

Article

Open Access

# Biodiversity metrics on ecological networks: Demonstrated with animal gastrointestinal microbiomes

Zhanshan (Sam) Ma<sup>1,2,\*</sup>, Lianwei Li<sup>1</sup>

<sup>1</sup> *Computational Biology and Medical Ecology Lab, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650201, China*

<sup>2</sup> *Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, Yunnan, 650201, China*

## ABSTRACT

Biodiversity has become a terminology familiar to virtually every citizen in modern societies. It is said that ecology studies the economy of nature, and economy studies the ecology of humans; then measuring biodiversity should be similar with measuring national wealth. Indeed, there have been many parallels between ecology and economics, actually beyond analogies. For example, arguably the second most widely used biodiversity metric, Simpson (1949)'s diversity index, is a function of familiar Gini-index in economics. One of the biggest challenges has been the high “diversity” of diversity indexes due to their excessive “speciation”—there are so many indexes, similar to each country's sovereign currency—leaving confused diversity practitioners in dilemma. In 1973, Hill introduced the concept of “numbers equivalent”, which is based on Renyi entropy and originated in economics, but possibly due to his abstruse interpretation of the concept, his message was not widely received by ecologists until nearly four decades later. What Hill suggested was similar to link the US dollar to gold at the rate of \$35 per ounce under the Bretton Woods system. The Hill numbers now are considered most appropriate biodiversity metrics system, unifying Shannon, Simpson and other diversity indexes. Here, we approach to another paradigmatic shift—measuring biodiversity on ecological networks—demonstrated with animal gastrointestinal microbiomes representing four major invertebrate classes and all six vertebrate classes. The network diversity can reveal the diversity of species interactions, which is a necessary step for understanding the spatial and temporal structures and dynamics of biodiversity across environmental gradients.

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright ©2024 Editorial Office of Zoological Research: Diversity and Conservation, Kunming Institute of Zoology, Chinese Academy of Sciences

**Keywords:** Biodiversity on network; Hill numbers; Animal gut microbiome; Network link diversity; Network species diversity; Network abundance-weighted link diversity

## INTRODUCTION

Biodiversity loss is one of the biggest challenges humans face today and in foreseeable future. Measuring biodiversity, as the foundation for dealing with biodiversity loss, seems to be deceptively simple. In retrospective, although started as early as 1940s, debates and confusions regarding the “diversity” (excessive number) of diversity indexes, continued until the beginning of the 21<sup>st</sup> century, which left the practitioners of biodiversity producing hardly comparable diversity numbers across literatures or regions. For example, arguable the top two widely used diversity indexes (Shannon entropy and Simpson diversity indexes) can be said to measure different aspects of diversity (Shannon & Weaver, 1949; Simpson, 1949), but both are hardly comparable and nor convertible. The rediscovery of Hill numbers during the last decade and so has resolved the major issues, and opened a new chapter for measuring diversity, but it is still not widely applied by practitioners. In last few years, researchers have begun to address another issue in measuring biodiversity—ignoring species interactions—which requires a paradigmatic shift, i.e., measuring biodiversity on ecological networks. The primary objective of this study is to port Ohlmann's et al. (2019) network diversity framework to microbial species co-occurrence networks of animal gastrointestinal microbiomes (with 4 903 samples of 318 animal species covering four major invertebrates and all six vertebrate classes). In consideration of the early debates on diversity indexes and emerging nature of network diversity research, we present a slightly more detailed review on the key literatures with a secondary objective to offer the practitioners a tutorial summary on the latest advances in measuring biodiversity.

Measuring biodiversity is of obvious importance both theoretically and practically. Nearly all existing measures (metrics) for biodiversity contain two key elements: one is

Received: 02 August 2023; Accepted: 25 August 2023; Online: 05 February 2024

Foundation items: This work was supported by the National Natural Science Foundation of China (31970116, 72274192)

\*Corresponding author, E-mail: [ma@vandals.uidaho.edu](mailto:ma@vandals.uidaho.edu)

variety and another is variability of the variety among living organisms in the environment (Chao et al., 2014b; Gaston, 1996). For example, species richness, arguably the simplest diversity index, is simply the number of species (varieties) in a community. For another example, with more sophisticated, familiar Shannon entropy ( $D$ ), the variety can be species, and the variability can be the abundance of each species or species abundance ( $p_i$ ),  $i=1, 2, \dots, S$ , where  $S$  is the number of species, then the diversity measured with Shannon entropy is  $D = -\sum_{i=1}^S p_i \ln(p_i)$ . It can be said that Shannon diversity index simply synthesizes the two components of diversity with Shannon entropy function. Shannon entropy is named after C. E. Shannon, who is the founder of modern information theory, but Shannon never studied biodiversity. The entropy function plays the role of computing the “uncertainty” (variability or lack of information) about species identities in a community (Jost, 2006). For instance, when a community only contains one species, then  $p_1=1$ ,  $S=1$ , and  $D=0$ , there is no uncertainty (variability) in the community when one randomly picks an individual. When there are two species, each with 50% relative abundance, then Shannon entropy  $D \approx 0.69$ . In contrast, assuming the same two species, but one species with 80% abundance and another with 20% abundance, then Shannon entropy  $D \approx 0.50$ . This example illustrates two points: (i) the species richness ( $R$ ) as the number of species in the community,  $R=2$ , is not useful in distinguishing the two communities with different species relative abundances ((50%, 50%) vs. (80%, 20%)) since two communities may have very different productivities; (ii) Shannon entropy measures the uncertainty, which is equivalent with lack of information or unevenness in the community, and turns out to be rather useful in differing the two community. The first community of two species with (50%, 50%) relative species abundances has higher evenness (uncertainty) than the second community that also has two species but with uneven species abundance distribution of (80%, 20%), i.e.,  $D_1 \approx 0.69 > D_2 \approx 0.50$ . The higher evenness (higher entropy) means that, when one randomly picks an individual, it is more difficult to correctly guess its species identity, i.e., higher uncertainty or less informative (low level of information). Therefore, while species richness ( $R$ ) as a rough estimate of biodiversity can be useful in certain occasions, Shannon entropy is obviously more appropriate in measuring biodiversity.

One would wonder why Shannon entropy, which was designed to measure the information associated with an event (more strictly information associated with a random variable), was chosen to measure biodiversity. For example, a convenient choice for measuring biodiversity could be statistical mean (average) of species abundances. Still with the previous example of two species community, if the average is used to measure biodiversity, then both communities will again have the same diversity since  $(20\%+80\%)/2 = (50\%+50\%)/2 = 0.5$ . In this case, statistical average failed to distinguish two obviously different communities. The failure has to do with the statistical distribution of species abundance frequency (distribution), which almost always follows highly skewed J-shaped distributions (e.g., lognormal, power-law distributions) and rarely follows symmetrical Gaussian distribution. For highly skewed distribution such as lognormal and power law distributions, statistical mean (average) is usually a poor statistic. For power law distribution, there is so-called no-average property, which means that the average of the

distribution cannot represent most data points in the distribution. Therefore, ecologists made a right choice for measuring biodiversity by selecting Shannon entropy rather than more familiar statistical average.

However, Shannon entropy is far from perfect, ecologists have been searching for better diversity measures since the introduction of Shannon entropy, familiar Simpson index (Simpson, 1949) was invented almost the same time as Shannon entropy index (Shannon, 1948). As commented in Southwood & Henderson (2000), “the result has been an ‘explosive speciation’ of diversity indices, which initially brought confusion to the subject; in addition, the ubiquitousness of some relationships and the apparent constancy of certain numerical values have added a measure of mystique”. Indeed, for quite a long period, the debates on diversity indexes had confused practitioners of biodiversity research and conservation, which prompted some scholars to publish consumer’s guide materials (e.g., Smith & Wilson, 1996). Furthermore, reviews and monographs were written to discuss deceptively simple diversity indexes (e.g., MacArthur, 1965; Magurran, 2004; Southwood & Henderson, 2000). The core issue of these debates around diversity measures lied in the excessively high “diversity” (of diversity indexes) from “explosive speciation” of diversity indices characterized by Southwood & Henderson (2000). The excessive “diversity” of diversity indexes implies high uncertainty—lack of consensus on the appropriateness of different diversity indexes, and what is even worse is that the values of different diversity indexes are hardly comparable. Using an analogy, measuring biodiversity is more like different countries (researchers) using different currencies (diversity indexes), and unlike currency exchanges, there was not a mechanism to convert the values of different diversity indexes until recent years.

It should be mentioned that there are other issues regarding diversity measures, most notably, the definitions of beta-diversity and gamma-diversity. All of the previous discussion on diversity indexes is, in default, about alpha diversity, which is the diversity of a single community. Gamma diversity refers to the total diversity of multiple (local) communities, equivalent to the alpha-diversity of the metacommunity consisting of the multiple local communities. While gamma diversity can be defined and measured from alpha-diversity in a relatively straightforward manner, defining and measuring beta-diversity turn out to be rather complex and have been a hotspot for debates. Beta-diversity conceptually refers to the difference between (among) two (multiple) communities. Even with the simplest case of two local communities, there are at least two ways to define beta-diversity: additive ( $\beta = \gamma - \alpha$ ) or multiplicative ( $\beta = \gamma/\alpha$ ). Intuitively, both schemes are reasonable and may have their own merits. However, the results from two schemes can be different, and may even generate paradox. Since their definitions ultimately rely on the adoption of basic diversity indexes discussed previously, the beta-diversity definitions inherit the previously identified issue—lack of consensus due to lack of comparability. Using an analogy, if “alpha-diversity” measures the “wealth evenness or unevenness” (diversity) of a country, “gamma-diversity” measures the wealth evenness of the world, then “beta-diversity” can be defined to measure the difference between two countries. In fact, the well-known Gini-index, which measures the wealth distribution (evenness or unevenness) of a country, is a function of Simpson’s diversity index in community ecology (Ma & Ellison, 2018). Obviously, a

beta-diversity of Gini-index for comparing two countries can be multiplicative or additive. For example, country A's Gini-index is twice the world average Gini-index would be multiplicative version, and the difference between country A's Gini-index is 0.2 or (20%) would be additive version. The two numbers from multiplicative and additive definitions *per se* are not comparable at all.

Ecology can be considered as nature's economy or economy of plants and animals, and economy can be considered as the ecology of humans; both ecological systems and economic systems consist of various production and consumption systems in the form of food webs (nature), supply chains (human industry), driven by fundamental processes such as competitions, cooperation, symbiosis, predation (war) etc. Therefore, measuring biodiversity in ecology is not unlike measuring wealth in human societies in economics. It was not coincidental that Gini-index in economics and Simpson's diversity index in ecology can be derived from each other. In fact, Ma & Ellison (2018, 2019) dominance concept in ecology is applicable to both population and community levels, and if applied to economics, then it means that one can define a Gini-index for a citizen, besides a Gini-index for a country. The underlying mechanism for their relationship is that dominance metrics, Simpson's diversity index, and Gini-index possess mathematical functional relationships (Ma & Ellison, 2018, 2019).

