Received: 12 February 2019 | Revised: 7 March 2019 | Accepted: 19 March 2019

DOI: 10.1111/1755-0998.13014

INVITED TECHNICAL REVIEW

MOLECULAR ECOLOGY RESOURCES WILEY

A guide to the application of Hill numbers to DNA-based diversity analyses

Antton Alberdi[1] | M. Thomas P. Gilbert[1,2]

[1]Section for Evolutionary Genomics, University of Copenhagen, Copenhagen, Denmark

[2]NTNU University Museum, Trondheim, Norway

Correspondence
Antton Alberdi, Section for Evolutionary Genomics, University of Copenhagen, 1350 Copenhagen, Denmark.
Email: antton.alberdi@snm.ku.dk

Funding information
Lundbeckfonden, Grant/Award Number: R250-2017-1351; ERC Consolidator Grant, Grant/Award Number: 681396

Abstract

With the advent of DNA sequencing-based techniques, the way we detect and measure biodiversity is undergoing a radical shift. There is also an increasing awareness of the need to employ intuitively meaningful diversity measures based on unified statistical frameworks, so that different results can be easily interpreted and compared. This article aimed to serve as a guide to implementing biodiversity assessment using the general statistical framework developed around Hill numbers into the analysis of systems characterized using DNA sequencing-based techniques (e.g., diet, microbiomes and ecosystem biodiversity). Specifically, we discuss (a) the DNA-based approaches for defining the types upon which diversity is measured, (b) how to weight the importance of each type, (c) the differences between abundance-based versus incidence-based approaches, (d) the implementation of phylogenetic information into diversity measurement, (e) hierarchical diversity partitioning, (f) dissimilarity and overlap measurement and (g) how to deal with zero-inflated, insufficient and biased data. All steps are reproduced with real data to also provide step-by-step bash and R scripts to enable straightforward implementation of the explained procedures.

KEYWORDS
beta diversity, biodiversity, dissimilarity coefficients, diversity partitioning, metabarcoding, niche breadth, niche overlap, numbers equivalents, phylodiversity

1 | INTRODUCTION

DNA sequencing-based tools are becoming one of the most popular approaches with which to detect and measure biodiversity in a wide range of natural systems (Bohmann et al., 2014). For example, studies spanning questions from dietary assessments (Alberdi et al., 2019) to microbiome analyses (Peterson et al., 2009) and ecosystem-level biodiversity studies (Bush et al., 2017) have adopted over the last decade molecular approaches based on metabarcoding (Deiner et al., 2017), and increasingly today, shotgun sequencing (Quince, Walker, Simpson, Loman, & Segata, 2017). This methodological shift has changed, or at least complemented, the way we measure biodiversity (Taberlet, Coissac, Pompanon, Brochmann, & Willerslev, 2012). For example, operational units for measuring diversity are no longer necessarily defined through taxonomic means (Blaxter et al., 2005), and it is easier than ever to account for the degree of (dis)similarity between the operational units detected in diversity measurement (Pavoine, Bonsall, Dupaix, Jacob, & Ricotta, 2017).

Researchers often need to quantify how diverse different systems are, for example, to assess ecosystem functioning (Cardinale, Palmer, & Collins, 2002) or to measure any species' niche breadth (Forister et al., 2015). It is also common to compare the composition of different systems, under experimental set-ups to measure the differences yielded by different treatments (e.g., Gevers et al., 2014), or in observational designs to assess whether and how much dietary niches differ (e.g., Kartzinel et al., 2015). A myriad of approaches and tools has been developed over the last century to perform such operations, each embedded within a statistical background, with popular examples including Shannon index (Shannon, 1948), Rao's

quadratic entropy (Rao, 1982), Pianka's niche overlap (Pianka, 1974) and Unifrac distances (Lozupone & Knight, 2005).

Regardless of the specific approach implemented, researchers need to make at least four essential choices when analysing the diversity of a biological system. First, they must define the unit that encompasses biologically alike entities to be used to measure diversity, hereafter referred to as the "type". Although community ecologists have traditionally measured diversity using the taxonomic species as the type (Pielou, 1966), with the implementation of molecular approaches, this is no longer a general rule (Blaxter et al., 2005). Second, it is necessary to consider how detections of these types are treated, either as presence/absence (=incidence) or quantitatively (=abundance), and if the latter, how abundant and rare detections are weighed (Jost, 2006). Third, researchers need to consider whether, and how, the phylogenetic—or ecological—relations between detected organisms will be accounted for when measuring diversity (Chao, Chiu, & Jost, 2014a). Finally, researchers should ideally assess whether the data are representative of the biological system studied, and if needed, take the necessary measures to correct the biases.

Making the right choice requires acknowledgement of the essential concepts and differences between the available approaches, because the final results might depend on the decisions researchers make to measure biodiversity. Given the plethora of metrics available, and since the incorporation of complex attributes like phylogenies increases the level of abstraction, there is an increasing awareness of the need to implement intuitively meaningful diversity measures rather than abstract indices, as the latter can be easily misused and misinterpreted (Jost, 2006; Leinster & Cobbold, 2012; Lucas et al., 2017; Tuomisto, 2010a). In this regard, considerable effort has recently been dedicated to creating frameworks that unify and generalize popular indices (Chao, Chiu, et al., 2014a; Jost, 2006; Tucker et al., 2017). One particularly relevant framework is the so-called "Hill numbers" which encompass the group of diversity measures that quantify diversity in units of equivalent numbers of equally abundant OTUs or species (Hill, 1973).

As explained throughout this article, Hill numbers provide a general statistical framework that is sufficiently robust and flexible to address a wide range of scientific questions that molecular ecologists regularly try to answer through measurement, estimation, partitioning and comparison of diversities (Chao, Chiu, et al., 2014a; Jost, 2006; Tuomisto, 2010a). Hence, this article aims to serve as a conceptual and practical guide for diversity analysis of systems characterized using DNA sequencing-based techniques, taking advantage of the benefits provided by molecular tools and diversity analyses based on Hill numbers. With an eye to facilitating its application by a wide range of researchers, example data files and functions are provided as Supporting information, and the scripts needed to reproduce the analyses explained throughout the article are presented and explained in the Appendix S1.

2 | DEFINING TYPES FOR DIVERSITY QUANTIFICATION

The first step researchers need to make when assessing the biodiversity of a system is to define the types (Krebs, 2014). In community ecology, individuals (i.e., recorded entities) have been traditionally classified into taxonomic species (i.e., types) (Figure 1). Therefore, diversity measurements have commonly been carried out at species level (e.g., species richness and species diversity), principally as determined based on morphological features (MacArthur, 1965; Pielou, 1966). The implementation of DNA-based molecular approaches now enables (in principle) diversity to be measured at a much finer scale—that of DNA sequence variation. Although genetic differentiation might sometimes be decoupled from ecological differences (Pavoine et al., 2017), multiple reasons render it an appropriate element with which to characterize biological diversity. First, there is a strong link between phylogeny and variation in morphological and functional traits (Felsenstein, 1985; Nipperess, Faith, & Barton, 2010). Second, genetic variation is the basis upon which evolution generates biological diversity (Mayr & Provine, 1998). Third, it confers the possibility to objectively and systematically define operational units (Blaxter et al., 2005), overcoming problems relating to subjective perception of morphological traits (Hey, Waples, Arnold, Butlin, & Harrison, 2003). Fourth, such methods increase the efficiency of the analysis of highly diverse systems, because laboratory procedures can be roboticized and bioinformatic steps automated (Alberdi et al., 2019). For all these reasons, molecularly defined types, broadly known as OTUs

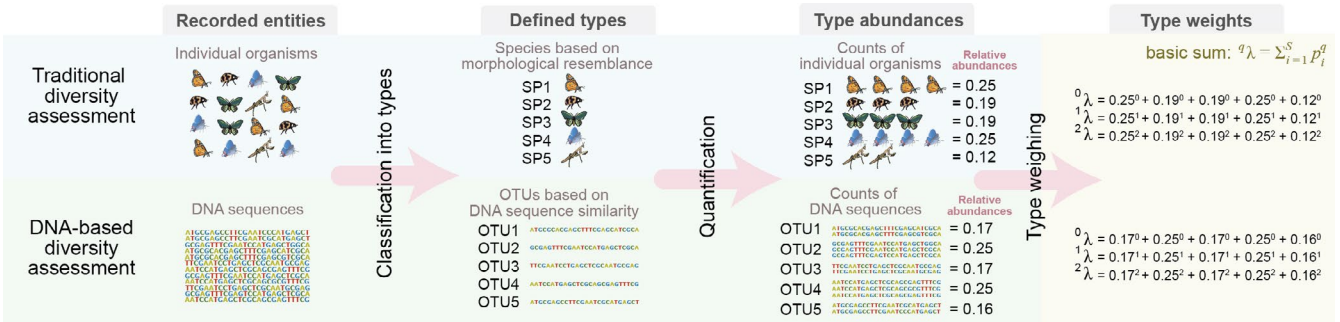

FIGURE 1 Diversity assessment procedures in traditional and DNA sequencing-based approaches. Recorded entities need to be classified into types, before each type is weighed according to its relative abundance and the order of diversity (q). Note the example refers to an abundance-based, rather than incidence-based, approach

or MOTUs (molecular operational taxonomic units, Blaxter et al., 2005), are becoming the preferred types with which to quantify diversity in many fields of the biological sciences.

