209	0.825	1.329	∞	∞	∞	∞	1227.9
400_0212	0.803	1.655	1303.2	∞	∞	∞	∞
400_0216	0.730	2.816	292.36	46.889	∞	236.7	552.24
400_0219	0.884	6.310	∞	∞	∞	∞	1918.2
400_0226	0.915	20.559	∞	∞	∞	∞	∞
400_0302	0.949	22.730	∞	3000.4	∞	∞	∞
400_0305	0.970	48.249	∞	∞	∞	∞	∞
400_0309	0.874	5.644	∞	1493.5	∞	∞	1866.8
400_0312	0.920	27.218	∞	1054.7	∞	∞	∞
400_0316	0.910	18.772	∞	1030.1	∞	∞	∞
400_0319	0.904	20.039	∞	∞	∞	∞	∞
400_0323	1.961	23.229	84.107	6.222	∞	84.11	∞
400_0326	2486.7	2486.7	0.997	1243.4	∞	∞	∞
400_0330	0.714	5.830	17.997	59.846	∞	∞	∞
400_0402	0.916	37.855	149.39	597.57	∞	∞	∞
400_0406	0.944	78.693	∞	∞	∞	∞	2223.1
400_0409	0.922	36.539	∞	∞	∞	∞	∞
400_0413	0.931	39.450	∞	∞	∞	∞	∞
400_0416	0.931	31.615	∞	2165.6	∞	∞	∞

Notes: Top three most abundant species are as follows: OTU#1 = *L. iners*, OTU#8 = *L. jensenii*, and OTU#28 = *Staphylococcus*. Two moderate abundant species are as follows: OTU#11 = *Anaerococcus* and OTU#115 = *Pseudomonas*. Two least abundant species (but excluded species with total reads <10) are as follows: OTU#57 = *Facklamia* and OTU#2 = *L. crispatus*. Theoretically, $D_{sd} \in (0, +\infty)$. When population abundance = 0, $D_{sd} = \infty$. Obviously, dominance rank and abundance rank can be very different because the most abundant species are not necessarily the most dominant species, and least abundant species are not necessarily least dominant.

Table 2. The *community dominance metric* (D_c) and *species dominance metric* (D_s) of seven species selected from the 29 longitudinal samples of Subject number (#) 400.

		D_s						
Sample ID	Community dominance D_c	OTU#1	OTU#8	OTU#28	OTU#11	OTU#115	OTU#57	OTU#2
400_0101	31.82	31.050	18.566	29.649	−1294.1	−3688.8	−3688.8	−3688.8
400_0105	46.36	45.471	−812.14	37.265	−3688.8	−3688.8	−3688.8	−3688.8
400_0108	54.46	53.507	−104.81	−3688.8	−2215.2	−3688.8	−3688.8	−2215.2
400_0112	53.03	52.090	−251.75	−3688.8	−3688.8	−3688.8	−3688.8	−3688.8
400_0119	51.85	50.918	16.577	−3688.8	−3688.8	−3688.8	−3688.8	−2108.5
400_0122	58.74	57.749	−386.21	−3688.8	−3688.8	−3688.8	−3688.8	−3688.8
400_0126	49.13	48.220	−860.62	−11.976	−3688.8	−3688.8	−3688.8	−3688.8
400_0129	11.19	−31.206	−175.37	7.138	−3688.8	7.224	−3688.8	−3688.8
400_0202	53.45	52.504	−79.495	−3688.8	−3688.8	−3688.8	−3688.8	−2173.4
400_0205	56.79	55.818	−29.239	−2309.0	−3688.8	−3688.8	−3688.8	−3688.8
400_0209	29.48	28.651	28.147	−3688.8	−3688.8	−3688.8	−3688.8	−1198.5
400_0212	31.27	30.470	29.618	−1271.9	−3688.8	−3688.8	−3688.8	−3688.8
400_0216	29.83	29.097	27.011	−262.54	−17.061	−3688.8	−206.849	−522.42
400_0219	46.04	45.156	39.731	−3688.8	−3688.8	−3688.8	−3688.8	−1872.1
400_0226	50.08	49.166	29.523	−3688.8	−3688.8	−3688.8	−3688.8	−3688.8
400_0302	54.01	53.063	31.282	−3688.8	−2946.3	−3688.8	−3688.8	−3688.8
400_0305	56.45	55.476	8.197	−3688.8	−3688.8	−3688.8	−3688.8	−3688.8
400_0309	44.80	43.929	39.159	−3688.8	−1448.6	−3688.8	−3688.8	−1822.0
400_0312	50.63	49.710	23.412	−3688.8	−1004.1	−3688.8	−3688.8	−3688.8
400_0316	49.45	48.540	30.678	−3688.8	−980.64	−3688.8	−3688.8	−3688.8
400_0319	48.82	47.916	28.781	−3688.8	−3688.8	−3688.8	−3688.8	−3688.8
400_0323	14.63	12.668	−8.601	−69.478	8.407	−3688.8	−69.478	−3688.8
400_0326	59.69	−2427.0	−2427.0	58.690	−1183.7	−3688.8	−3688.8	−3688.8
400_0330	29.79	29.081	23.964	11.797	−30.052	−3688.8	−3688.8	−3688.8
400_0402	50.18	49.265	12.326	−99.211	−547.39	−3688.8	−3688.8	−3688.8
400_0406	53.35	52.410	−25.339	−3688.8	−3688.8	−3688.8	−3688.8	−2169.7
400_0409	50.86	49.935	14.318	−3688.8	−3688.8	−3688.8	−3688.8	−3688.8
400_0413	51.84	50.912	12.392	−3688.8	−3688.8	−3688.8	−3688.8	−3688.8
400_0416	51.97	51.033	20.350	−3688.8	−2113.7	−3688.8	−3688.8	−3688.8

