

Defining, measuring, and partitioning species diversity

Author

Hanna Tuomisto

Department of Biology, University of Turku, FI-20014 University of Turku, Finland

Published in: Reference Module in Life Sciences 2017. Elsevier.

doi: 10.1016/B978-0-12-809633-8.02377-3

Article Outline

I. Introduction

II. Defining diversity

III. Practical issues in quantifying diversity

IV. Diversity indices

V. Partitioning diversity

VI. Resemblance indices related to diversity

Abstract

A practical definition of diversity is the effective number of types (such as species). This equals the number of equally-abundant types needed to get the observed mean proportional abundance of types. Diversity indices represent other phenomena, such as entropy and probability. Richness is the actual rather than effective number of types. Total (gamma) diversity can be partitioned multiplicatively into alpha diversity (mean diversity per subunit) and beta diversity (the effective number of compositionally distinct subunits) or additively into species diversity in an average subunit and species turnover. Resemblance indices can be derived as monotonic transformations of beta diversity or beta richness.

Keywords

alpha diversity, beta diversity, evenness, gamma diversity, Horn index, Jaccard index, Morisita-Horn index, richness, Shannon index, Simpson index, species turnover, Sørensen index

Glossary

alpha diversity The mean effective density of types (such as species) in a dataset, as expressed on a per compositional unit basis. See eqn [15].

beta diversity The effective number of compositional units in a dataset. See eqn [17].

- compositional unit** A virtual subunit that contains as many types (such as species) as the actual subunits do on average, but does not share any types with the other virtual subunits. See Figure 2.
- diversity** The effective number of types (such as species) into which the entities (such as individuals) that form the dataset of interest have been classified. Also known as 'true diversity' and 'Hill number'. See Figure 1.
- effective species** A virtual species that has the same proportional abundance as the actual species do on average. The number of effective species, or effective number of species, is species diversity. See Figure 1.
- evenness** The effective number of types expressed as a proportion of the actual number of types in the dataset. See eqn [9].
- gamma diversity** Total effective number of types (such as species) in a dataset. See eqn [10].
- generalized mean with exponent $q - 1$** A general equation that allows deriving different kinds of mean (e.g., arithmetic and geometric) by varying the parameter q . See eqn [2].
- Gini-Simpson index** An index that quantifies the probability that two entities (such as individuals) drawn at random from the dataset represent different types (such as species). Often used as a diversity index, but not to be confused with diversity itself. See the explanation after eqn [6].
- proportional abundance** The abundance (such as number of individuals) of one type (such as species) expressed as a proportion of the total abundance (number of individuals) of all types (species) in the dataset.
- richness** The actual number of types (such as species) into which the entities (such as individuals) that form the dataset of interest have been classified.
- Shannon entropy** An index that quantifies the uncertainty regarding which type (such as species) is picked when one entity (such as individual) is drawn at random from the dataset of interest. Often used as a diversity index, but not to be confused with diversity itself. See eqn [4].
- Simpson index** An index that quantifies the probability that two entities (such as individuals) drawn at random from the dataset represent the same type (such as species). Often used as a diversity index, but not to be confused with diversity itself. See the explanation after eqn [6].
- species turnover** A measure of the amount by which species composition changes among subunits in a dataset.

Body text

I. Introduction

Confusion around the concept 'diversity' has long hampered accurate communication among researchers. For example, Hurlbert (1971) lamented: "The term 'species diversity' has been defined in such various and disparate ways that it now conveys no information other than

'something to do with community structure'; species diversity has become a nonconcept". Hill (1973) argued for a logically coherent definition of diversity, which implies making a distinction between diversity itself and different kinds of diversity indices. Unfortunately, ecologists have continued to treat diversity indices as if they measured the same thing as diversity itself does, which makes much of the diversity literature very difficult to understand and interpret in a coherent way.

In recent years, awareness of the problems caused by inconsistent and often contradictory use of terminology has increased. Since Jost (2006, 2007) revived Hill's definition of diversity, it has become possible to appreciate that the concept of diversity is actually quite simple (Tuomisto 2010a, b, c, 2011). The simplicity emerges from three main sources. Firstly, the concepts become less ambiguous when an explicit difference is made between diversity itself and diversity indices. An index of something is just a surrogate for the thing itself, and for an index to be useful, it needs to be clear what it is an index of. This necessitates an unambiguous definition of diversity itself. Secondly, Hill's definition of diversity is understandable in simple, everyday terms, which makes diversity a much less abstract concept than ecologists have often thought. Thirdly, separating the conceptual problem of defining diversity from the practical problem of deciding how to adequately quantify diversity for a community of interest allows a clearer view on the issue. As a result, it has become a feasible goal to reach a consistent terminology for diversity and related concepts (Jurasinski et al., 2009; Moreno & Rodríguez 2010, 2011; Jurasinski & Koch 2011; Tuomisto 2011).