Beyond the previously identified two issues regarding biodiversity measures, i.e., excessive number of biodiversity indexes and the complexity of beta-diversity, a third widely known, also equally serious issue with the previous ones, is the so-termed sampling problem in diversity estimation. The issue is rooted in the hard reality that it is hardly possible to survey the whole community and diversity must be estimated from finite samples taken from the community under investigation. Therefore, errors in diversity estimates are hardly avoidable with finite sampling efforts, and the problem is simply to estimate the diversity of "population" (sensu statistics, not sensu ecology) from the diversity of "samples" (sensu statistics, not ecology). In standard statistics, variance, standard errors, and confidence intervals must be attached to the diversity estimates, but this practice has rarely been followed in diversity estimation. An even more serious part of the sampling problem is rooted in the previously mentioned, highly skewed distribution of species abundance distributions. The high-skewness makes it extremely costly for samples to fully cover low-abundance (rare) species, especially possibly many singletons (species represented by single individual) in a community. This also means that the diversity estimates are usually very sensitive to samples or sampling efforts, and it can be rather difficult to know when sufficient sampling efforts are made. To remedy the sampling problem, researchers have developed the so-termed rarefaction estimation approach to extrapolating the number of total species in the community, but the rarefaction itself is not a perfect solution either because it depends on another rather complex problem—Good's coverage problem, which was first studied by the founder of modern computer science, Alan Turing, who studied it but did not publish it and later published by I. J. Good with the permission from Turing (Good, 1953, 2000; Good & Toulmin, 1956).

A fourth issue in measuring biodiversity is to do with the scope of biodiversity. Fairly speaking, this issue is more to do with the scope of biodiversity research and is far less technical

than the previously discussed three types of problems. Much of biodiversity research has been focused on species diversity, especially the species diversity of plants and animals, until the recent two decades when microbial diversity receives increasing attention thanks to the revolutionary metagenomic sequencing technologies and the launch of Human Microbiome Project (HMP) and Earth Microbiome Project (EMP). However, biodiversity should not be limited to the diversity of organisms *per se* for scientific, technological or humanity advances (benefits). Intuitively, the elements of biodiversity, variety and variability (uncertainty) can be extended from organisms *per se* to any of their associated properties, most notably, genes (genome), metagenomes, functions, phylogeny, metabolic functions, etc. (Chao et al., 2014b; Ma, 2018c, 2023; Ma & Ellison, 2021; Ma & Li, 2018; Ma et al., 2020). Therefore, measuring biodiversity, as foundation for community ecology and as practical tools for biodiversity conservation has far broader scope (from gene, species, functions, etc.) and significant implications (conservation strategies, ecosystem services, global changes, etc.).

The four issues (problems) mentioned above are now largely resolved thanks to the introduction (Hill, 1973), actually the re-introduction (Jost 2007; Chao et al., 2012, 2014b; Ellison, 2010; Jost, 2010; Ma, 2018c), of a new metric system for biodiversity, i.e., the Hill numbers (Hill, 1973). Hill numbers are actually based on Rényi (1961) entropy. It can be said that Rényi (1961) entropy is of more mathematical generality than Shannon (1948) entropy, and therefore has achieved wider applications in many fields, including economics and quantum computing. Hill (1973) formulated its application for measuring biodiversity inspired by its success in economics, where the concept of "numbers equivalent of elements" was originated. According to Hill (1973), the diversity of a community should be measured with "numbers equivalent" (i.e., namely, Hill numbers)—the number of equally likely elements (such as individuals or species) necessary to generate the observed diversity being measured by a diversity index. It might be this scientifically rigorous but abstruse Hill numbers in ecology scientifically rigorous, but likely too abstruse to easily comprehend concept/statement that had prevented the wide adoption of Hill numbers in ecology. The Hill numbers (Hill, 1973) had not received attention it deserves until nearly four decades later when a group of scholars reintroduced it into ecology. It is estimated that during the last decade and so, more than a dozen core methodological papers exclusively devoted to the Hill numbers and its applications have been published in top-tiers ecological journals such as *Ecological Monographs*, *Ecology*, *Ecology Letters*, *Annual Review of Ecology and Systematics*, and *Methods in Ecology and Evolution*. Here, using less rigorous, analogical terminology, we further expose Hill (1973) concept of "numbers equivalent of elements" as follows: Basically, he was saying that, I have discovered that using Rényi entropy, we can establish a gold standard, in which all existing diversity indexes, possibly many other diversity indexes people may continue to propose in future, can be unified with my or Hill numbers. In other words, the values of other diversity indexes can be converted with my "gold standard" into a series of Hill numbers, for which I assign a diversity order  $q=0, 1, 2, \dots$  Using an analogy of economics again, what Hill (1973) wished to establish was not unlike to link the US dollar to gold at the rate of \$35 per ounce under Bretton Woods system. Under the Bretton Woods system, US

dollars were as good as gold, and all currencies could be pegged to US dollars, equivalently to gold (e.g., Mason & Asher, 1973). Similarly, with Hill numbers, all other diversity metrics, existing ones and possibly future ones, can be converted into Hill numbers (Hill, 1973). With Hill numbers as “gold” standard, the previously identified first issue as criticized by Southwood and Henderson (2000) as excessive “diversity” of diversity index, i.e., the lack of consensus on which diversity index is better, is a moot issue. That is, Hill numbers are “gold” standard, and all other diversity indexes can be converted into one of the Hill numbers corresponding to a diversity order ( $q$ ). Chao et al. (2012) further clarified Hill numbers as,  $D(q) = [\sum_{i=1}^S p_i^q]^{1/(1-q)}$ , where  $S$  is the number of species in the community,  $q=0, 1, 2, \dots$  is the diversity order,  $D(q)$  is the Hill numbers at diversity order  $q$ . When diversity order  $q=0$ ,  $D(0)=S$ , this is the traditional species richness. When  $q=1$ ,  $D(1) = \exp(D) = \exp\{-\sum_{i=1}^S p_i \ln(p_i)\}$ , which is the exponential function of previously defined Shannon entropy ( $D$ ), hence, Shannon entropy can be converted into the Hill number at the 1<sup>st</sup> diversity order ( $q=1$ ). When  $q=2$ ,  $D(2) = (1/\sum_{i=1}^S p_i^2)$ , which is the reciprocal of Simpson diversity index, again convertible to the gold standard of Hill numbers. In the Hill numbers system, different diversity indexes (different Hill numbers at different diversity order  $q$ ) actually possess more intuitive interpretations. For example, when  $q=0$ , species richness means that every species is treated equally, regardless of their abundances. Using an analogy in economics, it means that each person, regardless of their wealth, only contribute one count to the population number. When  $q=1$ , the Hill number (diversity) is the number of “typical species” in the community where each species is weighted in proportional to its abundance. In economics, this would mean that each person is counted in proportional to his or her wealth. When  $q=2$ , the Hill number is weighted in favor for more abundant species. In economics, this would mean that wealthy people get more weights, and poor people weigh in less and could be ignored. In other words, the Hill number of diversity order  $q=2$ , is more suitable for measuring the diversity of dominant species, when ignoring rare species is justified. Not coincidentally, the Gini-index, which can be converted with Simpson’s index or  $D(2)$ , is particularly suitable for revealing the effects of millionaires, the unevenness of wealth distribution. It can be said that evenness and unevenness (approximate to dominance or heterogeneity) are the both sides of same coin (Li & Reynolds, 1995; Ma, 2015, 2018a, 2018b, 2019a, 2019b, 2021c, 2022; Ma & Ellison, 2018, 2019; Ma & Taylor, 2020).

It was probably the above analyzed merits of Hill numbers, as well as other advantages of Hill numbers, a few consensus regarding the previously identified four issued have emerged during the last decade. The first consensus is that “numbers equivalent”, i.e., the Hill numbers is the most appropriate for measuring alpha diversity (Chao et al., 2012; Chao et al., 2014b; Ellison, 2010; Ma et al., 2019). Second, the multiplicative version (partition) of Hill numbers is more appropriate than additive partition for deriving beta-diversity (Chao et al., 2012, 2014b). Furthermore, to deal with the sampling problem in measuring diversity (i.e., the third issue), Chao et al. (2014a) extended rarefaction approach to Hill numbers (Chao & Jost, 2012, 2015; Chao et al., 2013, 2015a, 2015b). In addition, to measure other kinds of biodiversity beyond species diversity (the fourth issue mentioned previously), the Hill numbers for phylogeny (Chao et al.,

2014b; Chiu et al., 2014), genes, and metagenomes (Ma, 2018c, 2023; Ma & Li, 2018; Ma, et al., 2020) have been developed and applied.

The previous introduction seems to suggest that the major issues in measuring diversity have been resolved during the first decade of this century. It turned out that Hill numbers, while advantageous over virtually all existing diversity indexes, still cannot deal with another important issue in community ecology, that is species interactions. Back to the previous examples of two species communities, a community of wolf and bear and a community of wolf and deer would be rather different. In the community (habitat) of wolf and bear, both are apex predators and opportunistic carnivores and their relationship is largely competitive, but usually they are not each other’s prey. On the other hand, the community of wolf and deer, wolves may try to prey on as many as deer as they can. Traditional diversity indexes do not consider species interactions and offer little insights for comparing the previous two communities involving wolves. A solution to deal with such an issue requires a paradigmatic shift in measuring biodiversity, one venue is to move biodiversity concept on complex networks that captures species interactions. Defining and computing network diversity or moving biodiversity measures on to species interaction networks is an emerging field in recent years (Eagle et al 2010), and two pioneering studies are Ohlmann et al. (2019) and Luna et al. (2020). However, both approaches did not offer universally applicable approaches for building species interaction networks. The objective of the present study is to innovatively apply Ohlmann et al. (2019) approach to measure the microbial diversity on AGM (animal gastrointestinal microbiome) networks. Besides applying Ohlmann et al. (2019) definitions and procedures, we first construct AGM networks with SparCC algorithms and further developed statistical methods for comparing the network biodiversity. To the best of our knowledge, this study should be the first attempt to measure biodiversity on microbial networks, and we use big datasets including 4 903 AGM samples covering 319 animal species representing four major invertebrate classes and all six vertebrate classes, collected from 108 published studies on animal gut gastrointestinal microbiomes (Ma, 2021a, 2021b; Ma et al., 2022).