When using molecular approaches, the recorded entities are no longer individuals, but DNA sequences, and the classification into types is not any longer based on morphological features, but the level of dissimilarity between DNA sequences. Currently, metabarcoding is the most popular approach with which to characterize the diversity in samples containing DNA derived from complex communities (Deiner et al., 2017). Using this method, the amplification and subsequent sequencing of short (usually between 50 and 500 nucleotides) targeted molecular markers yields multiple presumably homologous sequences that are directly comparable to each other (but see section on zero-inflated, insufficient and biased data). Although each of the different sequences (recorded entities) could be considered a different type (Callahan et al., 2016), usually similar sequences are clustered according to user-defined similarity thresholds to generate OTUs that encompass sequences within a certain degree of similarity (Caporaso et al., 2010; Schloss et al., 2009). One of the aims of clustering is to adjust the defined types to approximately represent biological species, so as to yield diversity measurements that resemble traditional species level analyses (Hebert, Cywinska, Ball, & deWaard, 2003). Typically, an identity value of 97% has been suggested as a standard threshold for considering the average genetic dissimilarity level across species (Caporaso et al., 2010; Schloss et al., 2009), although clearly this does change across taxa and genetic marker (Alberdi, Aizpurua, Gilbert, & Bohmann, 2018; Kim, Oh, Park, & Chun, 2014). While clustering is useful for analyses in which intraspecific diversity is largely irrelevant (such as diet assessments), it entails a loss of potentially useful information for most biodiversity analyses (Pavoine & Izsák, 2014), as fine-scale variation can differentiate pathogenic from commensal bacterial strains (McElroy, Zagordi, Bull, Luciani, & Beerenwinkel, 2013) or be informative of population structure (Rosen, Davison, Bhaya, & Fisher, 2015), for instance.

With the implementation of molecular approaches, the definition of types becomes more objective, as types are defined based on mathematical criteria applied to simple DNA sequences, rather than assessments of complex morphological traits. However, as types can be defined using different algorithms and parameters (Rideout et al., 2014), and the employed approach has a large impact on the results (Alberdi et al., 2018), authors need to provide detailed information (ideally full codes) about the approaches and criteria used to define types in order for studies to be reproducible. If authors publish the raw sequence data, others can choose different criteria to define types and re-analyse the diversity of a system to ensure comparability.

3 | WEIGHING THE IMPORTANCE OF TYPES

Diversity measurements require assignment of an importance value to each of the detected types. In traditional community ecology, this has been done using metrics such as individual counts, biomass or spatial units, depending on the type of system, research question and fieldwork strategy. Molecular analyses provide a different type of data that could provide such information, namely the amount of DNA sequences assigned to each OTU (Deagle et al., 2019).

There are multiple approaches that enable differential weighing of abundant and rare OTUs. The simplest measure of diversity is OTU richness (McIntosh, 1967). As this only considers whether an OTU is present or absent in the system, abundant and rare OTUs are given the same weight. However, the multiple OTUs present in a system are seldom distributed evenly; thus, richness is rarely the best approach with which to reflect the diversity of a system. Consider for instance, a simple system characterized with 1,000 sequence reads, in which 990 belong to OTU1 and 10 to OTU2. This would yield a richness value of 2, even though the system is overwhelmingly dominated by OTU1. Thus, metrics such as the Shannon or the Simpson indices, which also account for the evenness of the system, are considered more representative of the diversity of a system. It is critical to note, however, that unlike richness, neither the Shannon index nor the Simpson index are actual measures of diversity. The former measures entropy thus yields the uncertainty in the OTU identity of a randomly chosen sequence in the system. The latter provides the probability that two randomly chosen DNA sequences actually belong to different OTUs (Chao, Chiu, et al., 2014a). Consequently, the values that Shannon and Simpson indices yield are difficult to interpret—the values in the previous example are 0.056 and 0.020, respectively, and do not exhibit the intuitive properties ecologists expect from a diversity measurement.

Specifically, our intuitive notion of diversity would expect that when doubling the number of OTUs in a system, then the diversity measured should also double. This is known as the "doubling property" or "replication principle" (Chao, Chiu, & Jost, 2010; Hill, 1973; Jost, 2006). For example, if the diet of one bat species is comprised of 15 moth species with even abundances, and the diet of another species encompasses 30 moths also with even abundances, intuitively we would conclude that the second bat's diet is twice as diverse (100% more diverse) as the first one. However, most diversity indices lack this basic property. The Shannon entropy only increases from 2.7 (15 species) to 3.4 (30 species), which might suggest a diversity gain of 26%, and the Simpson index only increases from 0.93 to 0.96, which might suggest a gain of barely 3%. Hence, treating diversity indices as diversity values has noticeable practical consequences, as they all vary in range and behaviour (Jost, 2006).

Fortunately, richness, Shannon index and Simpson index belong to a single statistical framework, as they all are monotonic functions of the *basic sum* ${}^{q}\lambda = \sum_{i=1}^{S} p_i^q$, that is, the sum of the relative abundances of the types (p_i) elevated to the q value (Jost, 2006; Keylock, 2005). This implies that Hill numbers (${}^{q}D$), or actual diversities, rather than entropies or probabilities, can be formulated in terms of the *basic sum* (${}^{q}\lambda$) and the parameter q (R scripts in Appendix S1):

$$ {}^{q}D = \left(\sum_{i=1}^{S} p_i^q\right)^{1/(1-q)} = \left({}^{q}\lambda\right)^{1/(1-q)}, q \neq 1 \quad (1) $$

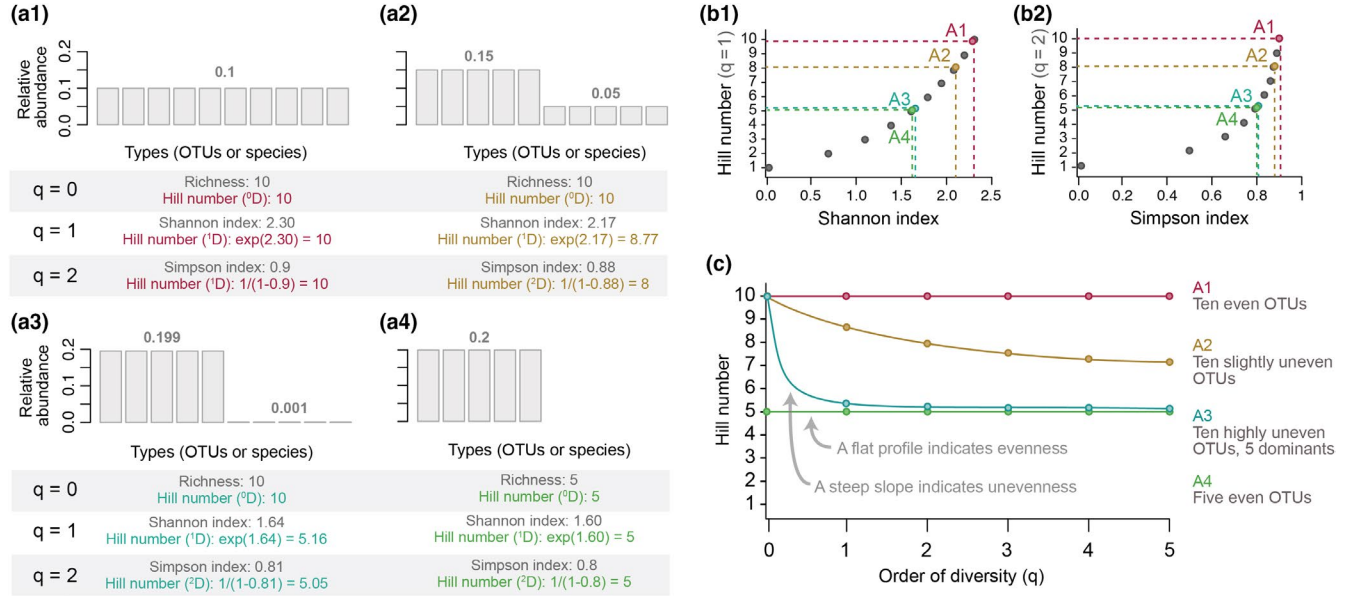

FIGURE 2 Relation between system composition, raw diversity indices and Hill numbers, and diversity profiles. (a) The raw diversity index values for q values 0, 1 and 2, and their respective Hill numbers of four theoretical systems with different abundance distributions. In a1 and a4, all types are evenly distributed; thus, the Hill numbers of all orders of diversity (q) are identical. In a2, as the relative abundances of the types differ, the Hill numbers of order q > 0 are lower than the value for q = 0, which overlooks relative abundances. In a3, there are five very abundant types and five very rare types. Consequently, the Hill numbers of order q > 0 are much lower than the value for q = 0, approximate the number of abundant types, and better reflect the abundance distribution of the system. Note that in a2 and a3, the Hill numbers decrease as the importance of abundant types is increased with higher q values. (b) The relation between the raw Shannon (b1) and Simpson (b2) index values and their respective Hill numbers. The position of the four theoretical systems (a1–4) is shown in both cases. (c) Diversity profiles that show the different components of the diversity (i.e., number of OTUs and their evenness) of the four theoretical systems at a glance

This expression was first discovered by Hill (1973), hence the use of the name "Hill numbers" to refer to the output of this formula. Hill numbers have two major advantages over diversity indices: (a) the interpretation of the measure and its measurement unit is always the same (Chao, Chiu, et al., 2014a; Tuomisto, 2010a), and ii) the sensitivity towards abundant and rare OTUs can be modulated with the parameter q. The expression yields a diversity measure in "effective number of OTUs", that is, the number of equally abundant OTUs that would be needed to give the same value of diversity (Hill, 1973; Jost, 2006). When all OTUs in a system have the same relative abundances, as in the moth example given above, the effective number of OTUs for all q values equals the actual number of OTUs, namely richness (Figure 2a1). When the relative abundances of the types vary however, then the effective number of OTUs for q > 0 values decreases (Figure 2a2). The higher the heterogeneity between types, the lower the effective number of OTUs. In extreme cases in which the system is dominated by a few equally abundant OTUs, the effective number of OTUs will approach the number of those abundant OTUs (Figure 2a3).