Notes: Top three most abundant species are as follows: OTU#1 = *L. iners*, OTU#8 = *L. jensenii*, and OTU#28 = *Staphylococcus*. Two moderate abundant species are as follows: OTU#11 = *Anaerococcus* and OTU#115 = *Pseudomonas*. Two least abundant species (but excluded species with total reads <10) are as follows: OTU#57 = *Facklamia* and OTU#2 = *L. crispatus*. Theoretically, $D_s \in (-\infty, +\infty)$. When population abundance = 0, $D_s = -\infty$. In practice, for each subject, we replace the ($-\infty$) with the *species dominance* value of the least dominant member in all time-series samples of the subject. Obviously, dominance rank and abundance rank can be very different because the most abundant species are not necessarily the most dominant species, and least abundant species are not necessarily least dominant.

Dominance metrics for very rare and very common species

Because $D_{sd} = +\infty$ and $D_s = -\infty$ when the abundance of the focal species = 0, we replace $D_s = -\infty$ in a particular sample with its lowest value in all the time-series samples of that subject. We note that at the most extreme value (abundance = 0, $D_s = -\infty$) that the distance $D_{sd} \to +\infty$ and the relevant species becomes "disconnected" from the community; it is locally (or temporarily) extinct. Of course, it is possible to artificially convert the value of each metric into a small range such as [0, 1], but we do not see any need or benefit to doing so. Rather, we view the capability to represent discontinuous points or local (temporary) extinctions as an advantage because temporary disappearance or local extinction of bacterial species frequently is observed in the HVMC (Gajer et al. 2012). We note that the most abundant species are not necessarily the most dominant species, and the least abundant species are not necessarily the least dominant, and vice versa, so there is no requirement that abundance and dominance rankings be perfectly aligned.

We also observe that values of D_s of rare species may have large fluctuations. Time series of D_s or D_{sd} illustrate this phenomenon (Fig. 2),