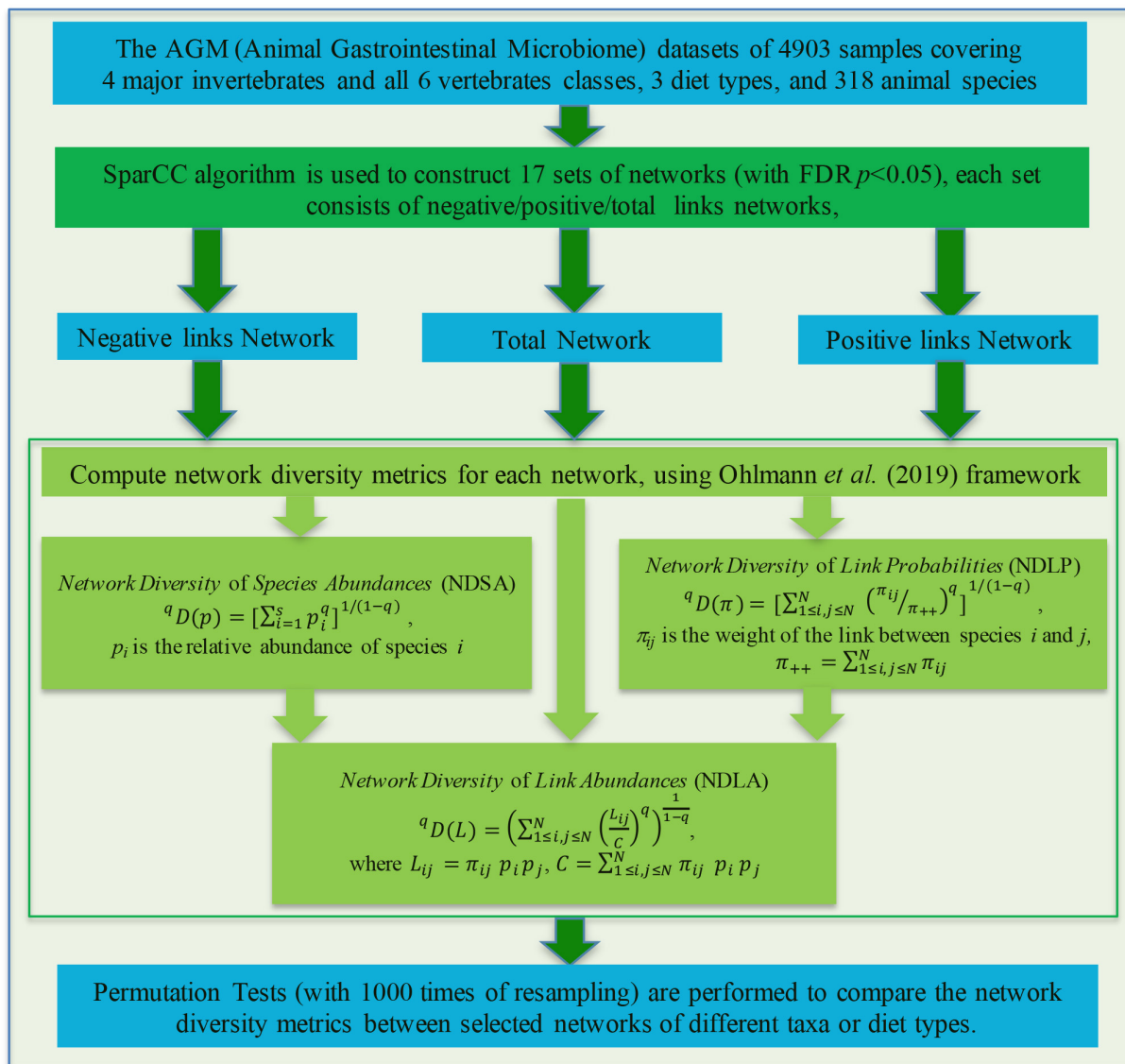
## MATERIALS AND METHODS

Figure 1 illustrates the study design and computational procedures used in this study. The computational codes we developed are provided in the online supplementary information. Besides using our own codes, Friedman & Alm (2012) SparCC algorithm (<https://bitbucket.org/yonatanf/sparcc>) (see introduction below) is used to build complex networks. Specifically, we use their C++ version of SparCC, named FastSpar (<https://github.com/scwatts/FastSpar>), which uses the same algorithm as SparCC but runs much faster.

### The AGM (Animal Gastrointestinal Microbiome) dataset

We surveyed 108 publications of animal gastrointestinal microbiomes (AGM) and collected 6 900 AGM samples using 16S-rRNA metagenomic sequencing technology. Enforcing quality control eliminated 1 997 samples and left 4 903 AGM samples spanning all three major animal phyla (i.e., Nematoda, Arthropoda, and Chordates), and ten major animal classes (including four major invertebrates and all six vertebrates), 59 orders, 142 families, 261 genera, and 318





**Figure 1** A Diagram showing the study design and computational procedures of this study

species. The ten animal classes covered by the 4 903 samples are: Chromadorea, Arachnida, Malacostraca, Insecta, Chondrichthyes, Actinopteri, Amphibia, Sauropsida, Aves, and Mammalia. Insects (76 species) and mammals (123 species) represented for the largest proportions of host species, and contributed 979 and 1 499 AGM samples, respectively. Both the classes occupy 62.6% of host species and 51% of AGM samples. These same datasets have been analyzed in our previous studies (Ma, 2020a, 2020b; Ma et al., 2022) with different research objectives from this study, and detailed information about these AGM samples can be found in the Supplementary Table S7 of Ma et al. (2022).

We further categorize the 4 903 AGM samples into three diet types, including 1 421 carnivores, 1 229 herbivores, and 1 473 omnivores groups, respectively. The remaining 780 AGM samples could not be categorized into any of the three major diet types and were excluded in the diet-type related modeling (but still included in regular host taxon-based analysis).

Besides enforcing the previously mentioned quality control (details outlined in Ma et al., 2022) that removed 1 997 (29%) AGM samples, we recomputed the OTU tables from the raw 16S-rRNA reads using QIIME-2 software (v.2018.6.0, Boylen et al., 2019). The OTUs from QIIME2 are species because it

uses the machine learning algorithm to cluster the reads of 100% similarity as an OTU, and the clustered OTU is then aligned to a taxon (species in our re-computation) against the Greengenes database. A total of 473 359 bacterial species were obtained from the recalculation of the 4 903 AGM samples.

#### Network diversity: measuring biodiversity on complex ecological networks

One motivation to measure biodiversity on complex networks is the previous introduced examples of two different communities (in the introduction section): community of (wolf and deer) vs. community of (wolf and bear). The species relationships (interactions) in both communities are obviously different. In the first community, wolves may eat up deer population, and the community diversity in terms of Shannon evenness entropy may approach to zero. While in the second community, wolves and bears may reach some kind of equilibriums and the community diversity is unlikely to be zero. In classic community ecology, species interactions are typically modelled with differential equations, known as Lotka-Volterra equations, pioneered by Alfred Lotka and Vito Volterra, most time actually independently (e.g., Lotka, 1925;

Volterra, 1931). Lotka-Volterra equations started with two-species interactions, and they can be extended to  $N$ -species interactions with differential equation system of  $N$ -equations. In fact, Lotka-Volterra equations have been the primary approach to investigate ecological community (Kingsland, 1995; May, 1973), in particular, species interactions in the settings of food webs, and the approach can be harnessed to investigate different species interactions including competition, predation, symbiosis, commensalism as well as their implications effects onto community stability (e.g., whether or not wolves will extinguish deer populations). However, the approach has certain limitations. First, the analytical solutions are generally too complex to be tractable even for communities of moderate number of species, and computer simulations must be used. Second, parameterization of large-scale differential equations systems with experimental data of plants and animals are usually practically infeasible. Metagenomic technologies, which makes it possible to obtain the species abundance data of microbial community samples, within days if not hours, offer unprecedented opportunities to parameterize large-scale model systems for microbial communities. In the meantime, the advances in complex network science have made the option of using Lotka-Volterra differential equation systems less attractive because the network approach is not only simpler but also produce excellent visualization of species interaction patterns. The network graphs with nodes representing for species and links (edges) representing for their interactions, are much more intuitive to most ecologists than phase portraits generated by Lotka-Volterra differential equations.

There have been extensive studies on ecological networks during the last few decades; however, to the best of our knowledge, we are only aware of a handful of studies that formally defined biodiversity on ecological networks (e.g., including: Bersier et al., 2002; Poisot et al., 2012a, 2016; Ohlmann et al., 2019; Luna et al., 2020). Here, we focused on the latest two studies. As a side note, there were earlier studies on network diversity, notably Rafols & Meyer (2010); Stirling (2007), which compute the diversity of network nodes, but often ignore the nodes interactions or lack comprehensive synthesis of node and link information as performed by the latest studies on network diversity mentioned previously (e.g., Ohlmann et al., 2019).

Ohlmann et al. (2019) proposed a generic model for ecological networks by extending the probabilistic network model by Poisot et al. (2016). The probabilistic network model, while theoretically solid as model of ecological networks, in our opinion, seems to be unnecessarily complex for the purpose of this study given that the extensive AGM datasets allow us to build realistic microbial species co-occurrence networks (SCN). Furthermore, we argue that the SCNs we are to build with SparCC correlation coefficients, are essentially an empirical implementation of their probabilistic network model, implemented with metagenomic species abundance distribution (SAD) data. With these considerations, we adopt Ohlmann et al. (2019) definitions for network diversity, but not their approach to building microbial networks for the AGM datasets. In Luna et al. (2020) study on network diversity, they defined network diversity mostly in the context of trophic food webs, and they did not offer approach to building complex ecological networks with large-scale metagenomic datasets. In consideration that they did not adopt Hill numbers, given the advantages of Hill numbers, we

skip their definitions of network diversity in this study. An advantage of Luna et al. (2020) paper is that they presented insightful discussion on network diversity, especially their implications to community ecology and biogeography, from which we draw excellent insights in our discussion section. Below, we outline the procedures to build the microbial SCN, and in the next subsection, we summarize the definitions and procedures for computing the network diversity for the SCNs built with the AGM datasets.

To build species co-occurrence network, the first step is to select an algorithm for computing species correlation coefficients, i.e., using correlation relationships to reveal cooccurrence relationships. We choose the SparCC algorithm developed by Friedman & Alm (2012) (<https://bitbucket.org/yonatanf/sparcc>), rather than using more traditional Pearson correlation coefficient and Spearman rank correlation coefficient, because it alleviates a fundamental issue associated with compositional species abundance data. The issue is that, with compositional data (Aitchison, 2003), many standard methods for estimating correlation coefficients including Pearson and Spearman correlations are biased since the relative abundances must be summed to 100% (Friedman & Alm, 2012). This constraint makes the relative abundances not independent, tending to be negatively correlated, regardless of the true correlations (Friedman & Alm, 2012). Theoretically, the true correlations can be accurately determined by absolute abundances, but the true absolute abundances are not available in metagenomic sequencing data. There is a C++ version of SparCC, named FastSpar (<https://github.com/scwatts/FastSpar>), which uses the same algorithm as SparCC but much faster. During the computation, to minimize the spurious effects of extremely rare OTUs, we filter out the extremely rare OTUs that occurs in less than or equal to 2% of the AGM samples for building the networks in this study.

In our opinion, an observation on SparCC algorithm is worthy of notice. The compositional effects Friedman & Alm (2012) aimed to alleviate with their SparCC algorithm is related to a suit of logical inference fallacies in statistics often under different names, most notably, association paradoxes, Simpson's paradox, aggregation, amalgamation, or reversal paradoxes (Bradford et al., 2014a, 2014b; Carlson, 2023). The problem was already noted by Pearson (Pearson, 1897, cited in Friedman & Alm, 2012), but obviously still needs attention today. Although statistician Edward Simpson's name was used, he only pointed out that the association paradoxes were widely known before his classic 1951 paper (Simpson, 1951), and obviously he never made the claim for the discovery. Of course, he did claim his Simpson's diversity index, which he published two years earlier (Simpson, 1949) and is arguably considered as the second most widely used diversity index, perhaps only after Shannon entropy (Shannon, 1948). Even with the Renyi-entropy based Hill numbers (Hill, 1973) that unified virtually all existing diversity indexes, Simpson's diversity index, as a reciprocal of Hill numbers at diversity order ( $q=2$ ), still occupies an important position. Simpson's paradox may not only occur in the association (correlation) analysis, as in Pearson or Spearman correlation coefficients with compositional data, but also in traditional biodiversity analysis, at least with species richness (e.g., Scheiner et al., 2000). However, detecting when Simpson's paradox may occur is usually not trivial, and avoiding its introduction in the methods for biodiversity analysis is of obvious significance, for

which we will have a brief elaboration in the discussion section.