The sensitivity towards abundant and rare OTUs can be modulated using the scaling parameter q, known as the "order" of diversity (Jost, 2006). The larger the q value, the higher the importance attributed to abundant OTUs. Three q values are particularly relevant, both for their significance, and their close relationship to popular diversity indices: q = 0, q = 1 and q = 2. When a diversity of order zero (q = 0) is applied to the formula, it becomes insensitive to OTU frequencies, thus yielding a richness value. As the relative abundances of OTUs are overlooked, rare OTUs are overweighed. A q value of 1 (in practical terms its limit, as the Hill number is undefined for q = 1) is the value that weighs OTUs by their frequency, without disproportionately favouring either rare or abundant ones (Jost, 2006). The value it yields is exactly the exponential of the Shannon index. In fact, q values under unity favour rare OTUs, while values above one favour abundant OTUs (Keylock, 2005). When a q value of 2 is applied, abundant OTUs are overweighed, and the formula yields the multiplicative inverse of the Simpson index. Indeed, common diversity indices can be transformed to Hill numbers (also known as numbers equivalents or true diversities sensu Jost, 2006), by applying simple mathematical transformations (Table 1).

The numbers equivalents of the Shannon and Simpson indices in the 2-OTU example given above (OTU1 = 990 sequences; OTU2 = 10 sequences) would therefore be exp(0.056) = 1.05 and 1/(1–0.020) = 1.02, respectively. These are the effective number of OTUs for orders q = 1 and q = 2, respectively. The fact that these effective OTU values are close to 1 demonstrates that the system is dominated by a single OTU (in this example 99% of the reads belong to OTU1). The higher the q value, the lower the diversity value, because the importance attributed to abundant OTUs increases. The relation

TABLE 1 The three most common diversity and phylodiversity indices and the transformations to be carried out to obtain (phylogenetic) Hill numbers of orders 0, 1 and 2. p_i refers to the relative abundance of OTU i, B_T refers to the tree depth (time interval if the tree is time-calibrated) considered, L_i refers to the length of branch i, and a_i refers to the sum of relative abundances of OTUs descended from branch *i*. Note that in the literature the diversity indices ${}^0H$, ${}^1H$ and ${}^2H$ are often represented as *S*, *H* and *D*, respectively. R scripts to execute all these equations are shown in Appendix S1

	q	Diversity index ${}^qH$	Hill number in terms of ${}^qH$	Hill number ${}^qD$
OTU diversity	0	Richness $^0H \equiv \sum_{i=1}^{S} p_i^0$	Richness $^0D = {}^0H$	$^0D = \left(\sum_{i=1}^{S} p_i^0\right)^{1/(1-0)}$
	1	Shannon entropy $^1H \equiv -\sum_{i=1}^{S} p_i^1 lnp_i^1$	Shannon diversity $^1D = exp\left({}^1H\right)$	$^1D = \left(\sum_{i=1}^{S} p_i \log p_i\right)$
	2	Simpson index $^2H \equiv 1 - \sum_{i=1}^{S} p_i^2$	Simpson diversity $^2D = 1/\left(1 - {}^2H\right)$	$^2D = \left(\sum_{i=1}^{S} p_i^2\right)^{1/(1-2)}$
Lineage diversity	0	Faith's PD $PD \equiv \sum_{i=1}^{S} L_i$	$^0\bar{D}\left(T\right) = PD/T$	$^0\bar{D}\left(T\right) = \left(\sum_{i \in B_T} \frac{L_i}{T} a_i^0\right)^{1/(1-0)}$
	1	Allen's H_P $H_P = -\sum_i L_i a_i log a_i$	$^1\bar{D}\left(T\right) = exp\left(H_P/T\right)$	$^1\bar{D}(T) = exp\left[-\sum_{i \in B_T} \frac{L_i}{T} a_i \log a_i\right]$
	2	Rao's Q $Q = \sum_{i,j} d_{i,j} p_i p_j$	$^2\bar{D}(T) = 1/\left(1 - Q/T\right)$	$^2\bar{D}(T) = \left(\sum_{i \in B_T} \frac{L_i}{T} a_i^2\right)^{1/(1-2)}$

between raw indices and diversities in more complex systems with different degree of heterogeneity is shown in Figure 2.

The ability to modulate the sensitivity towards abundant and rare OTUs by modifying a single parameter (*q*) is a useful means with which to adjust diversity measurements to the type of data and research question. For example, when rare types are considered to be of low importance (e.g., when attempting to define a core diet or microbiome), or when rare types are considered untrustworthy due to technical issues (e.g., PCR or sequencing errors), researchers might opt for using a high *q* value, for example, *q* = 2, which overweighs abundant OTUs. The result can be interpreted as the effective number of dominant OTUs in the system (Chao, Chiu, et al., 2014a). In contrast, if rare types are considered essential for the system, or researchers do not trust the relative abundance data due to potential technical biases, researchers might opt for using a *q* value of 0 that simply counts the number of types.

Hill numbers also enable *diversity profiles* of systems and subsystems to be plotted as continuous functions of the parameter *q* (Figure 2c). This is useful to characterize the OTU abundance distribution of a system, as different compositions and abundance distributions can yield the same value for a particular order of diversity (e.g., *q* = 1), but not for many of them (e.g., *q* = 0, *q* = 0.5 and *q* = 1). Hill numbers convey all information contained in a species abundance distribution at a glance (Chao, Chiu, et al., 2014a; Leinster & Cobbold, 2012).

4 | ABUNDANCE-BASED VERSUS INCIDENCE-BASED APPROACHES

Although the Hill number framework was originally developed to deal with relative abundance data (i.e., relative number of sequences assigned to each OTU), it has recently also been applied to incidence data. This means that the relative abundances of the types detected in each of the subsystems (=samples) belonging to a certain system are overlooked, and the diversity of the system is calculated by computing the relative number of detections of a given type across the whole system (Figure 3). Although incidence data are less informative than abundance data, it is both easier to collect, more comparable, and has been extensively used under the niche theory framework (Box 1). When dealing with DNA-derived data, incidence-based approaches have particular relevance, given the limited quantitative relationship that exists between the biomass in the actual system and the DNA sequences produced (Lamb et al., 2019), which might challenge the representativeness of abundance data. However, consensus has not been reached within the molecular ecology research community about which approach is the most appropriate, as simulations have shown that analyses based on incidence data often overestimate the importance of rare OTUs, and abundance data might provide a more accurate view of the diversity even with moderate recovery biases (Deagle et al., 2019).

While either approach might be valid depending on the research question and the study design, it is important to acknowledge the basic differences between abundance-based and incidence-based Hill numbers. In the abundance-based approach, the unit used to compute diversity is the count of DNA sequences assigned to each OTU. In contrast, in the incidence-based approach, the count of subsystems in which an OTU is present is used to compute diversity. Thus, abundance-based Hill numbers with different orders of diversity can be computed for both subsystems (=samples) and entire systems (see section about diversity partitioning for details), whereas incidence-based *q* > 0 Hill numbers are only meaningful for entire systems (=pool of samples). The diversity of a system

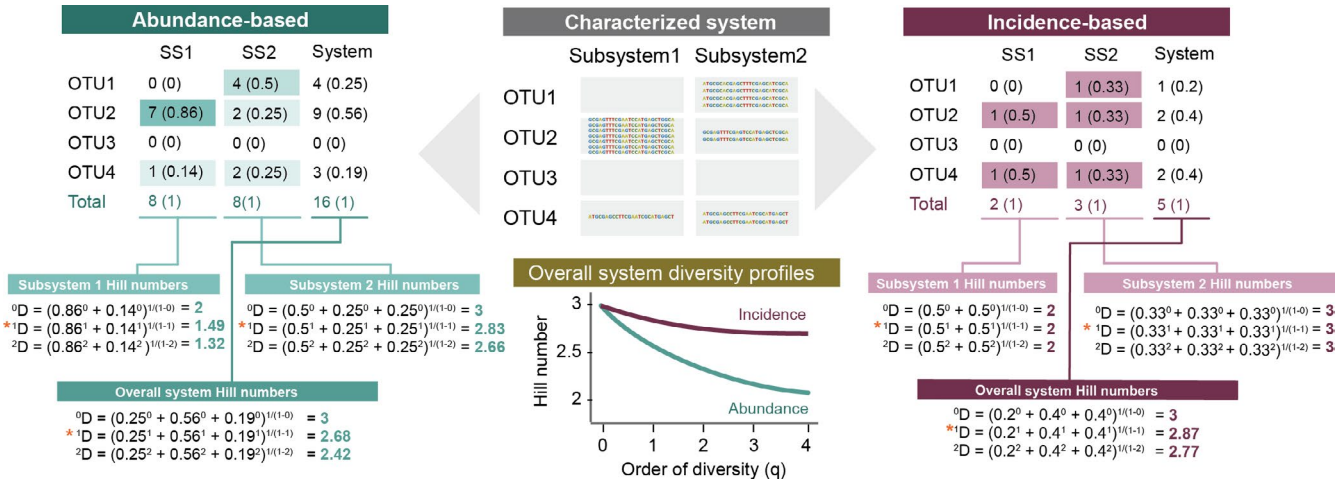

FIGURE 3 Differences between abundance-based and incidence-based Hill numbers. The Hill numbers yielded for the entire system are different depending on the approach employed. In abundance-based approaches, the DNA sequence is the unit that the diversity is computed on, while in incidence-based approaches, it is the sample the unit upon which the diversity is measured. (*) The asterisk indicates that the equations are undefined for $q = 1$, thus in practice either the ^1D formula shown in Table 1 or a limit of the unity must be used, for example, $q = 0.9999$. However, $q = 1$ is used for the sake of simplicity

Box 1 Incidence-based diversity analysis under the ecological niche framework

Dietary niche breadth and niche overlap are two concepts at the heart of the ecological niche theory. When diversity is measured within the framework of niche theory, it takes the name of niche breadth, niche width or niche size (Colwell & Futuyma, 1971), while the degree of similarity between two (or more) contrasting niches is known as niche overlap (Hurlbert, 1978). Surprisingly, Hill numbers are not explicitly used in the niche theory literature, although many approaches to measure niche breadth and overlap are closely related. Most statistical approaches employed within the niche framework rely on incidence, rather than abundance, data. In fact, one of the most popular indices to measure niche breadth, namely the Levins' index (Levins, 1968), is the incidence-based Hill number of order of diversity $q = 2$. Hence, the implementation of incidence-based Hill numbers in niche breadth characterization seems completely natural and useful, as it enables characterization of the observed diversity in a more complete and straightforward way than when using abstract indices, as explained throughout the main text. Naturally, the Hill number framework can also be used to compute niche overlap between two or more systems (e.g., predator species), and even explore phylogenetic relations between types (e.g., prey OTUs), by implementing the diversity partitioning approach explained in the main text, and deriving similarity from beta diversity values using the formulae shown in Table 2.

measured using abundance and incidence approaches converge when $q = 0$, but they yield different diversity values when $q > 0$ (Figure 3). This happens because the two approaches measure different properties. Abundance-based Hill numbers measure the effective number of equally abundant OTUs in the system, while incidence-based Hill numbers measure the effective number of equally frequent (across subsystems) OTUs in the system.