Before getting into the details, it is worth noting that diversity can be quantified for any dataset for which observed entities have been classified into types. The entities are often individuals, but with organisms whose individuals are difficult to delimit (such as clonal plants or colonial animals), the entities may also be something else, such as ramets, colonies, or units of cover or biomass.

If species diversity is of interest, the types into which the entities are classified are species. In other cases, the entities may be classified into, for example, genera, families, haplotypes, operational taxonomic units or functional types. To calculate the diversity of the dataset, all that is needed is knowledge of the proportional abundance of each of the relevant types within the dataset. Species richness is related to species diversity, but is not the same thing: richness does not take the proportional abundances into account. In the present paper, the letter *D* will be used for diversity and the letter *R* for richness. The traditional notation *S* refers specifically to species richness, and is not used here to avoid restricting the discussion to this specific kind of richness. For simplicity, the present chapter is nonetheless written in terms of species richness and species diversity, but the same logic applies to any types or categories of interest.

II. Defining diversity

Species richness (R) of a dataset can be defined simply as the number of species names in the species list that is obtained after every individual (or other entity) that belongs to the dataset has been identified to species. This can be called the actual number of species. The effective number of species, or species diversity (D), is a more complicated concept, because it not only takes into account how many species the dataset contains, but also their relative abundances. This can be done in many different ways, which has led to a variety of definitions for the concept 'diversity' (reviewed in Peet 1974; Magurran 2004; Jost 2006, 2007; Gauthier et al. 2010; Tuomisto 2010a). In the interest of accurate scientific communication, here the term 'species diversity' is only used to refer to the effective number of species, following Hill (1973).

This definition of diversity can be derived as follows. Let us first define the proportional abundance of species i in the dataset to equal $p_i = m_i/m$, where m_i is the absolute abundance of species i (e.g. number of individuals), and m is the total abundance of all species in the dataset. For simplicity, the present text is written with the number of individuals as the abundance measure, but the same principles apply when other units of abundance are used (e.g., grams of biomass or square metres of cover). When all species are equally abundant, then $m = m_i \times R$, and the proportional abundance p_i of each species equals $m_i/(m_i \times R)$, which simplifies to the reciprocal of richness $1/R$. The mean proportional species abundance then also equals $1/R$. When all species are not equally abundant, some of them have a larger and others a smaller proportional abundance, but the mean of the proportional abundances is nevertheless a single value \bar{p}_i , which can also be expressed $\bar{m}_i/(\bar{m}_i \times D) = 1/D$. Species diversity D can hereby be defined as the number of equally-abundant species that would be needed to obtain the same mean proportional species abundance as that observed in the dataset of interest. This is what the effective number of species means in practice (MacArthur 1964, 1965; Hill 1973; Routledge 1979; Fig. 1).

< Figure 1 near here >

A mean can be calculated in several different ways, which has implications for the quantification of diversity. The different kinds of mean yield different values when applied to the same dataset, so the diversities based on them also obtain different values. Hill's (1973) definition of diversity can be written in the general form

$${}^q D = 1 / {}^q \bar{p}_i \quad [1]$$

Here ${}^q D$ is diversity of order q , and ${}^q \bar{p}_i$ is the corresponding mean of the proportional species abundances. The mean here is the weighted generalized mean with exponent $q - 1$, which is calculated

$${}^q \bar{p}_i = {}^{q-1} \sqrt[{}^{q-1}]{\sum_{i=1}^R p_i p_i^{q-1}} \quad [2]$$

When calculating the mean, each p_i value is nominally weighted (i.e., multiplied) by itself, but the exponent $q - 1$ modifies the weighting. When $q = 1$, the final weights equal the nominal weights, which implies that each species contributes to the value of the mean in proportion to the amount of data (e.g. number of individuals) it contributed to the dataset. In other words, each individual has the same weight no matter which species it belongs to: each individual pulls the mean towards the p_i value of its own species by exactly the same weight irrespective of its species identity. Smaller values of q ($q < 1$) de-emphasise the abundance differences among species. This gives rare species more weight than implied by their p_i , and individuals of rare species hence pull the mean towards the p_i value of their own species with more weight than individuals of abundant species do. Conversely, larger values ($q > 1$) exaggerate the abundance differences among species and hence give abundant species (and individuals belonging to them) disproportionate weight. Specific values of q correspond to familiar kinds of means (Hill 1973):

$q \rightarrow -\infty$ minimum p_i value

$q = 0$ harmonic mean $= 1 / \sum_{i=1}^R p_i / p_i$

$q = 1$ geometric mean $= \prod_{i=1}^R p_i^{p_i}$

$q = 2$ arithmetic mean $= \sum_{i=1}^R p_i p_i$

$q \rightarrow \infty$ maximum p_i value

In practice, the parameter q determines where the mean proportional species abundance lands within the possible range between the minimum and maximum p_i observed in the dataset (see Fig. 1 for a numerical example).