#### Network diversity framework proposed by Ohlmann et al. (2019)

The design goal of Ohlmann et al. (2019) framework for network diversity is to understand the structure (topology) and composition of ecological networks across spatial and temporal scales and along environmental gradients (Ohlmann et al., 2019; Pellissier et al., 2017). The general principles they proposed can be summarized as following four points.

(i) First, Hill numbers, given the generally unifying power of the Hill numbers, they rightly choose to port the regular Hill numbers, which are briefly reviewed in the Introduction section, to ecological networks.

(ii) Second, the framework must consider both the probabilistic nature of biotic interactions and the abundances of species or groups. Ohlmann et al. (2019) argued that ecological networks should be analyzed across various species aggregation levels because species redundancy and trophic structures in communities (ecosystems) require so. Hence, in their framework, species and species groups (which can be different taxa, diet types, functional guilds, etc.) are treated as explicit units for defining and computing network diversity. With this notion, they distinguish microscopic, mesoscopic and macroscopic scales, depending on the levels of species aggregation adopted. The microscopic scale network refers to network built with all species, the mesoscopic scale network (also termed image network) refers to network built with species groups, and the macroscopic scale network refers to the single value of connectedness representing the probability of interactions between any two species. The macroscopic network only has one single node with abundance=100%, and a single link to itself, and it is essentially the case when image network degenerates to its limit of grouping all species as a single group.

(iii) Third, the framework is applicable for measuring network diversity locally (with alpha-diversity), regionally and globally (with gamma-diversity), and between local communities (networks) (with beta-diversity). The framework is hence across local networks and defined on the meta-network scale. The metanetwork is a weighted network, consisting of multiple ( $K$ ) realized local networks, and any grouping of species is defined on the meta-network, which in our opinion ensures the feasibility to implement the Hill numbers.

(iv) Fourth, the framework includes three kinds of network diversity: all are based on the Hill numbers, and all are applicable for species or species groups (i.e., from microscopic to macroscopic scale networks): (A) The network diversity in species (group) abundances (NDSA),  $D(p)$  is computed with species (group) relative abundance. (B) The network diversity in link probabilities (NDLP),  $D(\pi)$  is computed with adjacent matrix, in which link probabilities are not weighted by species (group) abundances and it assumes evenly distributed species abundances. (C) The network diversity in link abundances (NDLA) is computed with abundance weighted adjacent matrix. In the following, we briefly introduce the formulae of the three ND definitions based on Ohlmann et al. (2019) probabilistic network model at species level.

Assuming a given region that is inhabited by individuals of  $n$  different species with relative species abundances  $p=(p_1,$

$p_2, \dots, p_n)$ ,  $p_i$  is the probability of picking an individual belonging to species  $i$ . The occurrence probability of species interaction between two individuals of species  $i$  and  $j$  follows a Bernoulli law of parameter  $\pi_{ij}$ . The probabilistic species network is a weighted network  $G$ , in which there are node  $V_i$  (representing species  $i$ ) and  $V_j$  (representing species  $j$ ), with relative abundance  $p_i$  and  $p_j$  as well as the weight ( $\pi_{ij}$ ) of the link ( $V_i, V_j$ ). The probability of picking a link that connects (correlated) two individuals belonging to species  $i$  and  $j$ , respectively, should be:  $L_{ij} = \pi_{ij}p_i p_j$ .

Ohlmann et al. (2019) further proposed the concept of image network, a scheme to classify mesoscopic network of different species aggregation levels: Assuming that there are  $N$  species groups ( $C_1, C_2, \dots, C_N$ ) from the previously defined probabilistic species network  $G$  ( $N \leq n$ ), where  $n$  is the number of species in  $G$ . Obviously, when  $N=n$ , then both image network and species network are the same, i.e., each species is designated as a group. Also, obviously, when  $N=1$ , the image network becomes the macroscopic network which has single node, single link, and a connectedness value as mentioned previously. When  $1 < N < n$ , the image network has  $N$  species groups, and the image network is said to have a scale of  $S=N/n$ , which is further elaborated below.

With the image network, a new set of nodes is  $V=(V_1, \dots, V_N)$ . The relative abundances of  $V_i$  (or group  $i$ ) can be computed as:

$$p_i = \sum_{k \in C_i} p_k \quad (1)$$

The occurrence probability of interaction, or link probability, between individuals from group  $C_i$  and  $C_j$  is:

$$\pi_{ij} = \frac{\sum_{k \in C_i, l \in C_j} \pi_{kl} p_k p_l}{\sum_{k \in C_i} p_k \sum_{l \in C_j} p_l} \quad (2)$$

The link abundances between individuals of group  $i$  and  $j$  are defined as:

$$L_{ij} = \sum_{k \in C_i, l \in C_j} \pi_{kl} p_k p_l \quad (3)$$

which is the probability of picking a link that connects two individuals from group  $i$  and  $j$ , the counterpart of  $L_{ij}$  at species level (the species probabilistic network).

Based on the above definitions, and the definition of Hill numbers (Chao et al., 2012, 2014b; Hill, 1973; Rényi, 1961), Ohlmann et al. (2019) introduced the following three types of network diversity (ND). In the following three definitions,  $q$  is the diversity order of Hill numbers and was also given a new name of “viewpoint parameter” linked to the weight assigned to dominant species vs. rare species onto assemble rules (Chalmandrier et al., 2015).

The network diversity of species abundances (NDSA) is defined as:

$${}^q D(p) = \left( \sum_{i=1}^N p_i^q \right)^{\frac{1}{1-q}} \quad (4)$$

The network diversity of link probabilities (NDLP) is defined as:

$${}^q D(\pi) = \left( \sum_{1 \leq i, j \leq N} \left( \frac{\pi_{ij}}{\pi_{++}} \right)^q \right)^{\frac{1}{1-q}} \quad (5)$$

where

$$\pi_{++} = \sum_{1 \leq i, j \leq N} \pi_{ij}$$

The network diversity of link abundances (NDLA) is defined as:

$${}^qD(L) = \left( \sum_{1 \leq i, j \leq N} \left( \frac{L_{ij}}{C} \right)^q \right)^{\frac{1}{1-q}} \quad (6)$$

where

$$C = \sum_{1 \leq i, j \leq N} \pi_{ij} p_i p_j$$

The difference between NDLP (eqn. 5) and NDLA (eqn. 6), as mentioned previously, is that the former is not weighted by node abundance and assumes that the node abundances are evenly distributed, and the latter is weighted by node abundances. Their primary commonality is that both NDLP and NDLA are based on link probabilities.

Ohlmann et al. (2019) defined the scale of the image network as:  $S=N/n$ . If  $S=1$ , the network is considered to be at microscopic scale, or species level image network. If  $S=1/n$ , the network is considered to be at macroscopic scale. If  $1/n < S < 1$ , the network is considered to be at mesoscopic scale. In this study, we adopt the microscopic or original species-scale network, built with SparCC correlations (Friedman & Alm, 2012).

In summary, Ohlmann et al. (2019) network diversity has three metrics as defined by eqns. (4–6), network diversity of relative species abundances (NDSA) (i.e., diversity of nodes, eqn. 4), network diversity of link probabilities (NDLP) (i.e., diversity of link probabilities, eqn. 5), and network diversity of link abundances (NDLA) (i.e., diversity of nodes-links, eqn. 6). In our opinion, the most informative network diversity metric is the last one—the NDLA since it synthesizes the diversity information of both nodes and links on the network, while the other two NDs only quantify the diversity information of either nodes or links.

## RESULTS

### Building microbial co-occurrence networks for AGM datasets

Since the AGM datasets, with its 4 903 microbial samples covering 318 animal species from four major invertebrate classes and all six vertebrate classes, are so extensive that building microbial networks for all species or all taxa will be too computationally intensive, we selectively build 17 microbial networks to compute the network diversity. Specifically, at animal host species level, we randomly chose two animal species, i.e., *Apis mellifera* and *Bos taurus* to construct their gut microbial networks. At animal host class level, we built their microbial networks for all 10 animal classes, one for each class. At animal phylum level, we built two networks, one for invertebrates and another for vertebrates. Finally, we built three microbial networks for three diet types (carnivores, herbivores, and omnivores), respectively.

To build the 17 microbial networks for selected 17 animal taxa/diet types, we first compute the SparCC correlation coefficients with FDR control of  $P=0.05$  from the OTU tables of the AGM datasets. The SparCC correlation coefficients are advantageous over commonly used Pearson and Spearman

correlation coefficients because it overcomes their issues in processing compositional (abundance) data, but the computational load of SparCC is far heavier (about 1 000 times more) than that of Pearson or Spearman methods, which explains why we selectively built 17 networks, rather than for all possible animal species or taxa.

We further compare the network diversity of different taxa/diet-types by performing randomization (permutation) tests with 1 000 times of re-sampling and  $P$ -value threshold=0.05. This means that, for each comparison of both taxa or diet-types, 1 000 pairs of SparCC networks must be built for the two taxa or diet types (out of the previously counted 17 networks), rather than only two networks. A total of 9 000 microbial networks with SparCC correlation algorithm were constructed in this study to perform the permutation tests. Figure 2 illustrates the microbial network for the class of Amphibia to get a general glimpse of the networks built in this study.

### Computing the network diversities for microbial co-occurrence networks of AGMs datasets

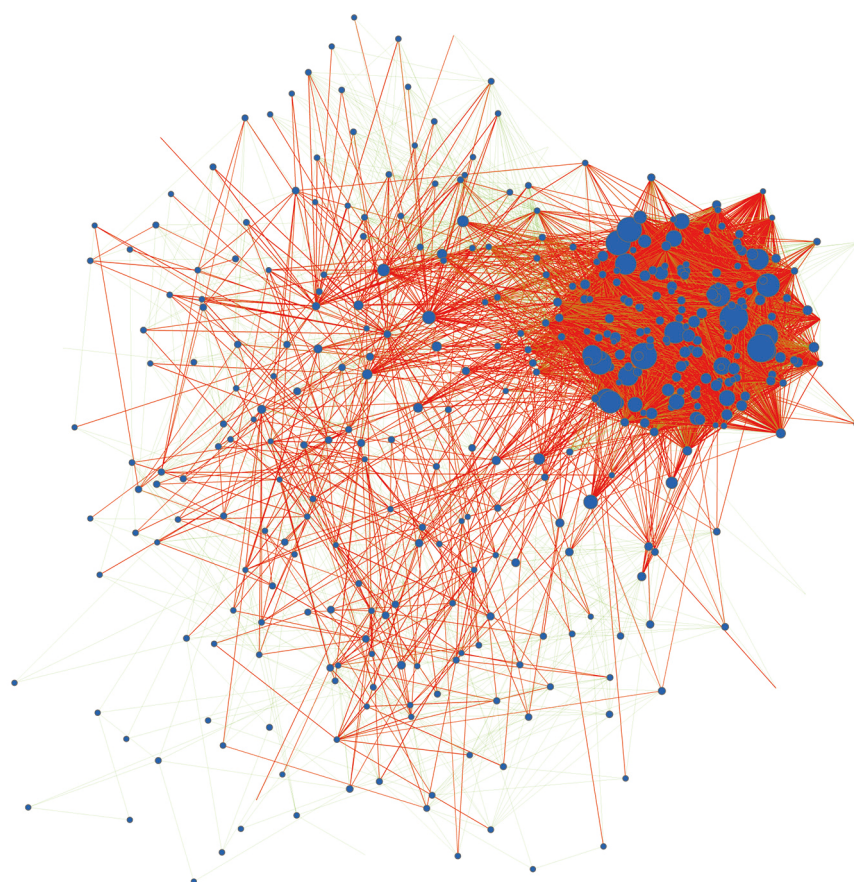
We computed Ohlmann's network diversity (OND) for the AGM microbial networks of the 17 taxa/diet- types hosts. Since there are two types of correlations (links), positive and negative, there can be three different ways to compute OND for each network. That is, computing the OND for positive links, negative links, and total links (using the absolute values of SparCC correlation coefficients).