5 | ACCOUNTING FOR RELATEDNESS BETWEEN TYPES

The traditional and still most broadly used diversity indices, as well as their corresponding Hill numbers, assume that all types are equally (dis)similar; that is, types are treated as discrete elements with no correlation among them. Such measures have been coined as *neutral* or *species-neutral* measures of diversity (Chao et al., 2010; Marcon & Hérault, 2015). This implies that, for example, a system (e.g., dietary sample) comprised of four *Noctua* moths (Insecta, Lepidoptera) with even abundances has the same diversity as another system evenly composed of two *Noctua* moths and two *Carabus* beetles (Insecta, Coleoptera). Obviously, in ecological terms, a system holding moths and beetles is more heterogeneous than a system composed solely of moths. Although ecologists realized this pitfall long before molecular tools were implemented into the study of community ecology (Pielou, 1975), it was only in the last decade that a general framework of *non-neutral* measures of diversities was developed, based on the trait-based or phylogenetic relations between detected OTUs (Chao, Chiu, et al., 2014a). In this article, we will focus on the latter due to its more straightforward applicability to molecularly characterized systems.

Each principal diversity index (e.g., richness, Shannon index and Simpson index) has an equivalent phylodiversity index (Chao et al.,

TABLE 2 Diversity and phylodiversity partitioning into α, β and γ components based on Hill numbers, and four similarity measures derived from the resulting β-diversities. In the diversity partitioning equations, *N* refers to the number of subsystems (samples), *j* refers to each subsystem, *S* refers to the number of OTUs, *i* refers to each OTU, w_j refers to the relative weight of subsystem j (in case of even weighs it is 1/*N*), p_i refers to the relative abundance of OTU *i*, *T* refers to the considered tree depth which for ultrametric trees is reduced to tree length, B_T refers to the considered tree depth (time interval if the tree is time-calibrated) considered, L_i refers to the length of branch i, and a_i refers to the sum of relative abundances of OTUs descended from branch *i*. Note that to calculate Hill number of order 1, the actual order of diversity to be used needs to be a limit of the unity rather than the number one itself, or a different equation must be used as shown elsewhere (Chao et al., 2010; Jost, 2006). R scripts to execute all these equations are shown in Appendix S1

Diversity partitioning		
Alpha	$^qD_\alpha = \frac{1}{N}\left\{\sum_{i=1}^{S}\sum_{j=1}^{N}\left(w_j p_{ij}\right)^q\right\}^{1/(1-q)}$	Jost (2007), Chao et al. (2012); Chiu et al. (2014)
Gamma	$^qD_\gamma = \left\{\sum_{i=1}^{S}\left(\sum_{j=1}^{N} w_j p_{ij}\right)^q\right\}^{1/(1-q)}$	
Beta	$^qD_\beta = {}^qD_\gamma / {}^qD_\alpha$	
Phylodiversity partitioning		
Alpha	$^q\bar{D}_\alpha(T) = \frac{1}{TN}\left[\sum_{i\in B_T} L_i \sum_{j=1}^{N}\left(\frac{w_j a_{ij}}{T}\right)^q\right]^{1/(1-q)}$	Chiu et al. (2014), Chao et al. (2016)
Gamma	$^q\bar{D}_\gamma(T) = \frac{1}{T}\left[\sum_{i\in B_T} L_i \left(\frac{\sum_{j=1}^{N} w_j a_{ij}}{T}\right)^q\right]^{1/(1-q)}$	
Beta	$^q\bar{D}_\beta(T) = {}^q\bar{D}_\gamma(T) / {}^q\bar{D}_\alpha(T)$	
Similarity measurement derived from β-diversities based on Hill numbers (for both diversity and phylodiversity). Dissimilarity measures are their one-complements ($1 - X_{qN}$)		
Sørensen-type overlap	$C_{qN} = \frac{\left[\left(1/{}^qD_\beta\right)^{q-1} - (1/N)^{q-1}\right]}{\left[1-(1/N)^{q-1}\right]}$	MacArthur (1965), Harrison et al. (1992), Jost (2006, 2007), Chiu et al. (2014), Chao et al. (2016)
Jaccard-type overlap	$U_{qN} = \frac{\left[\left(1/{}^qD_\beta\right)^{1-q} - (1/N)^{1-q}\right]}{\left[1-(1/N)^{1-q}\right]}$	
Sørensen-type turnover-complement	$V_{qN} = \frac{(N - {}^qD_\beta)}{(N-1)}$	
Jaccard-type turnover-complement	$S_{qN} = \frac{(1/{}^qD_\beta - 1/N)}{(1-1/N)}$	

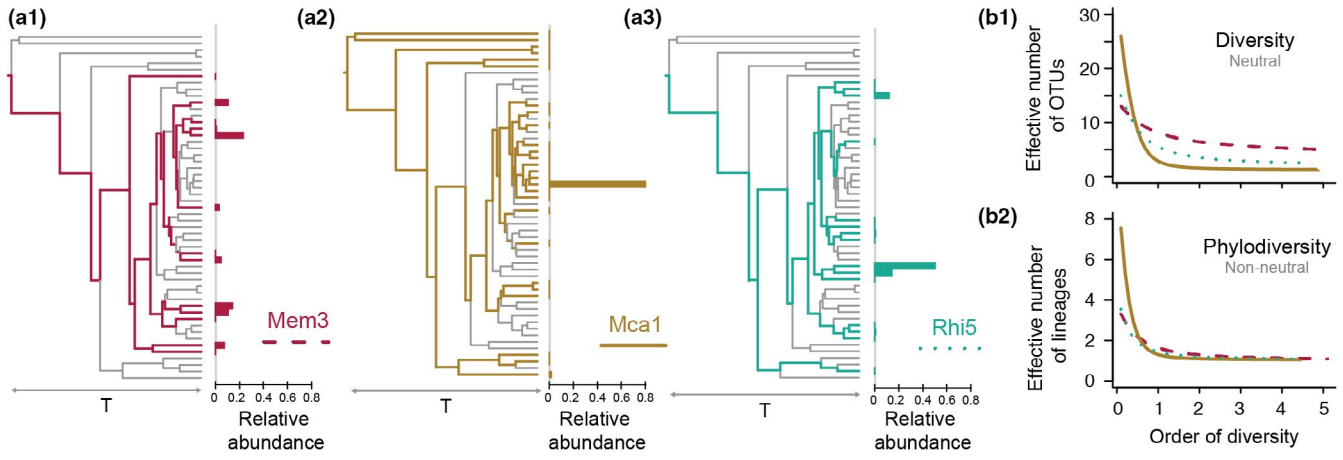

FIGURE 4 Phylogenetic trees and relative OTU abundances of samples and their corresponding diversity and phylodiversity profiles. (a) Relative abundances and phylogenetic relationships of the OTUs detected in three bat faecal samples. Phylogenetic trees include all the OTUs detected in the three samples, and the lineages detected in each sample are coloured and bolded. (b) Diversity and phylodiversity profiles that describe the three samples as a function of the order of diversity (*q*). Mca1 is the sample with the highest richness, although the dominance (relative abundance of 0.80) of one of the OTUs drops the diversity values when $q > 0$. In contrast, Mem3 contains less OTUs, but their even distribution yields higher $q > 0$ diversity values

2010; Pavoine, Love, & Bonsall, 2009). These are Faith's PD (Faith, 1992), Allen's Hp (Allen, Kon, & Bar-Yam, 2009) and Rao's Q (Rao, 1982), respectively (Table 1). Similar to Jost's (2006) work with diversity indices, Chao et al. (2010) unified the different phylodiversity indices around the Hill numbers, yielding the so-called phylogenetic Hill numbers. Within this framework, the units of diversity are neither OTUs nor species, but rather branch segments or lineages as defined by a phylogenetic tree (Figure 4a). In fact, the phylodiversity measure incorporates three types of information: the tree's branching pattern, the relative branch lengths and the relative abundances of each node/branch (see Chao et al., 2010 for a detailed explanation). It is formulated as follows:

$$ {}^{q}\bar{D}(T) = \left(\sum_{i \in B_T} \frac{L_i}{T} a_i^q \right)^{\frac{1}{1-q}}, q \neq 1 \quad (2) $$

In this equation, B_T is the set of all branches or the ones within the depth (T) considered for the analyses in the phylogenetic tree, L_i is the length of the branch i, and a_i is the sum of relative abundances of all OTUs descended from branch i. Hence, the abundance-based phylogenetic Hill numbers are measured in units of the effective number of equally abundant and equally distinct lineages (Chao et al., 2010; Chao, Gotelli, et al., 2014b). For two systems with identical number of types and relative abundances, the one with the deepest branches (largest phylogenetic differences across types) will be the one with the highest phylodiversity. Besides, the phylodiversity value will always be lower than its related OTU diversity (Chao et al., 2010), except in the imaginary case in which the phylogenetic tree is star-shaped, that is, with all equal-size branches radiating from the root, in that case the phylodiversity value is equal to the diversity value. It must be noted that ${}^{q}\bar{D}(T)$, which yields the effective number of lineages, expresses a generalized mean of Hill numbers rather than a genuine Hill number (Chao, Chiu, & Jost, 2016). Thus, it is more accurate to refer to it as the mean phylogenetic diversity (Chao et al., 2010). In order to obtain the total phylogenetic diversity, ${}^{q}\bar{D}(T)$ needs to be multiplied by tree depth (T), which if specified in years, yields a Hill number of order q during the time interval from T years ago to the present, that is, the effective number of lineage-years (Chao et al., 2010). Similar to the diversity profiles mentioned before (Figure 2b,c), it is also possible to plot the phylogenetic Hill numbers as a function of T (Figure S1), which indicates the distribution of the phylogenetic diversity across the OTU tree (Chao et al., 2010). If the OTU tree exhibits very recent radiation, the phylogenetic diversity will decrease rapidly (Figure S1a), while if the phylogenetic tree shows a deep branching pattern, the mean phylogenetic diversity will remain high until approaching the root of the three (Figure S1b).