By combining eqns [1] and [2], the definition of diversity can be written:

$${}^q D = 1 / \sqrt[q-1]{\sum_{i=1}^R p_i p_i^{q-1}} \quad [3]$$

If all p_i values are identical in a dataset, the number of effective species (${}^q D$) in that dataset equals the number of actual species (R) no matter which value of q is chosen. If different species have different p_i values in a dataset, increasing q causes mean p_i to increase and hence diversity ${}^q D$ to decrease for that dataset. At $q = 0$, the number of effective species (${}^0 D$) equals the number of actual species (R) even when all species are not equally abundant. At $q = 1$, diversity equals the exponential of the Shannon index, and at $q = 2$, diversity equals the reciprocal of the Simpson index (these diversity indices will be discussed in more detail below). When all species are not equally abundant, negative values of q cause the number of effective species to exceed the number of actual species (${}^q D > R$), and positive values of q

cause the number of effective species to become smaller than the number of actual species (${}^qD < R$).

In diversity studies, q is usually limited to non-negative values. This is because increasing the weight given to the rarest species makes the mean proportional species abundance more dependent on the total size of the dataset. The rarest species in a dataset are often singletons (represented by a single individual), and their proportional abundance is $1/m$. As q approaches negative infinity, the mean of the proportional abundances approaches this minimum value, and the effective number of species therefore approaches m , i.e. the total number of individuals in the dataset. It is difficult to think of practical applications where this property would be desirable in a diversity measure, and setting the restriction $q \geq 0$ ensures it does not cause problems of interpretation.

Obviously, qD can obtain very different values depending on the chosen value of the parameter q . However, the conceptual interpretation of qD does not change with q : it always quantifies the effective number of species (or other types of interest) in the dataset. Depending on the questions at hand, it might be appropriate to choose just a single mean to calculate qD , or several different means could be used in parallel (by choosing different values for the parameter q) to explore how the effective number of species varies when rare vs. abundant species are emphasised in different ways (Ricotta & Avena, 2002; Kindt et al., 2006; Tuomisto 2012; Tuomisto et al., 2014; Chao & Jost, 2015).

The concept of diversity as an effective number of types was apparently first used in ecology by MacArthur (1964, 1965), but it was more thoroughly discussed by Hill (1973). Due to Hill's influential paper on the topic, the diversity values have been known as 'Hill numbers' (or 'Hill's numbers') in the ecological literature. Although some researchers adopted Hill numbers relatively early (e.g. Routledge 1977, 1979), they have generally been treated as just another diversity index (e.g., Magurran 2004). However, it has been argued that the effective number of types should be used as the only definition of diversity (Jost 2006, 2007, 2009; Tuomisto 2010a, 2011, 2012), and Jost proposed to call it 'true diversity' to distinguish it from diversity indices that represent phenomena other than a count of effective species. This narrow definition of 'diversity' is gaining in influence, because it makes intuitive sense, and at the same time it considerably clarifies communication about diversity issues (e.g., de Bello et al., 2010; Gauthier et al., 2010; Gregorius, 2010; Jurasinski & Koch 2011; Moreno & Rodríguez 2011; Chao et al., 2014; Morante-Filho et al., 2016).

It can be noted that the use of the word 'true' in 'true diversity' is analogous to its use in 'true bugs' (Tuomisto 2011). In both cases, the word is used as a synonym of 'in the narrow sense'. In the case of bugs, 'true' indicates that the interest is in insects of the monophyletic order Heteroptera, rather than in the polyphyletic assemblage consisting of any insect-like animals and pathogenic microorganisms, which are also called 'bugs' in everyday contexts. In the case of diversity, 'true' indicates that the interest is in the effective number of types, rather than in the conceptually heterogeneous group of phenomena consisting of numbers of types, entropies, probabilities and differences, which have also been called 'diversity' in the ecological literature. Gregorius (2010) used the term 'explicit diversity' instead of 'true

diversity', possibly because the latter has raised negative reactions as some researchers have misinterpreted 'true' to mean 'best for any purpose' (Hoffman and Hoffman 2008; Anderson et al. 2011; Gorelick 2011).