Table 1 below exhibits the results of the OND for the total links, and Supplementary Table S1 in the online supplementary information (OSI) exhibits the corresponding results of the permutation tests for comparing the OND of different taxa/diet types. Supplementary Tables S2, S3 and Tables S4, S5 in the OSI exhibited the results of the OND and corresponding permutation tests for positive and negative links, respectively.

Comparing the OND results of total links, positive links and negative links, we realize that separating the positive and negative links did not introduce tangible benefits other than the additional computational complexity. For this, we focus on the results of the total links by using absolute values of SparCC correlation coefficients (Table 1), and we do not recommend the separate treatments for general purpose study of biodiversity on networks.

As shown in Table 1, the first metric of network diversity, namely network species (node) diversity, computed as the Hill numbers of species relative abundances, is essentially the same as traditional species diversity in community ecology, except that the species are conceptually called network nodes. Obviously, in the case of network species diversity, links (correlations) do not weigh in. The second metric of network diversity, namely network link diversity, computed as the Hill numbers of the link probability, is essentially the diversity of links in the network, and the species abundances do not weigh in here. The third, namely the network diversity of abundance-weighted links or simply "network abundance-link diversity" for short is obviously the most comprehensive network diversity metric among all three OND indexes, and both species abundances and links (correlations) weigh in for its definition. For this, we consider the third metric of network diversity of abundance-weighted links (the last four columns in Table 1 and Supplementary Tables S1–S5) as the most representative. Figures 3–5 illustrate the three types of





**Figure 2 Microbial species cooccurrence networks of Amphibia class, built with SparCC correlation coefficients (with FDR control of  $P$ -value=0.05)**

Green links and red links represent positive and negative correlations, respectively; the line thickness represents for the level of correlations, and circle (network node) size represents the level of connectedness (=network degree/abundance).

network diversity explained above for the 17 taxa/diet-types, using all links in the AGM networks (without distinguishing the positive and negative links).

Supplementary Tables S1, S3, and S5 showed the results ( $P$ -values) of the permutation tests from comparing the ONDs of commensurable taxa/diet types. It turned out that none of the compared taxa/diet-types exhibited significant differences in their network diversity. In our opinion, the network diversity, especially the network diversity of abundance-weighted links is essentially a measure of network “heterogeneity” as further discussed below.

## CONCLUSIONS

The primary objective of this article is to systematically analyze the microbial network diversity with extensive 4 903 AGM (animal gastrointestinal microbiome) samples covering four major invertebrate classes and all six vertebrate classes, as well as three major diet types. We first selectively constructed 17 microbial co-occurrence networks with SparCC correlation algorithm, as the implementations for Ohlmann et al. (2019) probabilistic network model. The 17 microbial networks span species (2 representative species), class (10 classes), phylum (invertebrate and vertebrate), diet types (carnivores, herbivores, and omnivores). For each of the 17 microbial cooccurrence networks, we computed their network diversities in terms of Ohlmann et al. (2019) three metrics. We separately computed the network diversity for positive, negative and total network links, but found that the separate

treatments have little benefits and using total links (or absolute values of SparCC correlation coefficients) is both simple and sufficient for measuring network biodiversity. These results, for the first time, sketched out the animal microbial diversity of biotic interactions across animal species, class, phylum scales from host perspective, and across community, metacommunity and landscape scales from microbial perspective.

A secondary objective of this article has something to do with the two realities in the field of measuring biodiversity: (i) The rediscovery of Hill numbers approximately a decade ago offers a unified approach to measuring biodiversity, and the unification establishes a “convertible common currency” or gold standard by default. However, the Hill numbers have not been widely used by practitioners, perhaps due to the interpretations of the Hill numbers are still not sufficiently clear. Recall, it took scientists of biodiversity research near four decades to correctly decode Hill (1973) message, and in fact, the field of biodiversity measure requires deep mathematical foundation and is deceptively simple. (ii) While Hill numbers resolved the four major issues associated with measuring diversity, as outlined in the introduction section, the Hill numbers *per se* in terms of species abundances cannot measure species interactions. This inability requires porting the Hill numbers on ecological networks, but this topic is still in its infancy and there are only a handful of studies (Bersier et al., 2002; Luna et al., 2020; Ohlmann et al., 2019; Poisot et al., 2012b, 2016). In consider of these realities, we presented reviews on the key literatures on measuring

**Table 1** The three metrics of network diversity, all measured in Hill numbers of different diversity orders ( $q=0, 1, 2, 3$ ) of AGM (animal gastrointestinal microbiome) networks of all links, built with SparCC correlation coefficients (FDR (False Discovery Rate) control of  $P$ -value=0.05)

Taxon level	Taxon	NDSA (network diversity in species (or group) abundances): computed with species (group) relative abundance $D(p)$				NDLP (network diversity in link probabilities): computed with adjacent matrix (i.e., matrix of correlation coefficients) $D(\pi)$				NDLA (network diversity in link abundances): computed with species (group) abundance-weighted adjacent matrix (i.e., matrix of link abundances) $D(L)$			
		$q=0$	$q=1$	$q=2$	$q=3$	$q=0$	$q=1$	$q=2$	$q=3$	$q=0$	$q=1$	$q=2$	$q=3$
Species	<i>Apis mellifera</i>	218	74.41	46.36	35.81	5427	4519.36	3437.32	2585.82	5427	235.02	81.77	54.90
	<i>Bos taurus</i>	1362	714.48	278.11	140.78	301000	284275.50	266705.76	249073.1	301000	50280.13	6583.18	2198.32
Class	Chromadorea	894	105.05	42.29	27.78	53722	46703.49	38876.75	31821.13	53722	885.81	184.09	89.91
	Arachnida	412	63.22	38.25	29.96	3667	3406.86	3090.41	2744.53	3667	74.06	25.65	16.63
	Malacostraca	576	39.40	14.83	10.67	37106	35836.17	34569.37	33312.37	37106	364.07	70.47	37.05
	Insecta	1140	127.82	33.13	20.59	40072	32730.32	26323.38	21159.26	40072	576.99	64.52	33.19
	Chondrichthyes	475	38.33	9.73	6.05	15655	15036.55	14419.30	13822.17	15655	300.45	66.60	37.72
	Actinopteri	1186	291.88	110.06	64.21	148809	115001.15	82478.72	60653.24	148809	6468.19	1189.99	585.70
	Amphibia	408	129.39	46.08	28.20	15968	15549.59	15136.68	14734.27	15968	1678.04	336.89	161.11
	Sauropsida	1689	425.34	127.09	69.04	222130	201525.25	179310.77	157683.7	222130	9890.82	1247.97	559.67
	Aves	1811	397.57	137.03	86.43	133688	115577.97	96241.29	78322.05	133688	3578.11	689.80	350.41
	Mammalia	4245	1157.71	299.76	166.03	228601	189925.94	156655.69	131124.0	228601	13343.53	1510.62	585.38
Phylum	Invertebrates	1565	249.85	68.49	38.97	51781	42382.64	34328.43	27887.94	51781	1409.42	129.21	56.31
	Vertebrates	3378	1269.24	475.36	274.46	143030	104040.89	74225.48	55291.42	143030	10938.46	1173.95	440.00
Diet types	Carnivore	2311	813.18	315.47	189.97	197864	146148.64	101829.03	74379.82	197864	8917.03	1318.69	583.28
	Herbivore	2390	354.60	81.30	45.11	96952	79281.92	64221.70	52273.98	96952	2399.03	388.10	196.51
	Omnivore	3201	922.23	255.16	119.74	161060	132017.77	107003.60	87555.73	161060	7306.79	269.03	91.73
Mean (SE)		1603.59 (283.74)	421.98 (98.06)	139.91 (32.78)	79.64 (18.15)	109207.76 (22155.09)	91997.65 (19453.08)	76403.16 (17271.93)	64377.91 (15601.85)	109207.76 (22155.09)	6979.17 (2904.78)	901.8 (377.54)	357.52 (127.28)
Range (min, max)		(218, 4245)	(38.33, 1269.24)	(9.73, 475.36)	(6.05, 274.46)	(3667, 301000)	(3406.86, 284275.5)	(3090.41, 266705.76)	(2585.82, 249073.1)	(3667, 301000)	(74.06, 50280.13)	(25.65, 6583.18)	(16.63, 2198.32)

biodiversity with slightly more details than usually presented in a research article, which can be considered as a secondary objective of this article—hoping to generate some tutorial effect.