The phylogenetic Hill numbers are based on trees (usually ultrametric) in which the relations between OTUs are established. A major advantage of metabarcoding is that phylogenetic trees can be directly built using the homologous DNA sequences or OTU representative sequences generated. This could liberate researchers from having to assign taxonomy to the DNA sequences in order to incorporate phylogenetic data to the diversity measurement and enables piping commands to automatize the process of the analysis of diversity, as shown in the Appendix S1. However, building phylogenies from metabarcoding data have its limitations, as discussed in Box 2.

6 | PARTITIONING DIVERSITY INTO α, β AND γ COMPONENTS

In ecology, the idea of diversity has been traditionally broken down into three components: alpha (α), beta (β) and gamma (γ) diversities (Whittaker, 1960). In general terms, α-diversity refers to the average diversity of subsystems or samples (although see discussion about the different α-diversities in Chao, Chiu, & Hsieh, 2012), β-diversity measures the differences between subsystems (although see discussion about the different β-diversities in Tuomisto, 2010a), while γ-diversity includes the entire diversity of the system (Figure 5). Despite the existence of different approaches for diversity partitioning, within the framework of Hill number diversity partitioning responds to a multiplicative definition ${}^{q}D_{\gamma} = {}^{q}D_{\alpha} \times {}^{q}D_{\beta}$ (Chao et al., 2012; Jost, 2007); that is, beta diversity is obtained by dividing gamma diversity by alpha diversity. This formulation has three properties that ecologists intuitively expected from a diversity measurement: (a) alpha and beta components are unrelated; thus, a high value of alpha does not force the beta component to be high and vice versa, (b) gamma is completely determined by alpha and beta, and (c) alpha is never greater than gamma. The alpha, gamma and beta components based on Hill numbers of order q are obtained using the equations shown in Table 2 (Chao et al., 2012; Jost, 2006), which despite their apparent complexity, can be easily computed using the R functions shown in Appendix S1. Alpha, beta and gamma diversities can also be computed including phylogenetic information (Chiu, Jost, & Chao, 2014) and can be used to decompose biological systems with multiple hierarchical levels (Gaggiotti et al., 2018).

It must be highlighted though that the alpha diversity is not obtained by averaging the Hill numbers of the subsystems, but computing the Hill numbers from the averaged basic sums of the subsystems (Chao et al., 2012). In contrast, gamma diversity is obtained by taking the average of OTU relative abundances across subsystems and then computing the Hill numbers of the pooled system. Although beta diversity is often used to vaguely refer to any kind of compositional heterogeneity among systems (Barwell, Isaac, & Kunin, 2015; Chao, Chiu, et al., 2014a; Tuomisto, 2010a,2010b), when diversity partitioning is carried out using Hill numbers, beta diversity is an actual diversity value that measures the effective number of equally large and completely distinct subsystems in a system. However, the Hill number beta diversity can also be interpreted as a unitless scalar that quantifies the ratio of diversities between two levels (alpha and gamma) of observation; thus, it also quantifies how many times richer an entire system is in effective OTUs (gamma diversity) than its constituent subsystems are on average (alpha diversity) (Figure 5).

Box 2 Obtaining reliable phylodiversity measurements from metabarcoding data

Metabarcoding enables phylogenetic diversities to be computed directly from OTU representative sequences. However, at least three important issues must be considered so as to obtain reliable phylogenetic diversity values. First, some OTUs might not be of interest for diversity measurement of, for instance, microbiomes or diets. Even when primers are taxon-specific, non-targeted taxa might also be amplified and sequenced (Alberdi et al., 2018). On the one hand, there are organisms that usually appear at low abundance; examples could be eukaryotic intestinal parasites when characterizing microbiomes from gut contents, or skin acari when studying diet from faeces. These organisms do not usually introduce large distortion to $q > 0$ diversity measurements, as their relative abundances tend to be low, but if they are not detected and excluded, they can considerably inflate the phylodiversity measures, given that they tend to be distantly related to the intended targets. On the other hand, some metabarcoding primers might amplify and sequence host or predator DNA from which samples have been acquired, which in certain cases can account for the majority of the sequence reads (Alberdi et al., 2018; Galan et al., 2018), thus completely distorting diversity and phylodiversity measurements in analyses with $q > 0$ values. Hence, we strongly advise taxonomic assignment to the OTU sequences, and application of a relaxed filter based on low similarity values (e.g., 90%) to reference sequences, so as to exclude OTU sequences assigned to non-targeted taxonomic groups.

Another major issue is the low robustness of phylogenetic trees generated from metabarcoding sequences. The combination of short DNA sequences (often < 200 bp) and abundant OTUs (often > 1,000), as usually generated in metabarcoding studies, yields phylogenetic trees with very low node support, indicating high phylogenetic uncertainty (Douady, Delsuc, Boucher, Doolittle, & Douzery, 2003). Phylogenetic uncertainty means that multiple trees have shared probabilities of reflecting the evolutionary history of the organisms. As each of these trees might yield different phylodiversity values (Figure S3a,b), it is advisable to implement approaches that account for the uncertainty of the phylogenetic reconstructions used for diversity measurement. This enables probability distributions to be generated for the different orders of diversity (Figure S3c), as well as plotting of diversity profiles with confidence intervals (Figure S3d).

One option to account for phylogenetic uncertainty is to rely on Bayesian inference approaches (e.g., MrBayes, BEAST), in which node support is based on a posterior probability distribution of trees. In a Bayesian analysis, a Markov chain Monte Carlo (MCMC) usually with > 1–10 million steps begins from an initial tree (usually built using simple neighbour-joining or parsimony methods) and moves through the parameter space searching for high-probability regions of the posterior. Trees are sampled at fixed intervals, and the posterior probability of a given tree is approximated by the proportion of time that the chains visit it (Nascimento, Reis, & Yang, 2017). In the long run, a successful Markov chain should reach a stationary distribution (Figure S4). Software such as MrBayes or BEAST output all the trees samples across the MCMC chain, usually every 1,000 steps. Hence, as explained in Appendix S1, it is possible to skip the trees sampled before the MCMC reached the stationary phase (known as burn-in), randomly sample a subset of trees from the resulting MCMC chain, compute phylodiversity measurements based on different randomly samples trees and obtain the average and standard error value that indicates phylogenetic uncertainty (Figure S3d). If the MCMC has reached a stationary distribution with low variance, the sampled trees will be similar to each other, so they will be the phylodiversity results. In contrast, if the MCMC chain has not reached a stationary distribution and/or the variance is high, the sampled trees will be different to each other, and hence, the variance of the phylodiversity results will also increase.

Finally, it is essential to acknowledge that different markers might yield different phylogenetic trees, more so if their power for phylogenetic inference is low, as is the case for most markers employed for metabarcoding. This means that an identical community characterized with different molecular markers, even in the ideal (yet likely unrealistic) case without PCR biases, could yield different phylodiversity values. It is therefore advisable to avoid comparisons between the diversities of systems characterized using different molecular markers.

The Hill number beta diversity always ranges from 1 (when all subsystems are identical) to the actual number of subsystems (when all subsystems are completely different) (Chao, Chiu, et al., 2014a; Chiu et al., 2014).

7 | MEASURING DISSIMILARITY

Researchers often need to quantify the (dis)similarity between subsystems (e.g., dietary overlap between species, microbiome differences between intestinal sections and community differences between habitats). Computing (dis)similarities is also a necessary step prior to popular statistical methods such as NMDS or ANOSIM. Dissimilarity indices range between 0 and 1; 0 indicates that the subsystems compared are identical, while 1 indicates that they are completely different. As the beta diversity lies in between 1 and the total number of subsystems, the Hill number beta diversity cannot directly be used to compute dissimilarities. However, it is possible—and desirable—to remove the dependence on the number of subsystems and compute dissimilarity measures by applying simple transformations to beta diversity, both for diversities (Chao et al., 2012; Jost, 2007) as well as phylodiversities (Chao, Chiu, et al., 2014a; Chiu et al., 2014). Four classes of similarity measures derived from Hill number beta diversities have

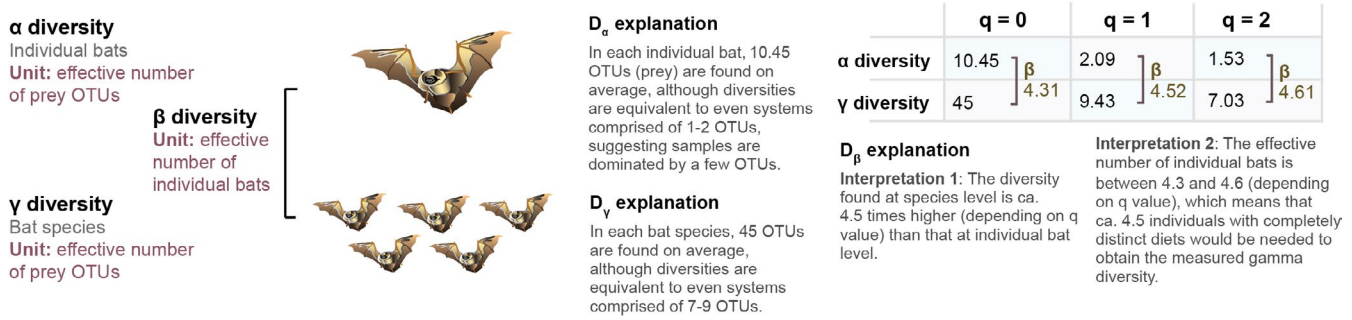

FIGURE 5 Explanatory example of hierarchical diversity partitioning. The system is comprised of 40 individual bats; thus, it has two hierarchical levels: individual bats and overall community

been proposed (Table 2), from which dissimilarity measures can be obtained by calculating their one-complements (1-X_{qN}). The Sørensen-type classes quantify similarity from the perspective of the subsystem, while the Jaccard-type classes quantify similarity from the perspective of the overall system (Chao et al., 2019; Chiu et al., 2014).