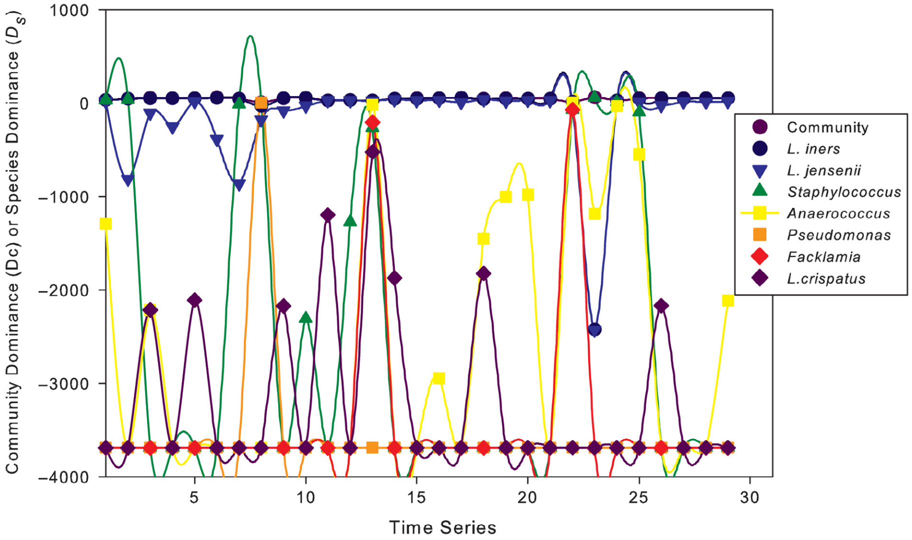

Fig. 2. Community dominance (D_c) and species dominance (D_s) metrics of seven representative species selected from the time-series data of Subject number 400, including the three most abundant species (*Lactobacillus iners*, *Lactobacillus jensenii*, and *Staphylococcus*), two moderately abundant species (*Anaerococcus* and *Pseudomonas*), and the two least abundant (rare) species (*Facklamia* and *Lactobacillus crispatus*), raw data in Table 1. One phenomenon this graph reveals is that (rare species) may have even larger fluctuations including local or temporal extinctions, which may help identify potential associations between rare microbes (such as opportunistic pathogens) and pathological changes (such the occurrence of bacterial vaginosis).

and we think that such illustrations (e.g., using exploratory data analysis) could help identify potential associations between rare microbes and pathological changes such as the occurrence of BV (e.g., opportunistic pathogens should be rare, at least initially). Indeed, this was one of the major motivations for our development of dominance metrics that could be used at both community and species levels. Fig. 2 shows the community dominance as well as the species dominance of seven selected species including the three most abundant species, two moderately abundant species, and two rare species.

Fig. 3, which illustrates community and species dominance of the three most abundant species using polar coordinates, reveals that community dominance seems to be “controlled” by one of the top three species (two circles representing the community dominance and the “master” species overlapped), but the “control” is dynamic and may be transferred from one species to another with time progression (change of angular coordinate degree). This possible control mechanism is more obvious in species dominance network analysis constructed with our species dominance metric (Z. Ma and A. M. Ellison, *unpublished manuscript*).

It could be argued that other diversity or evenness indices could be extended similarly, but we suspect that such extensions would be difficult. For example, the Berger–Parker (1970) diversity index yields the same value if the most abundant species in two communities are equally abundant in each community, regardless of whether they are the same species. However, two communities may be dominated by different species with equal abundances.

Phenomenological modeling of community dominance–stability relationships

We illustrate qualitative patterns of stability and its relationship with measures of dominance as modeled with the linear, linear–quadratic (L–Q), quadratic–quadratic (Q–Q), logistic, and logistic-sine models. For each model, we examined how well it fit the 32-healthy cohort HVMC data and whether it revealed biologically interpretable patterns. As with most statistical models, our approach is a compromise between realism and simplicity and includes biological