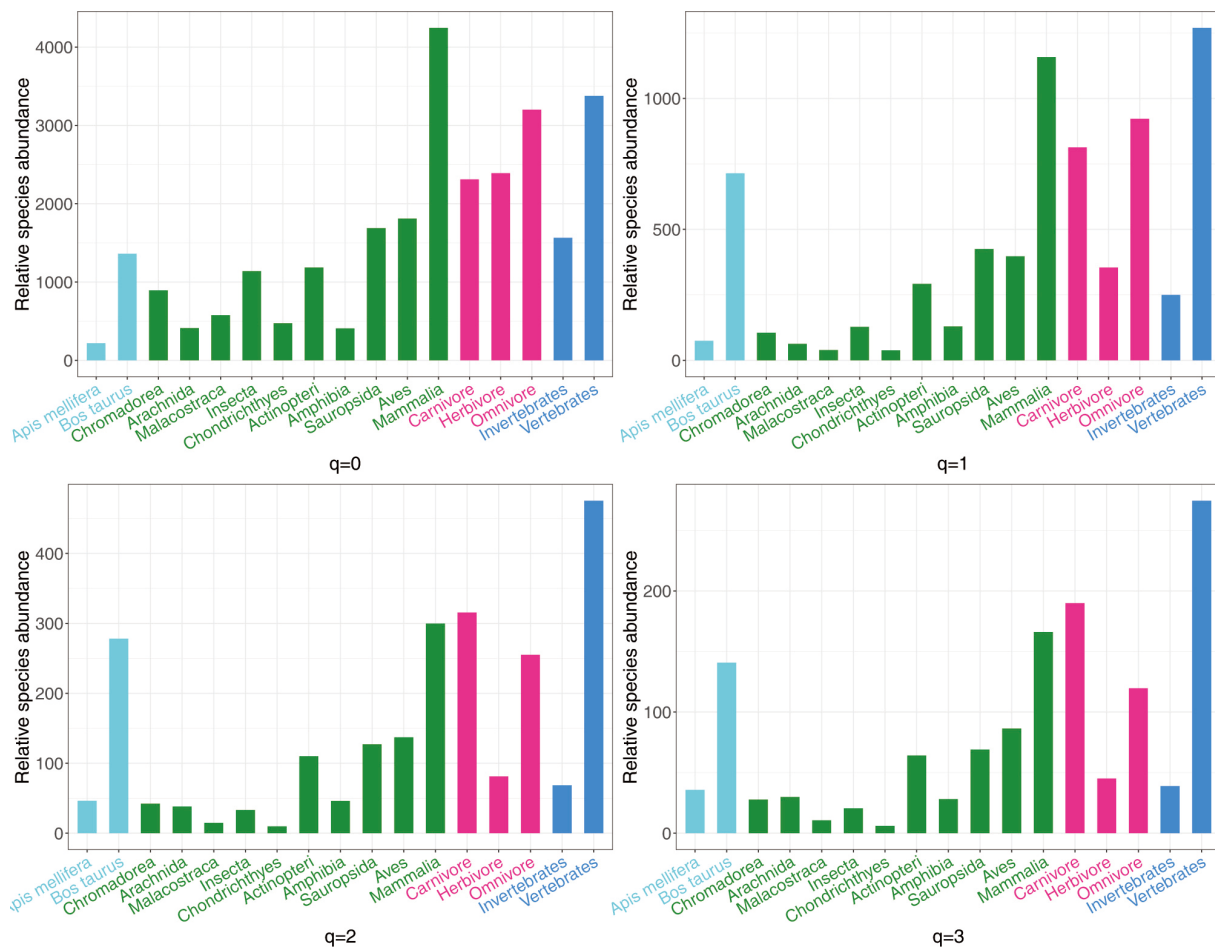
## DISCUSSION

Measuring microbial diversity on networks or network diversity is still an emerging research topic, and to the best of our knowledge, this article should be the first comprehensive estimation of the network diversity of animal microbiomes with extensive metagenomic samples. Still, our study is at the stage of estimating the network diversity, and a consequence of such an exploratory study is that the implications (significance) of our results are not self-evident. That is, the significance requires future studies. For this, we can only speculate the implications of our results. In the remainder of this article, we suggest three types of possible implications from our study, including (i) dynamics (including extinctions) of species interactions may possess importance significance in conservation biology and agriculture (e.g., crop pollinations)—both rare species and rare interactions matter; (ii) network diversity may be helpful for resolving previously mentioned Simpson's paradox; (iii) network diversity and ecological heterogeneity.

A general consensus is that dominant and more abundant species interact more than less abundant and rare species (Dee et al., 2019; Luna et al., 2020), and our AGM datasets support this consensus in a pre-experiment data analysis (in which there is a statistically significant positive correlation between network node degrees and node abundances). The dominant and abundant species are assumed to play more

important roles in maintaining ecosystem functions than rare species and their interactions (Luna et al., 2020). As argued by Luna et al. (2020), since majority of species are rare and are often specialists, determining how the rare species interact with common/dominant species and with other rare species may help to broaden our understanding of community structure and dynamics. Endangered species are not only rare in abundances, but also, usually, rare in species interactions. As Daniel Jansen once warned: “*What escapes the eye, however, is a much more insidious kind of extinction: the extinction of ecological interactions*” (Janzen, 1974; cited by Luna et al., 2020). In crop-pollinator mutualism networks, the breakup of the links may jeopardize the crop flowering and influence crop production (Luna et al., 2020). The suppression of biological invasions in agriculture and forestry by foreign species should obviously pay attention to the establishment of new links (interactions) with native species.

Next, we revisit the previously mentioned Simpson's paradox, which has various aliases such as association paradox, Yule-Simpson effect in statistics. Simpson's paradox refers to a logical contradiction that may occur when the marginal association between two categorical variables is qualitatively different (e.g., the opposite trend or relationship) from the partial association between the same two variable when a third associated, but previously unobserved or controlled variable, is introduced (Carlson, 2023). Simpson's paradox was often illustrated with a simple example of sex bias during the graduate student admission. Inspired by the example of graduate admission (Carlson, 2023) and a recently debated preprint on the relationship between COVID-19 mortality and the consumption of vegetable kinds (Fonseca

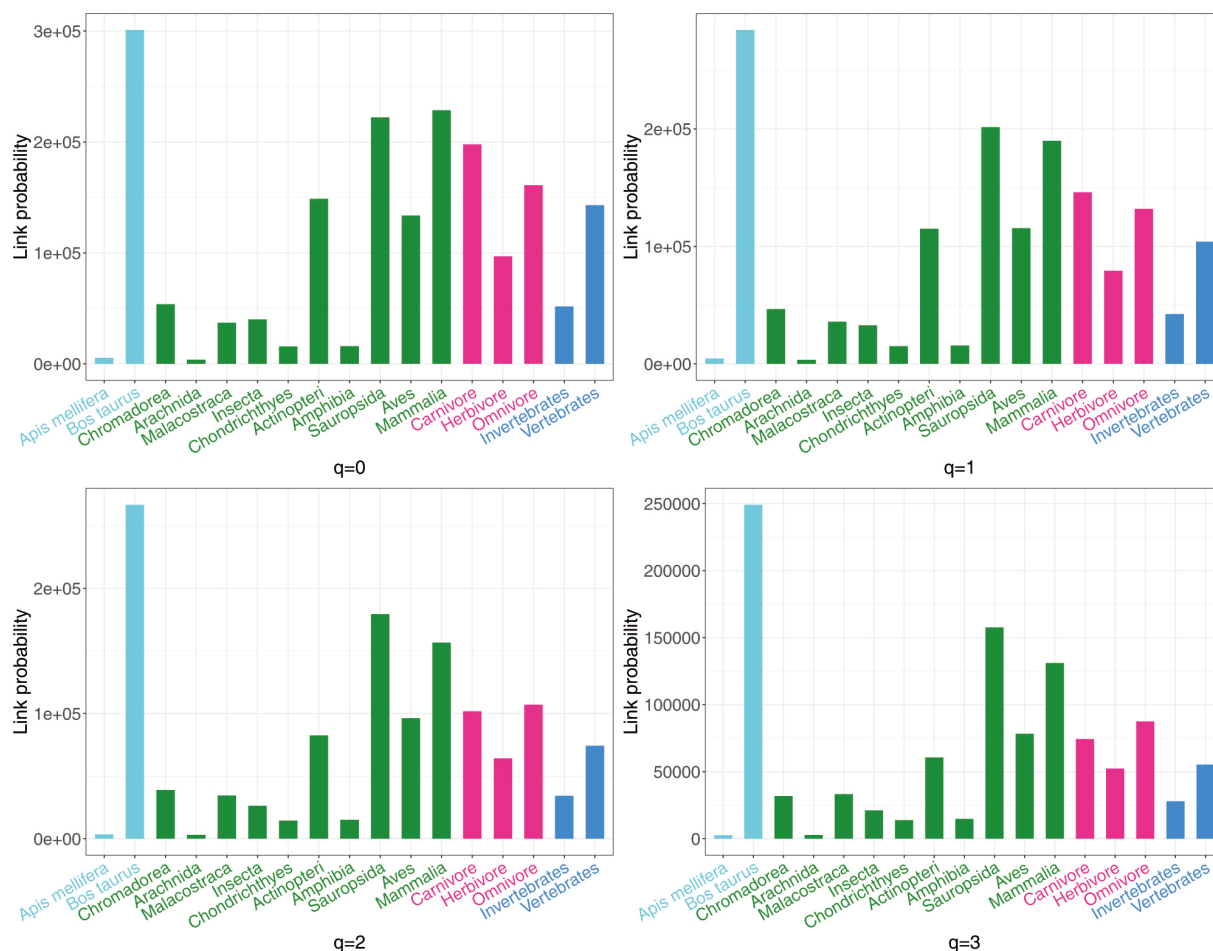


**Figure 3** The network species (node) diversity measured in Hill numbers at diversity orders ( $q=0, 1, 2, 3$ ), computed with relative species (node) abundances on SparCC species cooccurrence networks of all links

et al., 2020), here, we contrive a fictitious scenario with the association between COVID-19 severity and the usage of probiotics. For example, one may try to test a hypothesis that probiotics may strengthen one's immune system by maintain a healthy gut microbiome, and consequently could lower the mortality of COVID patients. It is possible to find a negative correlation between probiotics usages and mortality rates. However, if the researcher forgot to control the nutrition of the participants. For example, it might be the case that the group with probiotics may have better appetite than the control group, and if the difference in food intake is significant, then a data reanalysis by regrouping in terms of the level of food intake, the previously observed probiotics-mortality relationship may become statistically insignificant or even reversed. In other words, the confounding effects of food intake, as a third variable, may create a paradox with previous trend when it is ignored. In this case, it is hard to distinguish the effects of probiotics with better nursing care, possibly with a better chef. Perhaps, it is for similar reason, the previously mentioned study on the COVID-19 mortality and consumption kinds of vegetables in different European countries had received extensive attention among ecologists, given that the authors started their paper with: "Many foods have an antioxidant activity, and nutrition may mitigate COVID-19. To test the potential role of vegetables in COVID-19 mortality in Europe, we performed an ecological study" (Fonseca et al., 2020). We are not in a position to cast support or doubt to their study. It is the wide discussion on their study in social media (the paper was tweeted more than 2 000 times within a

few months) that prompted us to make our own fictitious example for explaining the fallacy that does not seem to be uncommon.

As mentioned previously, Simpson's paradox may happen in the inferences of species correlation coefficients, if traditional Pearson or Spearman correlation coefficients were used. For this reason, we adopted Friedman & Alm (2012) SparCC correlations. Simpson's paradox can also occur with the simplest diversity index—species richness (as revealed by Scheiner et al., (2000)). Another ecological scenario that Simpson's paradox could occur can be the studies of species interactions when indirect effects are ignored. Obviously pairwise species interactions are simplified views, and indirect effects among species interactions are likely to be ubiquitous. In a still simplified view of three species system, consisting of donor, transmitter and receiver, the donor's effect is transmitted to the received indirectly through the transmitter. The indirect effects may act through so-called "interaction chain indirect effect", in which a species indirectly affect others by influencing the abundance of an intermediate transmitter species. Alternatively, the indirect effects may act through so-called "interaction modification indirect effect", in which the donor species alters the per capita effect of the intermediate transmitter species on the receiver species, but without changing the abundance of transmitter species (Morin, 1999). These indirect effects may generate so-terms apparent competition, consumptive competition, indirect mutualism, indirect commensalism, keystone predation, trophic cascade, etc. (Morin, 1999). Determining indirect effects *per se* itself is



**Figure 4** The network link diversity measured in Hill numbers at diversity orders ( $q=0, 1, 2, 3$ ), computed with the link probability on SparCC species cooccurrence networks of all links