1. The Sørensen-type overlap (C_{qN} for diversity/$\bar{C}_{qN}$ for phylodiversity) quantifies the effective average proportion of a subsystem's OTUs (or lineages in the case of phylodiversities) that is shared across all subsystems. This is thus a metric that quantifies overlap from the subsystem's perspective. Its corresponding dissimilarity measure (1 − C_{qN}) quantifies the effective average proportion of nonshared OTUs or lineages in a system (Chao et al., 2012; Chao, Jost, Chiang, Jiang, & Chazdon, 2008).
2. The Jaccard-type overlap ($U_{qN}/\bar{U}_{qN}$) quantifies the effective proportion of OTUs or lineages in a system that are shared across all subsystems. Hence, this metric quantifies overlap from the perspective of the overall system. Its corresponding dissimilarity (1 − U_{qN}) quantifies the effective proportion of nonshared OTUs or lineages in the overall system.
3. The Sørensen-type turnover-complement ($V_{qN}/\bar{V}_{qN}$) is the complement of the Sørensen-type turnover, which quantifies the normalized OTU turnover rate with respect to the average subsystem (i.e., alpha), thus provides the proportion of a typical subsystem that changes across subsystems (Harrison, Ross, & Lawton, 1992; Jost, 2007).
4. The Jaccard-type turnover-complement ($S_{qN}/\bar{S}_{qN}$) is the complement of the Jaccard-type turnover, which quantifies the normalized OTU turnover rate with respect to the whole system (i.e. gamma).

These generalizations encompass, as special cases some of the most popular (dis)similarity measures used in ecology (Chao, Chiu, et al., 2014a; Chao et al., 2016; Jost, 2007). For instance, C_{02} (the Sørensen-type overlap between two systems [N = 2] when OTU phylogenies are not considered and q = 0) produces the Sørensen similarity index, while C_{22} (idem but q = 2) yields the Morisita-Horn index. Another noteworthy example is that the measure 1 − $\bar{U}_{02}$ (the one-complement of the Jaccard-type overlap when OTU phylogenies are considered, q = 0 and N = 2) is identical to the UniFrac distance (Lozupone & Knight, 2005). Further relations between these four (dis)similarity measures and other popular indices can be found elsewhere (e.g., Jost, 2007, Chao et al., 2012, Chiu et al., 2014). If researchers opt for basing diversity measurements on Hill numbers, as advocated in this article, it is also advisable to frame dissimilarity measurements within the same scheme. Basing dissimilarity measurements on beta diversities derived from Hill numbers enables logical consistency to be kept with the conclusions based on Hill numbers (Chao et al., 2012; Jost, 2007). Furthermore, as all measures are continuous as q ranges from zero to infinity, (dis)similarity profiles can be made for any of them (Chiu et al., 2014).

8 | DEALING WITH ZERO-INFLATED, INSUFFICIENT AND BIASED DATA

All measures introduced throughout the article assume that the parameters of the analysed system are well-known, that is, that the OTU counts and relative abundances in the data set perfectly mirror the biological system under study. In practice, however, due to the high diversity of biological systems, their spatio-temporal heterogeneity and the complexity of the data processing methods, molecularly analysed systems are seldom characterized perfectly (Alberdi et al., 2019). Hence, there is a high chance that the Hill numbers calculated from the data differ from the actual diversity of the biological system (Chao, Gotelli, et al., 2014b). In DNA-based diversity assessments, there are three major interrelated sources of distortion that need to be assessed and, if possible, addressed: (a) zero-inflation, (b) insufficient and biased sample characterization effort (low sequencing depth) and sample size, and (c) varying OTU-detection probability.

The tendency to contain a large proportion of zero values is a distinguishing characteristic of many ecological data sets (Martin et al., 2005), but it is particularly relevant for DNA-derived data, due to the capacity to recover very high diversities (often > 1,000 OTUs). Data sets that exhibit a distribution of OTU counts biased towards a few dominant ones, and a "tail" of rare OTUs represented by a few counts, are common in molecular data sets. While the dominant OTUs tend to appear in many samples, the rare OTUs are often sparsely distributed, yielding data sets with high number of zeros

(Figure S2). When the high number of zeros produces the data do not readily fit standard distributions (e.g., normal, Poisson, binomial, negative-binomial), and thus complicate statistical analyses, the data set is referred to as zero-inflated (Heilbron, 1994).

One of the main challenges of zero-inflated data is the unknown nature of the zero values. Some might be true zeros due to the absence of the OTU in the system, yet others might be false zeros derived from imperfect characterization. Temporal and spatial variability of the distribution of OTUs can result in OTUs that actually occur in a system, being absent during the sampling period. The high incidence of such false zeros is the result of an insufficient or incorrect sampling design, and it is very relevant in diet analyses, because each sample usually contains just a fraction of the total prey range of a predator (Aizpurua et al., 2018). Another source of false zeros is the imperfect characterization of samples; that is, the OTU is present in the sample but not detected. This can occur because sequencing was not deep enough to recover a DNA sequence with very low abundance in the system, or because the bottlenecks during sample processing remove the rare DNA sequences (Alberdi et al., 2019). Zero-inflation is not the only issue of molecular diversity analyses though. Recently conducted research has clearly documented how many biological and technical distortion factors introduce numerous biases that break the relation between the actual biomass distribution in the system and the relative amount of DNA sequences obtained in the final results (Barnes & Turner, 2016; Lamb et al., 2019). This is partly due to primer amplification biases, and thus sequencing probability, due to the different binding affinity between primers and target sequences (Piñol, Senar, & Symondson, 2019). Finally, both PCR sequencing and DNA sequencing can generate artifactual DNA sequences that do not exist in the actual biological system, which results in increased false positive rate and inflated diversity (Alberdi et al., 2019). Consequently, in order to account for insufficient, biased and zero-inflated data, and thus produce reliable diversity assessments from molecularly characterized samples using Hill numbers, we encourage researchers to consider the following recommendations: (1) acknowledge, (2) assess, (3) correct and (4) model.

The essential initial step is to acknowledge the expected properties of the biological system to be studied, as well as the technical and statistical procedures to be employed, so that a correct study design is implemented. For example, in intraspecific analyses diet data tend to be more sparse than gut microbiome data at the OTU level, because while the intestinal microorganism community barely changes from one day to another, the diet at the OTU level can be completely different (Figure S2). In contrast, each sample in a diet data set tends to be less diverse than in a microbiome study. These differences will require different study designs and methodological procedures (e.g., incidence-based vs. abundance-based) to be employed, to ensure appropriate characterization of each system. Acknowledging the potential bias and error sources of the technical procedures is also essential, because the optimal study design will also depend on the procedures employed. Additionally, PCRs are known to exhibit certain degree of stochasticity (Alberdi et al., 2018) and the effect of sequencing runs can be as large, or even larger, than the actual biological signal (Chase et al., 2016).

Once the data are generated it is necessary to assess whether it is correct to address the research question. An essential procedure is to assess the completeness of the data both within each sample (sequencing depth, i.e., number of DNA sequences used to characterize each sample) and across samples (sample size, i.e., number of sample units used to characterize the system). The R packages iNext (Hsieh, Ma, & Chao, 2016) and iNextPD (Hsieh & Chao, 2017) offer tools based on Hill numbers and phylogenetic Hill numbers, respectively, to perform such operations efficiently with abundance and incidence data. The abundance approach is useful for assessing whether the sequencing depth of each sample is adequate, and the incidence approach enables sample size completeness estimations, although without considering the within-sample abundance distributions of the OTUs. An alternative approach to assess sample size completeness based on abundance (rather than incidence) data is to rarefy the gamma diversities of the study system at different sample sizes. This can be applied to both Hill numbers and phylogenetic Hill numbers, although the latter might require extensive computation time if the number of OTUs and samples is high.

The critical assessment of the data should drive researchers to take action to correct the data. If resources are available, increasing sequencing effort in undersequenced samples is usually possible, as only a fraction of the library is often sent to sequencing. This can minimize the impact of false zeros to a certain degree, yet it is essential to highlight that increased sequencing effort cannot recover diversity lost due to laboratory bottlenecks or PCR stochasticity (discussed in Alberdi et al., 2019). Sample size limitations and biases are more difficult to overcome, as obtaining more samples is often very costly or impossible. The impact of PCR stochasticity can be minimized by performing PCR replicates, and primer amplification biases can be moderated by lowering annealing temperatures or complementing the data with another primer set known to have different inherent biases (Alberdi et al., 2018). Multiple strategies to minimize the impact of artifactual sequences have been proposed, including the removal of DNA sequences below a certain count threshold, the use of PCR and sequencing replicates to enable constraining analyses only to sequences present in multiple replicates (Alberdi et al., 2018), and identification and removal of artifactual sequences by statistical means, either while bioinformatic processing of the sequences (Callahan et al., 2016), or when carrying out diversity analyses (Chiu & Chao, 2016).