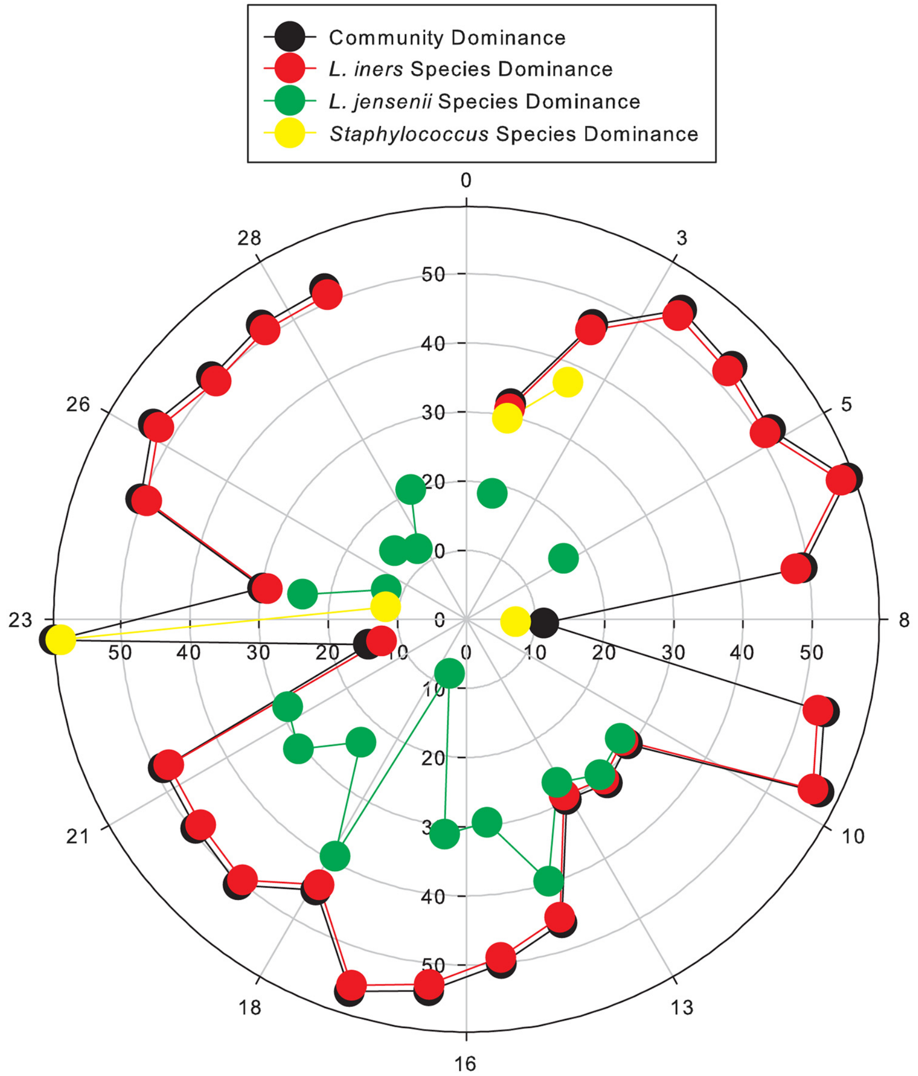

Fig. 3. A polar coordinate graph showing community dominance and the species dominance metrics of the three most abundant species in the vaginal microbial community of Subject number 400: The graphs reveal that the community dominance seems to be “controlled” by one of the top three species (two circles representing the community dominance and the “master” species overlapped), but this “control” is dynamic and may be transferred from one species to another with time progression (change of angular coordinate degree).

interpretations of model parameters, statistical tests (coefficient of determination r^2 and standard errors of parameters), and an appeal to parsimony (details of model selection are discussed further in Appendix S1).

The most appropriate model for each subject is summarized in Tables 3–7. Both the logistic model and the linear model can capture only one kind of diversity–stability mechanisms with a single model, depending on the sign of the

Table 3. A summary of the single best-fitted model (logistic model) for individuals in the 32-healthy cohort, summarized from Appendix S1: Table S2.

Subject ID	K	r	a	R^2	n
#400	4.741	−0.206	0.026	0.91	28
#412	4.168	−0.647	0.00002	0.99	27
#412	4.168	−0.647	0.00002	0.99	27

Note: Dominance-dependent stability with an asymptotic equilibrium line when $D_c \to \infty$.

parameter b (linear model) or r (logistic model). In contrast, each of the other three models can capture all three diversity–stability mechanisms, that is, the same community may exhibit alternately the three diversity–stability mechanisms (DDS, DID, and DIS). Detailed information about the model selection criteria and process is exposed in Appendix S1: Tables S2–S6.

Table 4. A summary of the single best-fitted model (logistic-sine model) for individuals in the 32-healthy cohort, summarized from Appendix S1: Table S3.

Subject ID	K	r	a	R^2	n
#420	−0.294	0.048	−1.677	0.71	27
#431	−0.142	0.060	−1.822	0.36	26

Note: Dominance-dependent stability and dominance-independent stability alternate periodically.

Table 5. A summary of the single best-fitted model (linear model) for individuals in the 32-healthy cohort, summarized from Appendix S1: Table S4.

Subject ID	a	b	R	n
#401	0.594	−0.012	0.53	29
#403	1.361	−0.047	0.80	31
#404	0.488	−0.010	0.37	29
#405	1.551	−0.033	0.85	30
#406	0.800	−0.014	0.52	30
#407	1.783	−0.021	0.56	27
#410	0.642	−0.012	0.49	28
#413	0.789	−0.028	0.45	27
#414	0.941	−0.056	0.57	31
#424	1.260	−0.016	0.68	27
#430	0.539	−0.037	0.45	26
#432	1.269	−0.017	0.63	28
#443	0.718	−0.008	0.43	27
#444	0.486	−0.014	0.39	26

Note: Globally dominance-dependent stability, but the mechanism may be complex locally.

The parameters (r) of the two logistic models and the slopes (b) of the linear models were negative for all subjects; hence, there is no need to note the sign of the model parameters in Table 5. For the HVMC data, the dominance–stability relationships modeled by the logistic and linear models were dominance-dependent, that is, the higher the dominance, the more stable the community. Equivalently, when diversity was high (i.e., dominance was low), stability declined. However, we emphasize that not all the HVMCs exhibited DDS, as suggested by the other three models (i.e., logistic-sine, L–Q, and Q–Q models).

Additional examination of the slopes (b) of the linear models (Appendix S1: Table S4) suggested further nuances in the relationship between dominance (diversity) and stability in the HVMC. The slope b is the derivative of the linear stability function, that is, the rate of change of stability with respect to dominance. This slope is a measure of resilience: the speed at which a community returns to local equilibrium after perturbation.