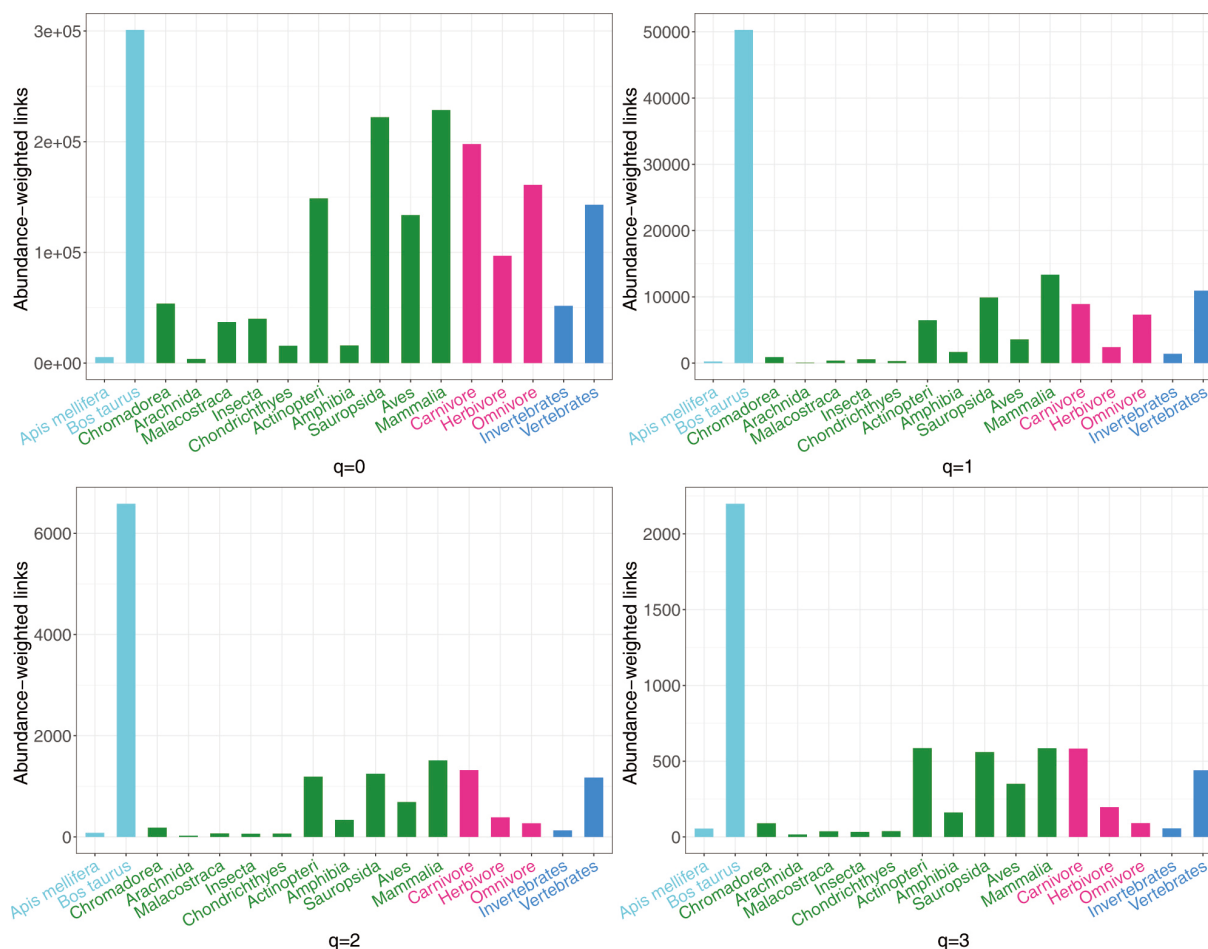
hardly possible without well designed experimental studies, not to mention of detection of potential Simpson's paradox in data analysis. On the positive side, the unified approaches such as Ohlmann et al. (2019) network diversity framework with the Hill numbers, supported by SparCC correlation algorithms, is likely to lower the risk of falling in the trap of the paradox, thanks to their rigorous schemes used to aggregate species, constructing subnetworks, and computing interaction probabilities. Of course, we do not claim that these approaches can eliminate the risks of Simpson's paradox, which ultimately depends on properly designed experiments for data collection.

If we recognized a fundamental different between network diversity such as Ohlmann et al. (2019) and Luna et al. (2020) with traditional biodiversity is the inclusion of species interactions with network diversity, then network diversity actually reaches out to another fundamental concept in ecology, i.e., the ecological heterogeneity. One may argue that heterogeneity and evenness (a major aspect of diversity, and the other is species richness) can be considered both sides of same coin since literally heterogeneity and unevenness are close to be synonyms. Nevertheless, this analogy highlighted their similarity but seemed to ignore their difference. By pointing out this negligence, we do not mean to criticize the notion, and in fact, we have made similar analogies previously in our own publications (Ma, 2020a,2020b; Ma & Ellison, 2018, 2019). In our opinion, except for a handful of exceptions, from academia through societies to many cultures, diversity and heterogeneity are

often not rigorously distinguished, if not used interchangeably. One exception we are aware of distinguish diversity from heterogeneity is Shavit et al. (2016) quotes of Robert Frost (1916) "*Two roads diverged in a wood, and I took the one less traveled by, and that has made all the difference.*" We concur with them that heterogeneity has received relatively little attention than diversity, especially in community ecology where diversity research has occupied a center position for long time. Another exception is a motto by Aaron Ellison "Zoos are diversity, and natural ecosystems are heterogenous" (personal communication, cited in Ma (2020a)). A fundamental difference between heterogeneity and diversity identified by Shavit & Ellison et al. (2016), is that heterogeneity must consider interactions in a group context (group behavior), while diversity is measured in numbers or relative abundances (more strictly equivalent numbers with Hill numbers). From this perspective, the network diversities in link probabilities or link abundances, demonstrated in this study, actually measure heterogeneity. It should be noted that the concept of heterogeneity is not limited to ecology, and its usages are as wide as diversity, if not more. In fact, in many fields, diversity and heterogeneity are interwoven and sometimes crossly measured. This recent monograph (Guajardo, 2023) that mixes diversity and heterogeneity concepts and uses Simpson's diversity index to measure heterogeneity is an example to demonstrate the enormous challenge to distinguish the both.

Finally, with gratitude, we would like to mention certain limitations of this study, which were prompted by anonymous





**Figure 5** The network diversity measured in Hill numbers at diversity orders ( $q=0, 1, 2, 3$ ), computed with the abundance-weighted link probability on SparCC species cooccurrence networks of all links

expert reviewers of this article. First, although we previously claim that the diversity partition problem is largely solved in traditional community ecology, and furthermore, Ohlmann et al. (2019) also developed the diversity partition solution for their network diversity indexes by distinguishing single (local) network (alpha diversity) vs. meta-network (gamma diversity), the diversity partition in a network setting is much more complex than that in traditional community ecology. For example, to build a network, usually multiple community samples from same or different (local) communities must be taken, which may involve meta-community if the samples are taken from different communities. In the present study, our measures of network diversity are limited to alpha diversity, given that a single network is built for each taxa level. Theoretically, with the AGM datasets and Ohlmann et al. (2019) framework, it is possible to build hierarchical networks corresponding to different taxa levels, with each subnetwork corresponding to a taxa level. Then, it is possible to apply Ohlmann et al. (2019) partition scheme for network diversity. Second, in ecological networks, network specialization (or interaction unevenness, see Luna et al. (2020)) quantifies the degree of how specialized the interactions between species are. As with network diversity, network specialization is a multi-faceted concept and can be measured in different ways (see an excellent review by Poisot et al. (2012a)). A comparative study of network specialization with network diversity would be a very interesting research topic. Third, although we previously claim that the sampling problem in conventional community ecology is largely solved, the

sampling problem also arises in networks when interaction frequency data are used to measure network diversity. On this issue, two important studies exist (Chacoff et al., 2012; Jordano, 2016). Fourth, although we suggest previously that SparCC algorithm (Friedman & Alm, 2012) is one of the most appropriate algorithms for building species co-occurrence network and is used in this study given its advantage in processing compositional data, other alternative algorithms can produce different results even for same datasets. Currently there is not yet a consensus on possible standard method for building ecological networks, and the issue is obviously worthy of further investigations. Fifth, the approach we implemented and demonstrated in this article can be applied to other general animal microbiome datasets such as Cui et al. (2019), Li & Ma (2019), Li et al. (2020), Xiong et al. (2019), Zhang et al. (2019), Zhu et al. (2015). Further testing with other animal microbiome datasets may be helpful not only for testing the approach, but also for revealing biological insights of those specific studies.

#### DATA AVAILABILITY

All data re-analyzed in this study are available in the public domain. Summary information can be found in Ma et al. (2022) Microbiome–host–phylogeny relationships in animal gastrointestinal tract microbiomes, *FEMS Microbiology Ecology*, Vol. 98(2), fiac021, <https://doi.org/10.1093/femsec/fiac021>

#### SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

## COMPETING INTERESTS

The authors declare that they have no competing interests.

## AUTHORS' CONTRIBUTION

Z.M defined the research objective. L.L conducted data analysis. ZSM wrote the manuscript and revised the manuscript. All authors read and approved the final version of the manuscript.

## ACKNOWLEDGMENTS

We are deeply indebted to the editor-in-chief, Prof. Yao Yong-Gang, and two anonymous expert reviewers for their insightful comments and suggestions for improving our paper.