In most cases, data sets cannot be completely corrected or completed; hence, researchers need to rely on modelling to obtain approximations for real diversity values. Diversity partitioning based on Hill numbers enables unequal sample sizes to be accounted for when comparing different systems, through applying different weights to the subsystems (Table 2). If the data are proven to be insufficient, the aforementioned R packages iNext and iNextPD enable extrapolation of rarefaction curves based on Hill numbers, to estimate the actual diversity of each sample (using abundance approach) or the whole system (using incidence approach) for orders of diversity 0, 1 and 2. Alternatively, another method to perform such estimations on any low order of diversity ($q \leq 3$), and hence enable ploting continuous diversity profiles, is implemented in the R package SpadeR (Chao & Jost, 2015).

These approaches enable some shortcomings of insufficient data to be addressed, although the issue of imperfect detection remains. It is well known that the probability to detect organisms depends on their abundances, but also on biological properties inherent to each organism. Occupancy-modelling approaches have traditionally been applied to overcome such detectability biases (MacKenzie, Nichols, Hines, Knutson, & Franklin, 2003). In molecularly characterized systems, the issue of detectability is even more complex, because in addition to the biological distortions of environmental DNA (Barnes & Turner, 2016), there is another important source of bias produced by uneven primer amplification rates (Piñol et al., 2019). While eDNA representativeness assessment might be too complex to model (Alberdi et al., 2019; Barnes & Turner, 2016), amplification biases can be measured *in silico* (Piñol et al., 2019) and using mock communities (Lamb et al., 2019). This enables implementing occupancy-modelling approaches that account for the relative amplification probability of OTUs (Ficetola et al., 2015), which have also been implemented within the framework of Hill numbers (Broms, Hooten, & Fitzpatrick, 2015; Iknayan, Tingley, Furnas, & Beissinger, 2014). These approaches are still in its infancy, but are likely to undergo a rapid development in the upcoming years.

9 | CONCLUSIONS

The extensive framework recently developed around Hill numbers provides a powerful toolset for the integrative analysis of multiple aspects of biological diversity. This approach could be applied in a wide range of scientific fields, under different ecological frameworks, such as community ecology or niche theory. As explained in this article, molecular data enable exploiting the full potential of this framework in a rather simple and straightforward way. It must be noted that the methods explained here are simply a practical overview of a more complex and broader statistical framework, which can be consulted in detail in several excellent reviews (e.g., Chao, Gotelli, et al., 2014b; Chao et al., 2016). Overall however, we hope that our piece and associated examples will encourage molecular ecologists to take advantage of Hill numbers, and in doing so be able to generate more logical, intuitive and reproducible results that will serve to improve the reliability and usefulness of their research.

ACKNOWLEDGEMENTS

We thank Orly Razgour, Iñaki Odriozola and David Moreno, as well as the anonymous reviewers, for their insightful comments on the manuscript. A.A. was supported by Lundbeckfonden (Grant R250-2017-1351), and M.T.P.G. acknowledges ERC Consolidator Grant (681396-Extinction Genomics).

AUTHOR CONTRIBUTION

A.A. conceived the idea and wrote the first drafts of the article and the supplementary documents. T.G. contributed to the discussion, manuscript writing and its revisions.

DATA AVAILABILITY STATEMENT

The data and scripts needed to reproduce the analyses mentioned in the article are available as Supporting Information.

ORCID

Antton Alberdi https://orcid.org/0000-0002-2875-6446

REFERENCES

Aizpurua, O., Budinski, I., Georgiakakis, P., Gopalakrishnan, S., Ibañez, I., Mata, V., … Alberdi, A. (2018). Agriculture shapes the trophic niche of a bat preying on multiple pest arthropods across Europe: Evidence from DNA metabarcoding. *Molecular Ecology*, *27*(3), 815–825. https://doi.org/10.1111/mec.14474

Alberdi, A., Aizpurua, O., Bohmann, K., Gopalakrishnan, S., Lynggaard, C., Nielsen, M., & Gilbert, M. T. P. (2019). Promises and pitfalls of using high-throughput sequencing for diet analysis. *Molecular Ecology Resources*, *19*(2), 327–348. https://doi.org/10.1111/1755-0998.12960

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, *9*, 134–147. https://doi.org/10.1111/2041-210X.12849

Allen, B., Kon, M., & Bar-Yam, Y. (2009). A new phylogenetic diversity measure generalizing the Shannon index and its application to phyllostomid bats. *American Naturalist*, *174*, 236–243. https://doi.org/10.1086/600101

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, *17*, 1–17. https://doi.org/10.1007/s10592-015-0775-4

Barwell, L., Isaac, N. J. B., & Kunin, W. E. (2015). Measuring beta-diversity with abundance data. *Journal of Animal Ecology*, *84*, 1112–1122.

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R., & Abebe, E. (2005). Defining operational taxonomic units using DNA barcode data.*Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*, 1935–1943. https://doi.org/10.1098/rstb.2005.1725

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., … de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution*, *29*, 358–367. https://doi.org/10.1016/j.tree.2014.04.003

Broms, K. M., Hooten, M. B., & Fitzpatrick, R. M. (2015). Accounting for imperfect detection in Hill numbers for biodiversity studies. *Methods in Ecology and Evolution*, *6*, 99–108. https://doi.org/10.1111/2041-210X.12296

Bush, A., Sollmann, R., Wilting, A., Bohmann, K., Cole, B., Balzter, H., … Yu, D. W. (2017). Connecting Earth observation to high-throughput biodiversity data. *Nature Ecology & Evolution*, *1*, 176. https://doi.org/10.1038/s41559-017-0176

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*, 581–583. https://doi.org/10.1038/nmeth.3869

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., … Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*, 335–336. https://doi.org/10.1038/nmeth.f.303

Cardinale, B. J., Palmer, M. A., & Collins, S. L. (2002). Species diversity enhances ecosystem functioning through interspecific facilitation. *Nature*, *415*, 426–429. https://doi.org/10.1038/415426a

Chao, A., Chiu, C.-H., & Hsieh, T. C. (2012). Proposing a resolution to debates on diversity partitioning. *Ecology*, *93*, 2037–2051. https://doi.org/10.1890/11-1817.1

Chao, A., Chiu, C.-H., & Jost, L. (2010). Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*, 3599–3609. https://doi.org/10.1098/rstb.2010.0272

Chao, A., Chiu, C.-H., & Jost, L. (2014a). 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, A., Chiu, C.-H., & Jost, L. (2016). Phylogenetic diversity measures and their decomposition: A framework based on hill numbers. In R. Pellens, & P. Grandcolas (Eds.), *Biodiversity conservation and phylogenetic systematics: Preserving our evolutionary heritage in an extinction crisis* (pp. 141–172). Cham: Springer International Publishing.

Chao, A., Chiu, C.-H., Villéger, S., Sun, I.-F., Thorn, S., Lin, Y.-C., … Sherwin, W. B. (2019). An attribute-diversity approach to functional diversity, functional beta diversity, and related (dis)similarity measures. *Ecological Monographs*. https://doi.org/10.1002/ecm.1343

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014b). Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, *84*, 45–67.

Chao, A., & Jost, L. (2015). Estimating diversity and entropy profiles via discovery rates of new species. *Methods in Ecology and Evolution*, *6*(8), 873–882. https://doi.org/10.1111/2041-210X.12349

Chao, A., Jost, L., Chiang, S. C., Jiang, Y.-H., & Chazdon, R. L. (2008). A two-stage probabilistic approach to multiple-community similarity indices. *Biometrics*, *64*, 1178–1186. https://doi.org/10.1111/j.1541-0420.2008.01010.x

Chase, J., Fouquier, J., Zare, M., Sonderegger, D. L., Knight, R., Kelley, S. T., … Caporaso, J. G. (2016). Geography and location are the primary drivers of office microbiome composition. *mSystems*, *1*(2), e00022–e116. https://doi.org/10.1128/mSystems.00022-16

Chiu, C. H., & Chao, A. (2016). Estimating and comparing microbial diversity in the presence of sequencing errors. *PeerJ*, *4*, e1634. https://doi.org/10.7717/peerj.1634

Chiu, C.-H., Jost, L., & Chao, A. (2014). Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs*, *84*, 21–44. https://doi.org/10.1890/12-0960.1

Colwell, R. K., & Futuyma, D. J. (1971). On the measurement of niche breadth and overlap. *Ecology*, *52*, 567–576. https://doi.org/10.2307/1934144

Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., & Eveson, J. P. (2019). Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular Ecology*, *28*(2), 391–406.