Simulations

We used mobsim—an R package for the simulation and measurement of biodiversity across spatial scales, developed by May et al. (2017)—to simulate three commonly used species abundance distributions, that is, the lognormal distribution, log-series distribution, and power law

Table 6. A summary of the single best-fitted model (linear–quadratic [L–Q] model) for individuals in the 32-healthy cohort, summarized from Appendix S1: Table S5.

Subject ID	a	b	c	d	e	R^2	$b_1 = b - e$	$c_2 = 2c$
#402	0.331	0.014	0.00033	28.341	−0.108	0.68	0.122	0.0007
#416	2.210	−0.059	0.00099	23.740	−0.149	0.74	0.09	0.00198
#429	−28.13	0.728	0.00058	38.752	−0.888	0.67	1.616	0.0011
#408	49.023	−5.947	0.00067	8.187	5.862	0.85	−11.8	0.0013
#445	1.608	−0.091	0.00032	19.639	0.016	0.66	−0.107	0.00064
#411	2.001	−0.067	−0.00027	27.018	0.106	0.68	−0.173	−0.0005
#423	−0.371	0.020	−0.00029	61.448	0.071	0.77	−0.051	−0.00058
#435	1.714	−0.054	−0.00042	28.569	0.119	0.73	−0.173	−0.00084
#436	−5.741	0.171	−0.00142	51.078	0.223	0.47	−0.052	−0.0028
#437	−4.670	0.127	−0.00094	66.958	0.180	0.54	−0.053	−0.00188

Notes: DDS, dominance-dependent stability; DIS, dominance-independent stability. When $b_1 > 0$, $c_2 > 0$, DIS followed by DDS, a possible stable equilibrium and DIS (#402, #416, #429). When $b_1 < 0$, $c_2 > 0$, DDS followed by a possible equilibrium and DIS (#408, #445). When $b_1 < 0$, $c_2 < 0$, DDS followed by possibly two equilibriums and DDS (#411, #423, #435, #436, #437).

Table 7. A summary of the single best-fitted model (quadratic–quadratic model) for individuals in the 32-healthy cohort, summarized from Appendix S1: Table S6.

Subject ID	a	b	c	d	e	f	R^2	$c_1 = c - e$	$c_2 = c + e$
#415	−3.361	0.269	−0.0007	18.353	0.0017	−0.385	0.73	−0.0024	0.001
#418	−1.945	0.176	−0.0032	43.002	0.0032	−0.168	0.86	−0.0064	−0.00001
#446	−12.71	0.572	−0.0066	43.061	−0.0027	0.333	0.59	−0.0039	−0.0093

Notes: DDS, dominance-dependent stability; DIS, dominance-independent stability. When $c_1 < 0$, $c_2 > 0$, DDS and DIS alternate, two parabolas connected at $D_c = d \approx 18$, with a possible stable equilibrium (#415). When $c_1 < 0$, $c_2 < 0$, DDS and DIS alternate, two parabolas connected at $D_c = d \approx 43$, stability of equilibriums is uncertain (#418, #446).

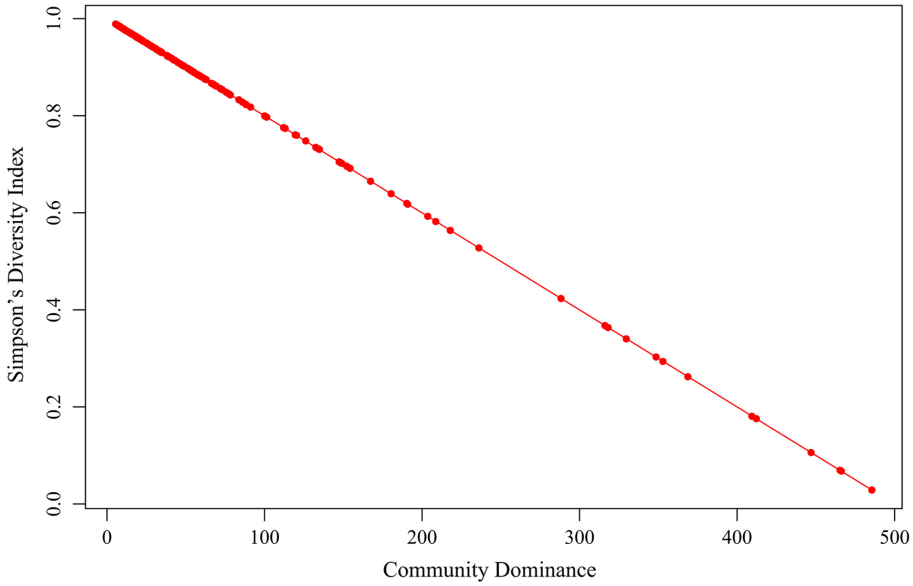

Fig. 4. The linear relationship (Eq. 4) between community dominance (x-axis) and Simpson's diversity index (y-axis), plotted with the simulated data from power law distribution with $K = 2$.

distribution. Overall, the results from the simulations supported our use of the dominance concept and findings discussed in previous sections. For example, Fig. 4 shows the linear relationship (Eq. 4) between community dominance and Simpson's diversity index, confirmed by the simulated data. More detailed simulation results are discussed in Appendix S1.