## REFERENCES

- Aitchison J. 2003. The Statistical Analysis of Compositional Data. Caldwell: Blackburn Press, 416.
- Bersier LF, Banasek-Richter C, Cattin MF. 2002. Quantitative descriptors of food-web matrices. *Ecology*, **83**(9): 2394–2407.
- Bradford MA, Wood SA, Bardgett RD, et al. 2014a. Discontinuity in the responses of ecosystem processes and multifunctionality to altered soil community composition. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(40): 14478–14483.
- Bradford MA, Wood SA, Bardgett RD, et al. 2014b. Reply to Byrnes et al. : Aggregation can obscure understanding of ecosystem multifunctionality. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(51): E5491.
- Carlson BW. 2023(2023-10-31). Simpson's paradox. <https://www.britannica.com/topic/Simpsons-paradox>.
- Chacoff NP, Vázquez DP, Lomáscolo SB, et al. 2012. Evaluating sampling completeness in a desert plant-pollinator network. *Journal of Animal Ecology*, **81**(1): 190–200.
- Chalmandrier L, Münkemüller T, Lavergne S, et al. 2015. Effects of species' similarity and dominance on the functional and phylogenetic structure of a plant meta-community. *Ecology*, **96**(1): 143–153.
- Chao AN, Chiu CH, Hsieh TC. 2012. Proposing a resolution to debates on diversity partitioning. *Ecology*, **93**(9): 2037–2051.
- Chao AN, Chiu CH, Hsieh TC, et al. 2015a. Rarefaction and extrapolation of phylogenetic diversity. *Methods in Ecology and Evolution*, **6**(4): 380–388.
- Chao AN, Chiu CH, Jost L. 2014b. Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through Hill numbers. *Annual Review of Ecology, Evolution, and Systematics*, **45**: 297–324.
- Chao AN, Gotelli NJ, Hsieh TC, et al. 2014a. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs*, **84**(1): 45–67.
- Chao AN, Hsieh TC, Robin L, et al. 2015b. Unveiling the species-rank abundance distribution by generalizing the Good-Turing sample coverage theory. *Ecology*, **96**(5): 1189–1201.
- Chao AN, Jost L. 2012. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, **93**(12): 2533–2547.
- Chao AN, Jost L. 2015. Estimating diversity and entropy profiles via discovery rates of new species. *Methods in Ecology and Evolution*, **6**(8): 873–882.
- Chao AN, Wang YT, Jost L. 2013. Entropy and the species accumulation curve: a novel entropy estimator via discovery rates of new species. *Methods in Ecology and Evolution*, **4**(11): 1091–1100.
- Chiu CH, Jost L, Chao AN. 2014. Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs*, **84**(1): 21–44.
- Cui YF, Wang FJ, Yu L, et al. 2019. Metagenomic comparison of the rectal microbiota between rhesus macaques (*Macaca mulatta*) and cynomolgus macaques (*Macaca fascicularis*). *Zoological Research*, **40**(2): 89–93.
- Dee LE, Cowles J, Isbell F, et al. 2019. When do ecosystem services depend on rare species?. *Trends in Ecology & Evolution*, **34**(8): 746–758.
- Eagle N, Macy M, Claxton R. 2010. Network diversity and economic development. *Science*, **328**(5981): 1029–1031.
- Ellison AM. 2010. Partitioning diversity. *Ecology*, **91**(7): 1962–1963.
- Fonseca SC, Rivas I, Romaguera D, et al. 2020. Association between consumption of vegetables and COVID-19 mortality at a country level in Europe. *medRxiv*, doi:<https://doi.org/10.1101/2020.07.17.20155846>.
- Friedman J, Alm EJ. 2012. Inferring correlation networks from genomic survey data. *PLoS Computational Biology*, **8**(9): e1002687.
- Gaston KJ. 1996. Biodiversity: A Biology of Numbers and Difference. Oxford: Blackwell.
- Good IJ. 1953. The population frequencies of species and the estimation of population parameters. *Biometrika*, **40**(3-4): 237–264.
- Good IJ. 2000. Turing's anticipation of empirical Bayes in connection with the cryptanalysis of the naval enigma. *Journal of Statistical Computation and Simulation*, **66**(2): 101–111.
- Good IJ, Toulmin GH. 1956. The number of new species, and the increase in population coverage, when a sample is increased. *Biometrika*, **43**(1-2): 45–63.
- Guajardo SA. 2023. Assessing Organizational Diversity with the Simpson Index. Cambridge Scholars Publishing.
- Hill MO. 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology*, **54**(2): 427–432.
- Janzen DH. 1974. The deflowering of Central America. *Natural History*, **83**: 48–53.
- Jordano P. 2016. Sampling networks of ecological interactions. *Functional Ecology*, **30**(12): 1883–1893.
- Jost L. 2006. Entropy and diversity. *Oikos*, **113**(2): 363–375.
- Jost L. 2007. Partitioning diversity into independent alpha and beta components. *Ecology*, **88**(10): 2427–2439.
- Jost L. 2010. Independence of alpha and beta diversities. *Ecology*, **91**(7): 1969–1974.
- Kingsland SE. 1995. Modeling Nature: Episodes in the History of Population Ecology. Chicago: University of Chicago Press.
- Li H, Reynolds JF. 1995. On definition and quantification of heterogeneity. *Oikos*, **73**(2): 280–284.
- Li HZ, Li N, Wang JJ, et al. 2020. Dysbiosis of gut microbiome affecting small intestine morphology and immune balance: a rhesus macaque model. *Zoological Research*, **41**(1): 20–31.
- Li W, Ma ZS. 2019. Diversity scaling of human vaginal microbial communities. *Zoological Research*, **40**(6): 587–594.
- Lotka AJ. 1925. Elements of Physical Biology. Baltimore: Williams and Wilkins.
- Luna P, Corro EJ, Antoniazzi R, et al. 2020. Measuring and linking the missing part of biodiversity and ecosystem function: the diversity of biotic interactions. *Diversity*, **12**(3): 86.
- Ma ZS. 2015. Power law analysis of the human microbiome. *Molecular Ecology*, **24**(21): 5428–5445.
- Ma ZS. 2018a. DAR (diversity–area relationship): extending classic SAR (species–area relationship) for biodiversity and biogeography analyses. *Ecology and Evolution*, **8**(20): 10023–10038.
- Ma ZS. 2018b. Diversity time-period and diversity-time-area relationships exemplified by the human microbiome. *Scientific Reports*, **8**(1): 7214.
- Ma ZS. 2018c. Measuring microbiome diversity and similarity with Hill Numbers. In: Nagarajan M. Metagenomics: Perspectives, Methods, and Applications. Amsterdam: Academic Press.
- Ma ZS. 2019a. Sketching the human microbiome biogeography with DAR (diversity–area relationship) profiles. *Microbial Ecology*, **77**(3): 821–838.
- Ma ZS. 2019b. A new DTAR (diversity–time–area relationship) model

- demonstrated with the indoor microbiome. *Journal of Biogeography*, **46**(9): 2024–2041.
- Ma ZS. 2020a. Assessing and interpreting the metagenome heterogeneity with power law. *Frontiers in Microbiology*, **11**: 648.
- Ma ZS. 2020b. Predicting the outbreak risks and inflection points of COVID - 19 pandemic with classic ecological theories. *Advanced Science*, **7**(21): 2001530.
- Ma ZS. 2021a. Cross-scale analyses of animal and human gut microbiome assemblies from metacommunity to global landscape. *mSystems*, **6**(4): e0063321.
- Ma ZS. 2021b. Philosophical skepticism concerning the neutral theory or randomness: misplaced or misconceived? A reply to Madison, “Stochasticity and randomness in community assembly: real or as-if?”. *mSystems*, **6**(5): e01014–21.
- Ma ZS. 2021c. Coupling power laws offers a powerful method for problems such as biodiversity and COVID-19 fatality predictions. *arXiv*, doi:<https://doi.org/10.48550/arXiv.2105.11002>.
- Ma ZS. 2022. Coupling power laws offers a powerful modeling approach to certain prediction/estimation problems with quantified uncertainty. *Frontiers in Applied Mathematics and Statistics*, **8**: 801830.
- Ma ZS. 2023. Measuring microbiome diversity and diversity-scaling analysis with hill numbers. In: Nagarajan M. *Metagenomics: Perspectives, Methods, and Applications*. 2<sup>nd</sup> ed. Amsterdam: Academic Press (in press).
- Ma ZS, Ellison AM. 2018. A unified concept of dominance applicable at both community and species scales. *Ecosphere*, **9**(11): e02477.
- Ma ZS, Ellison AM. 2019. Dominance network analysis provides a new framework for studying the diversity-stability relationship. *Ecological Monographs*, **89**(2): e01358.
- Ma ZS, Ellison AM. 2021. Toward a unified diversity–area relationship (DAR) of species and gene diversity illustrated with the human gut metagenome. *Ecosphere*, **12**(11): e03807.
- Ma ZS, Li LW. 2018. Measuring metagenome diversity and similarity with Hill numbers. *Molecular Ecology Resources*, **18**(6): 1339–1355.
- Ma ZS, Li LW, Gotelli NJ. 2019. Diversity-disease relationships and shared species analyses for human microbiome-associated diseases. *The ISME Journal*, **13**(8): 1911–1919.
- Ma ZS, Li LW, Zhang YP. 2020. Defining individual-level genetic diversity and similarity profiles. *Scientific Reports*, **10**(1): 5805.
- Ma ZS, Li W, Shi P. 2022. Microbiome–host-phylogeny relationships in animal gastrointestinal tract microbiomes. *FEMS Microbiology Ecology*, **98**(2): fiac021.
- Ma ZS, Taylor RAJ. 2020. Human reproductive system microbiomes exhibited significantly different heterogeneity scaling with gut microbiome, but the intra-system scaling is invariant. *Oikos*, **129**(6): 903–911.
- MacArthur RH. 1965. Patterns of species diversity. *Biological Reviews*, **40**(4): 510–533.
- Magurran AE. 2004. *Measuring Biological Diversity*. Malden: Blackwell Publishing.
- Mason ES, Asher RE. 1973. *The World Bank Since Bretton Woods: The Origins, Policies, Operations and Impact of the International Bank for Reconstruction*. Washington: Brookings Institution.
- May RM. 1973. *Stability and Complexity in Model Ecosystems*. Princeton: Princeton University Press.
- Morin PJ. 1999. *Community Ecology*. 3<sup>rd</sup> ed. Malden: Blackwell Science.
- Ohlmann M, Miele V, Dray S, et al. 2019. Diversity indices for ecological networks: a unifying framework using Hill numbers. *Ecology Letters*, **22**(4): 737–747.
- Pearson K. 1897. Mathematical contributions to the theory of evolution. —On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London*, **60**(359-367): 489–498.
- Pellissier L, Albouy C, Bascompte J, et al. 2017. Comparing species interaction networks along environmental gradients. *Biological Reviews*, **93**(2): 785–800.
- Poisot T, Canard E, Mouillot D, et al. 2012a. The dissimilarity of species interaction networks. *Ecology Letters*, **15**(12): 1353–1361.
- Poisot T, Canard E, Mouquet N, et al. 2012b. A comparative study of ecological specialization estimators. *Methods in Ecology and Evolution*, **3**(3): 537–544.
- Poisot T, Cirtwill AR, Cazelles K, et al. 2016. The structure of probabilistic networks. *Methods in Ecology and Evolution*, **7**(3): 303–312.
- Rafols I, Meyer M. 2010. Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience. *Scientometrics*, **82**(2): 263–287.
- Rényi A. 1961. On Measures of Entropy and Information. *Berkeley Symp. on Math. Statist. and Prob.*, **1961**: 547–561.
- Scheiner SM, Cox SB, Willig M, et al. 2000. Species richness, species–area curves and Simpson’s paradox. *Evolutionary Ecology Research*, **2**: 791–802.
- Shannon CE. 1948. A Mathematical Theory of Communication. *The Bell System Technical Journal*, **27**: 379–423.
- Shannon CE, Weaver W. 1949. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.
- Shavit A, Kolumbus A, Ellison AM. 2016(2016-10-04). Two Roads Diverge in a Wood: Indifference to the Difference Between ‘Diversity’ and ‘Heterogeneity’ Should Be Resisted on Epistemic and Moral Grounds. <http://philsci-archive.pitt.edu/12432/>.
- Simpson EH. 1949. Measurement of diversity. *Nature*, **163**(4148): 688.
- Simpson EH. 1951. The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, **13**(2): 238–241.
- Smith B, Wilson JB. 1996. A consumer’s guide to evenness indices. *Oikos*, **76**(1): 70–82.
- Southwood TRE, Henderson PA. 2000. *Ecological Methods*. 3<sup>rd</sup> ed. Oxford: Blackwell Science.
- Stirling A. 2007. A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, **4**(15): 707–719.
- Volterra V. 1931. Variations and fluctuations of the number of individuals in animal species living together. In: Chapman RN. *Animal Ecology*. New York: McGraw–Hill.
- Xiong JB, Nie L, Chen J. 2019. Current understanding on the roles of gut microbiota in fish disease and immunity. *Zoological Research*, **40**(2): 70–76.
- Zhang MX, Song TZ, Zheng HY, et al. 2019. Superior intestinal integrity and limited microbial translocation are associated with lower immune activation in SIVmac239-infected northern pig-tailed macaques (*Macaca leonina*). *Zoological Research*, **40**(6): 522–531.
- Zhu L, Lei AH, Zheng HY, et al. 2015. Longitudinal analysis reveals characteristically high proportions of bacterial vaginosis-associated bacteria and temporal variability of vaginal microbiota in northern pig-tailed macaques (*Macaca leonina*). *Zoological Research*, **36**(5): 285–298.