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*, 5872–5895. https://doi.org/10.1111/mec.14350

Douady, C. J., Delsuc, F., Boucher, Y., Doolittle, W. F., & Douzery, E. J. P. (2003). Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Molecular Biology and Evolution*, *20*, 248–254. https://doi.org/10.1093/molbev/msg042

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, *61*, 1–10. https://doi.org/10.1016/0006-3207(92)91201-3

Felsenstein, J. (1985). Phylogenies and the comparative method. *American Naturalist*, *125*, 1–15. https://doi.org/10.1086/284325

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., … Taberlet, P. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, *15*(3), 543–556. https://doi.org/10.1111/1755-0998.12338

Forister, M. L., Novotny, V., Panorska, A. K., Baje, L., Basset, Y., Butterill, P. T., … Dyer, L. A. (2015). The global distribution of diet breadth in insect herbivores. *Proceedings of the National Academy of Sciences, USA*, *112*, 442–447. https://doi.org/10.1073/pnas.1423042112

Gaggiotti, O. E., Chao, A., Peres-Neto, P., Chiu, C. H., Edwards, C., Fortin, M. J., … Selkoe, K. A. (2018). Diversity from genes to ecosystems: A unifying framework to study variation across biological metrics and scales. *Evolutionary Applications*, *11*, 1176–1193. https://doi.org/10.1111/eva.12593

Galan, M., Pons, J.-B., Tournayre, O., Pierre, É., Leuchtmann, M., Pontier, D., & Charbonnel, N. (2018). Metabarcoding for the parallel identification of several hundred predators and their prey: Application to bat species diet analysis. *Molecular Ecology Resources*, *18*, 474–489. https://doi.org/10.1111/1755-0998.12749

Gevers, D., Kugathasan, S., Denson, L. A., Vázquez-Baeza, Y., Van Treuren, W., Ren, B., … Xavier, R. J. (2014). The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host & Microbe*, *15*, 382–392. https://doi.org/10.1016/j.chom.2014.02.005

Harrison, S., Ross, S. J., & Lawton, J. H. (1992). Beta diversity on geographic gradients in Britain. *Journal of Animal Ecology*, *61*, 151–158. https://doi.org/10.2307/5518

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*, 313–321. https://doi.org/10.1098/rspb.2002.2218

Heilbron, D. C. (1994). Zero-altered and other regression models for count data with added zeros. *Biometrical Journal*, *36*, 531–547. https://doi.org/10.1002/bimj.4710360505

Hey, J., Waples, R. S., Arnold, M. L., Butlin, R. K., & Harrison, R. G. (2003). Understanding and confronting species uncertainty in biology and conservation. *Trends in Ecology & Evolution*, *18*, 597–603. https://doi.org/10.1016/j.tree.2003.08.014

Hill, M. O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, *54*, 427–432. https://doi.org/10.2307/1934352

Hsieh, T. C., & Chao, A. (2017). Rarefaction and extrapolation: Making fair comparison of abundance-sensitive phylogenetic diversity among multiple assemblages. *Systematic Biology*, *66*(1), 100–111.

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, *7*, 1451–1456. https://doi.org/10.1111/2041-210X.12613

Hurlbert, S. H. (1978). The measurement of niche overlap and some relatives. *Ecology*, *59*(1), 67–77. https://doi.org/10.2307/1936632

Iknayan, K. J., Tingley, M. W., Furnas, B. J., & Beissinger, S. R. (2014). Detecting diversity: Emerging methods to estimate species diversity. *Trends in Ecology and Evolution*, *29*(2), 97–106. https://doi.org/10.1016/j.tree.2013.10.012

Jost, L. (2006). Entropy and diversity. *Oikos*, *113*, 363–375.

Jost, L. (2007). Partitioning diversity into independent alpha and beta components. *Ecology*, *88*, 2427–2439. https://doi.org/10.1890/06-1736.1

Kartzinel, T. R., Chen, P. A., Coverdale, T. C., Erickson, D. L., Kress, W. J., Kuzmina, M. L., … Pringle, R. M. (2015). DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. *Proceedings of the National Academy of Sciences, USA*, *112*, 8019–8024. https://doi.org/10.1073/pnas.1503283112

Keylock, C. J. (2005). Simpson diversity and the Shannon-Wiener index as special cases of a generalized entropy. *Oikos*, *109*, 203–207. https://doi.org/10.1111/j.0030-1299.2005.13735.x

Kim, M., Oh, H.-S., Park, S.-C., & Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, *64*, 346–351. https://doi.org/10.1099/ijs.0.059774-0

Krebs, C. J. (2014). Ecological Methodology. http://www.zoology.ubc.ca/~krebs/books.html

Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G., & Taylor, M. I. (2019). How quantitative is metabarcoding: A

meta-analytical approach. *Molecular Ecology*, *28*(2), 420–430. https://doi.org/10.1111/mec.14920

Leinster, T., & Cobbold, C. A. (2012). Measuring diversity: The importance of species similarity. *Ecology*, *93*, 477–489. https://doi.org/10.1890/10-2402.1

Levins, R. (1968). *Evolution in changing environments: Some theoretical explorations*. Princeton, USA: Princeton University Press.

Lozupone, C., & Knight, R. (2005). UniFrac: A new phylogenetic method for comparing microbial communities. *Applied and Environment Microbiology*, *71*, 8228–8235. https://doi.org/10.1128/AEM.71.12.8228-8235.2005

Lucas, R., Groeneveld, J., Harms, H., Johst, K., Frank, K., & Kleinsteuber, S. (2017). A critical evaluation of ecological indices for the comparative analysis of microbial communities based on molecular datasets. *FEMS Microbiology Ecology*, *93*, fiw209. https://doi.org/10.1093/femsec/fiw209

MacArthur, R. H. (1965). Patterns of species diversity. *Biological Reviews of the Cambridge Philosophical Society*, *40*, 510–533. https://doi.org/10.1111/j.1469-185X.1965.tb00815.x

MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G., & Franklin, A. B. (2003). Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology*, *84*(8), 2200–2207. https://doi.org/10.1890/02-3090

Marcon, E., & Hérault, B. (2015). Decomposing phylodiversity. *Methods in Ecology and Evolution*, *6*, 333–339. https://doi.org/10.1111/2041-210X.12323

Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., … Possingham, H. P. (2005). Zero tolerance ecology: Improving ecological inference by modelling the source of zero observations. *Ecology Letters*, *8*(11), 1235–1246. https://doi.org/10.1111/j.1461-0248.2005.00826.x

Mayr, E., & Provine, W. B. (1998). *The evolutionary synthesis: Perspectives on the unification of biology*. Cambridge, USA: Harvard University Press.

McElroy, K., Zagordi, O., Bull, R., Luciani, F., & Beerenwinkel, N. (2013). Accurate single nucleotide variant detection in viral populations by combining probabilistic clustering with a statistical test of strand bias. *BMC Genomics*, *14*, 501. https://doi.org/10.1186/1471-2164-14-501

McIntosh, R. P. (1967). An index of diversity and the relation of certain concepts to diversity. *Ecology*, *48*, 392–404. https://doi.org/10.2307/1932674

Nascimento, F. F., Reis, M. D., & Yang, Z. (2017). A biologist's guide to Bayesian phylogenetic analysis. *Nature Ecology & Evolution*, *1*, 1446–1454. https://doi.org/10.1038/s41559-017-0280-x

Nipperess, D. A., Faith, D. P., & Barton, K. (2010). Resemblance in phylogenetic diversity among ecological assemblages. *Journal of Vegetation Science*, *21*, 809–820. https://doi.org/10.1111/j.1654-1103.2010.01192.x

Pavoine, S., Bonsall, M. B., Dupaix, A., Jacob, U., & Ricotta, C. (2017). From phylogenetic to functional originality: Guide through indices and new developments. *Ecological Indicators*, *82*, 196–205. https://doi.org/10.1016/j.ecolind.2017.06.056

Pavoine, S., & Izsák, J. (2014). New biodiversity measure that includes consistent interspecific and intraspecific components. *Methods in Ecology and Evolution*, *5*, 165–172. https://doi.org/10.1111/2041-210X.12142

Pavoine, S., Love, M. S., & Bonsall, M. B. (2009). Hierarchical partitioning of evolutionary and ecological patterns in the organization of phylogenetically-structured species assemblages: Application to rockfish (genus: Sebastes) in the Southern California Bight. *Ecology Letters*, *12*, 898–908.

Peterson, J., Garges, S., Giovanni, M., McInnes, P., Wang, L., Schloss, J. A., … Guyer, M. (2009). The NIH human microbiome project. *Genome Research*, *19*, 2317–2323. https://doi.org/10.1101/gr.096651.109

Pianka, E. R. (1974). Niche overlap and diffuse competition. *Proceedings of the National Academy of Sciences of the United States of America*, *71*, 2141–2145. https://doi.org/10.1073/pnas.71.5.2141

Pielou, E. C. (1966). Species-diversity and pattern-diversity in the study of ecological succession. *Journal of Theoretical Biology*, *10*, 370–383. https://doi.org/10.1016/0022-5193(66)90133-0

Pielou, E. C. (1975). *Ecological diversity*. New York, USA: Wiley & Sons.

Piñol, J., Senar, M. A., & Symondson, W. O. C. (2019). The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. *Molecular Ecology*, *28*(2), 407–419.

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology*, *35*, 833–844. https://doi.org/10.1038/nbt.3935

Rao, C. R. (1982). Diversity and dissimilarity coefficients: A unified approach. *Theoretical Population Biology*, *21*, 24–43. https://doi.org/10.1016/0040-5809(82)90004-1

Rideout, J. R., He, Y., Navas-Molina, J. A., Walters, W. A., Ursell, L. K., Gibbons, S. M., … Caporaso, J. G. (2014). Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ*, *2*, e545. https://doi.org/10.7717/peerj.545

Rosen, M. J., Davison, M., Bhaya, D., & Fisher, D. S. (2015). Microbial diversity. Fine-scale diversity and extensive recombination in a quasisexual bacterial population occupying a broad niche. *Science*, *348*, 1019–1023. https://doi.org/10.1126/science.aaa4456

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., … Weber, C. F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environment Microbiology*, *75*, 7537–7541. https://doi.org/10.1128/AEM.01541-09

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*, 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., … Mazel, F. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews of the Cambridge Philosophical Society*, *92*, 698–715. https://doi.org/10.1111/brv.12252

Tuomisto, H. (2010a). A diversity of beta diversities: Straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, *33*, 2–22.

Tuomisto, H. (2010b). A diversity of beta diversities: Straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography*, *33*, 23–45.

Whittaker, R. H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*, *30*, 279–338. https://doi.org/10.2307/1943563

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Alberdi A, Gilbert MTP. A guide to the application of Hill numbers to DNA-based diversity analyses. *Mol Ecol Resour*. 2019;19:804–817. https://doi.org/10.1111/1755-0998.13014