Discussion

In late 1990s, clinical microbiologists (Sobel 1999) using culture-dependent technology applied ecological interpretations—notably community diversity, species dominance, and the diversity–stability relationship—to interpret BV etiology. With the advent of metagenomic sequencing technology, a much larger number of uncultivable microbial OTUs have been detected in the HVMC and associated with BV (Ravel and Gordon 2011, Gajer et al. 2012). Whether considering culturable or unculturable microbes in the HVMC, many more questions are raised than answers are available. The state-of-the-art remains: "BV remains a riddle, wrapped in a mystery, and inside an enigma" (Fredricks 2011); and "BV is not a single entity, but a syndrome linked to various community types that cause somewhat similar physiological symptoms. This suggests that a yet unknown common community function may account for BV and the differing responses to antibiotic therapies" (Ma et al. 2012).

Until recently, the prevailing opinion was that more diverse (higher diversity) HVMCs are less stable and prone to BV. Although Ma et al. (2012) rightly rejected this opinion by citing counterexamples, they did not present a mechanistic explanation for why the diversity–instability relationship should fail to result in BV. On first glance, our identification of DDS mechanism as displayed by the standard logistic and the linear model (Appendix S1: Tables S2, S4) appears to support the prevalent opinion about diversity–instability leading to BV because dominance-dependence mechanism predicts that lower dominance (higher diversity) corresponds to lower stability. However, more careful consideration of the slope (b) of the linear models provides a counterargument. The slope b is the derivative of the linear stability function, that is, the rate of change of stability with respect to dominance. A community that has a steeper slope (b) should be easier to stabilize with the same units of dominance increase than a community with a less steep slope.

The range of differences in slopes (b) among communities in Appendix S1: Table S4 exceeded 15-fold (smallest $b = -0.123$ for Subject number 412, and largest $b = -0.008$ for Subject number 443). This suggests that a diverse community—the community that usually lacks apparent dominant species—is not necessarily inherently unstable because it can be quicker in stabilizing itself than a counterpart that is with highly dominant species.

The above apparent contradiction can be resolved by a careful distinction between the stability and resilience of a community. There are numerous definitions of stability (Grimm and Wissel 1997), and resilience is often treated simply as one component or dimension of stability (Ma 2012*b*). We defined community stability as $S_c(t)$ (Eq. 7), but in the case of linear model, the slope (b) represents resilience: the speed at which a community returns to local equilibrium after perturbation. It should be noted that in our definition of community stability, time (t) is implicitly included in the stability function (Eqs. 7–9), so the slope (b) of the linear model can be interpreted as a measure of resilience.

By distinguishing community stability from resilience in the context of the linear stability function, we can draw the following insights from Appendix S1: Table S4: The DDS mechanism suggests that high-diversity (low dominance) community can be less stable than low-diversity (high dominance) community, but the former, if its slope is steeper than the slope of the latter, may have higher resilience than the latter. This distinction resolves the apparent paradox regarding the diversity–stability relationship in the case of HVMC and presents a more comprehensive, cohesive, and quantitative argument to support Ma et al.'s (2012) rejection of the dominance–instability mechanism for BV.

In summary, when discussing community stability, it is critical to distinguish between stability and resilience. Both are needed to describe accurately the diversity–stability relationship in HVMC. A high-diversity community may have lower stability in terms of the magnitude of community dominance change, as suggested by the prevalent opinion on the stability of HVMC, but the community may still be resilient. In other words, a high-diversity community may be able to stabilize itself more efficiently or effectively, as is suggested by more general ecological paradigms of diversity and stability. For example, high-diversity communities may have higher connectivity among its species that results in more efficient stabilization following perturbations.

We note in closing that our dominance–stability analysis is imperfect. First, we chose what are arguably the simplest definitions of stability and resilience to avoid opening the Pandora's box of stability definitions (Grimm and Wissel 1997, Green et al. 2006, Ma 2012*b*). Second, our modeling strategy followed the principle of parsimony and was essentially a compromise between realism and simplicity. Third, our modeling approach was phenomenological and lacks rigorous theoretical assumptions (May 1973, 1975, Allesina and Tang 2012). Overcoming these limitations is beyond the scope of the present article, but we begin to address them in a follow-up paper in which we apply our species-level dominance metric to construct a species dominance network (SDN) of the HVMC (Z. Ma and A. M. Ellison, *unpublished manuscript*).

ACKNOWLEDGMENTS

We thank Prof. Larry Forney, University of Idaho, and Prof. Jacques Ravel, University of Maryland, for discussion of BV at various occasions. We are also indebted to D. D. Ye and L. W. Li, the Chinese Academy of Sciences, for verifying the computational results. The authors declare no conflict of interests.

LITERATURE CITED

Allesina, S., and S. Tang. 2012. Stability criteria for complex ecosystems. Nature 483:205–208.

Berger, W. H., and F. L. Parker. 1970. Diversity of planktonic foraminifera in deep-sea sediments. Science 168:1345–1347.

Berryman, A. A., M. Lima Arce, and B. A. Hawkins. 2002. Population regulation, emergent properties, and a requiem for density dependence. Oikos 99:600–606.

Caporaso, J. G., et al. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7:335–336.

Cohen, J. E., and M. Xu. 2015. Random sampling of skewed distributions implies Taylor's power law of fluctuation scaling. Proceedings of the National Academy of Sciences USA 112:7749–7754.

Crase, B., P. A. Vesk, A. Liedloff, and B. A. Wintle. 2015. Modelling both dominance and species distribution provides a more complete picture of changes to mangrove ecosystems under climate change. Global Change Biology 21:3005–3020.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460.

Edgar, R. C., B. J. Haas, J. C. Clemente, C. Quince, and R. Knight. 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194.

Ellison, A. M., M. S. Bank, B. D. Clinton, E. A. Colburn, K. Elliott, C. R. Ford, D. R. Foster, B. D. Kloeppel, J. D. Knoepp, and G. M. Lovett. 2005. Loss of foundation species: consequences for the structure and dynamics of forested ecosystems. Frontiers in Ecology & the Environment 3:479–486.

Fredricks, D. N. 2011. Molecular methods to describe the spectrum and dynamics of the vaginal microbiota. Anaerobe 17:191–195.

Gajer, P., R. M. Brotman, G. Bai, J. Sakamoto, U. M. Schütte, X. Zhong, S. S. Koenig, L. Fu, Z. S. Ma, and X. Zhou. 2012. Temporal dynamics of the human vaginal microbiota. Science Translational Medicine 132:132ra52.

Green, D. G., N. Klomp, G. Rimmington, and S. Sadedin. 2006. Complexity in landscape ecology. Springer, Dordrecht, The Netherlands.

Grimm, V., and C. Wissel. 1997. Babel, or the ecological stability discussions: an inventory and analysis of terminology and a guide for avoiding confusion. Oecologia 109:323–334.

Hillebrand, H., D. M. Bennett, and M. W. Cadotte. 2008. Consequences of dominance: a review of evenness effects on local and regional ecosystem processes. Ecology 89:1510–1520.

Hooper, D. U., F. S. Chapin III, J. J. Ewel, A. Hector, P. Inchausti, S. Lavorel, J. H. Lawton, D. M. Lodge, M. Loreau, and S. Naeem. 2005. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecological Monographs 75:3–35.

ISI [Institute for Scientific Information]. 1986. This week's citation classic. Current Contents 4:14.

Iwao, S. 1968. A new regression method for analyzing the aggregation pattern of animal populations. Researches on Population Ecology 10:1–20.

Kot, M. 2001. Elements of mathematical ecology. Cambridge University Press, Cambridge, UK.

Kuno, E. 1991. Sampling and analysis of insect populations. Annual Review of Entomology 36:285–304.

Lloyd, M. 1967. Mean crowding. Journal of Animal Ecology 36:1–30.

Lozupone, C. A., J. I. Stombaugh, J. I. Gordon, J. K. Jansson, and R. Knight. 2012. Diversity, stability and resilience of the human gut microbiota. Nature 489:220–230.

Ma, Z. S. 1989. M^*-M power law and its application to the sampling of insect populations. Journal of Beijing Forestry University 11:92–95 [In Chinese with English Abstract.].

Ma, Z. S. 1991. Further interpreted Taylor's power law and population aggregation critical density. Transactions of the Ecological Society of China 1:284–288 [In Chinese with English Abstract.].

Ma, Z. S. 2012*a*. A Note on extending Taylor's power law for characterizing human microbial communities. http://adsabs.harvard.edu/abs/2012arXiv1205.3504M

Ma, Z. S. 2012*b*. A unified definition for reliability, survivability and resilience inspired by the handicap principle and ecological stability. International Journal of Critical Infrastructures 8:242–272.

Ma, Z. S. 2013. Stochastic populations, power law and fitness aggregation in genetic algorithms. Fundamenta Informaticae 122:173–206.

Ma, Z. S. 2015. Power law analysis of the human microbiome. Molecular Ecology 24:5428–5445.

Ma, Z. S., Z. Abdo, and L. J. Forney. 2011. Caring about trees in the forest: incorporating frailty in risk analysis for personalized medicine. Personalized Medicine 8:681–688.

Ma, B., L. J. Forney, and J. Ravel. 2012. Vaginal microbiome: rethinking health and disease. Annual Review of Microbiology 66:371–389.

Magurran, A. E. 2004. Measuring biological diversity. Wiley-Blackwell, Hoboken, New Jersey, USA.

May, R. M. 1973. Stability and complexity in model ecosystems. Princeton University Press, Princeton, New Jersey, USA.

May, R. M. 1975. Patterns of species abundance and diversity. Pages 81–120 *in* M. L. Cody and J. M. Diamond, editors. Ecology and evolution of communities. Harvard University Press, Cambridge, Massachusetts, USA.

May, F., K. Gerstner, D. J. McGlinn, X. Xiao, and J. M. Chase. 2017. mobsim: an R package for the simulation and measurement of biodiversity across spatial scales. Methods in Ecology & Evolution. https://doi.org/10.1111/2041-210x.12986

Moya, A., and M. Ferrer. 2016. Functional redundancy-induced stability of gut microbiota subjected to disturbance. Trends in Microbiology 24:402–413.

Oh, J., A. L. Byrd, M. Park, H. H. Kong, and J. A. Segre. 2016. Temporal stability of the human skin microbiome. Cell 165:854–866.

Pastor, J. 2008. Mathematical ecology of populations and ecosystems. Wiley-Blackwell, Hoboken, New Jersey, USA.

Ravel, J., and J. I. Gordon. 2011. Vaginal microbiome of reproductive-age women. Proceedings of the National Academy of Sciences USA 108(Suppl 1):4680–4687.

Rodriguez, S., A. P. Martin, I. Sousa-Pinto, and F. Areans. 2015. Biodiversity effects on macroalgal productivity: exploring the roles of richness, evenness, and species traits. Marine Ecology Progress Series 562:79–91.

Romiguier, J., P. Gayral, M. Ballenghien, A. Bernard, V. Cahais, A. Chenuil, Y. Chiari, R. Dernat, L. Duret, and N. Faivre. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature 515:261–263.

Sechi, V., R. G. M. De Goede, M. Rutgers, L. Brussaard, and C. Mulder. 2017. A community trait-based approach to ecosystem functioning in soil. Agriculture Ecosystems & Environment 239:265–273.

Simpson, E. 1949. Measurement of diversity. Nature 163:688.

Sobel, J. D. 1999. Is there a protective role for vaginal flora? Current Infectious Disease Reports 1:379.

Taylor, L. R. 1961. Aggregation, variance and the mean. Nature 189:732–735.

Taylor, L. R. 1984. Assessing and interpreting the spatial distributions of insect populations. Annual Review of Entomology 29:321–357.

Taylor, L. R. 1986. Synoptic dynamics, migration and the Rothamsted insect survey. Journal of Animal Ecology 1:1–38.

Thibaut, L. M., and S. R. Connolly. 2013. Understanding diversity-stability relationships: towards a unified model of portfolio effects. Ecology Letters 16:140–150.

Valls, A., M. Coll, and V. Christensen. 2015. Keystone species: toward an operational concept for marine biodiversity conservation. Ecological Monographs 85:29–47.

Wang, Q., G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied & Environmental Microbiology 73:5261.

Wang, S., and M. Loreau. 2016. Biodiversity and ecosystem stability across scales in metacommunities. Ecology Letters 19:510–518.

Young, V. B. 2017. The role of the microbiome in human health and disease: an introduction for clinicians. BMJ 356:j831.

Zhang, Z., J. Geng, X. Tang, F. Hong, J. Xu, X. Wen, Z. Ma, and S. Peng. 2014. Spatial heterogeneity and co-occurrence patterns of human mucosal-associated intestinal microbiota. ISME Journal 8: 881.

DATA ACCESSIBILITY

The HVMC dataset is available at http://stm.sciencemag.org/content/4/132/132ra52.full

SUPPORTING INFORMATION

Additional Supporting Information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecs2.2477/full