The greater the largest p_i values in a given dataset, the larger the mean p_i and hence the smaller the number of equally-abundant species that would yield the same mean p_i value. These hypothetical, equally-abundant species can be called effective species (${}^q\text{sp}_E$) and they are the unit of measurement of species diversity qD (note the use of the superscript q to specify which mean is used). In contrast, species richness R is measured in units of actual species (sp; Tuomisto, 2010a, c). When the individuals (or other observed entities) are classified into types other than species (such as genera, families, OTUs, functional types, haplotypes or size classes), the measurement units of richness and diversity change accordingly (e.g., actual and effective genera, respectively).

Both richness and diversity follow the important replication principle: if each of the original species is replicated to form r different species, each of which has the same absolute abundance as the original, both richness and diversity of the dataset increase to r times their original values (Hill, 1973; Jost, 2006, 2007). The replication principle corresponds to an intuitive sense of how diversity is expected to behave. The traditional diversity indices (such as the Shannon and Gini-Simpson indices) do not follow the replication principle, which is one reason why their values are difficult to interpret and indeed are often misinterpreted (Jost, 2006, 2007, 2009).

III. Practical issues in quantifying diversity

Often ecologists are interested in the species richness or diversity of a community, but are able to sample only a tiny fraction of the individuals belonging to that community. Since any one of the unseen individuals could represent a species not recorded in the sample, the species richness of the community is generally (and often considerably) underestimated by the recorded data. Therefore, richness and diversity values that are accurate for an existing dataset provide no more than estimates for the community of interest. How accurate such estimates are depends on how representative the dataset is of the community. In general, the more species a community contains, the more individuals need to be observed for the sample to become representative. Different kinds of sampling and data analysis methods can be used to improve the accuracy of the estimate, just as in other situations where results are extrapolated beyond the actual observations (e.g., Colwell and Coddington 1994; Plotkin and Muller-Landau 2002; Chao et al. 2006, 2013; Beck and Schwanghart 2010; Gauthier et al. 2010; Chao & Jost 2015).

Much of the discussion in the past on how to define diversity has actually focused on how to accurately estimate diversity for a community of interest from a sample of that community. For example, it has often been recommended that diversity is best measured using diversity indices that stabilise at small sample sizes (e.g., Lande 1996; Magurran 2004; Beck and

Schwanghart 2010). By this criterion, species richness is a very bad index of diversity, the Shannon index clearly better, and the Simpson index superior to both. However, the conceptual definition of a term should not depend on context-specific details such as sampling decisions, the ease of measurement or the questions of interest in a particular study. Consider observing a mixed-species flock of birds. The definition of the concept 'species diversity' should not depend on how difficult it is to measure the species diversity of the flock, just as the definition of the concept 'volume' should not depend on how difficult it is to estimate the volume of the space occupied by the flock. Of course, decisions such as which individuals are considered to belong to the flock and which of those can be identified to species determine whether the measurement result is appropriate and sufficiently accurate for the intended purpose. However, this is a separate issue that has no bearing on the conceptual definition of the phenomenon itself (whether diversity or volume).

In the present chapter, the focus is on conceptual definitions, so richness and diversity will be treated simply as properties of an existing dataset (see also Smith and Wilson 1996, Gregorius & Gillet 2008; Tuomisto 2010a, b, c). Before diversity and richness can be quantified, a decision therefore has to be made on the limits of the dataset, i.e. which individuals (or other entities) belong to it. Every different set of individuals in effect corresponds to a different dataset, and consequently may yield different richness and diversity values.

Each individual that is added to a dataset may represent a species that was not yet present in the dataset before. Therefore, drawing meaningful ecological conclusions from comparisons of richness and diversity values among datasets necessitates that the methods by which individuals have been selected for the datasets are comparable, which in practice implies that sampling efforts must have been standardised across the datasets in an appropriate way (Gotelli and Colwell 2001; Gauthier et al. 2010; Tuomisto 2010b; Chao et al., 2013). For example, bird inventories might be based on making observations in a fixed observation point either for a pre-defined amount of time, or until a pre-defined number of individuals has been recorded. Tree inventories could be based on recording all stems of a pre-defined minimum diameter in a plot of a pre-defined shape and surface area, or on recording a pre-defined number of stems using some plotless method. Inventory of marine benthic fauna could be based on animals captured by sieving cores of a specific volume through a mesh of a specific size.

If the communities of interest vary widely in species diversity, standardising relative sampling effort (representativeness or coverage) may sometimes be more important than standardising absolute sampling effort (e.g. number of individuals). A relevant measure of sampling effort in such cases would pay attention to species accumulation curves and/or the proportion of singletons in the data (Good and Toulmin, 1956; Chao et al., 2013).

If all individuals within a fixed surface area have not actually been observed and included in the dataset, then surface area cannot be used as a measure of sampling effort. For example, the nominal sampling unit size in macroecological studies may be fixed (e.g., 50 km by 50 km grid cells), but the actual sampling effort invested in each sampling unit (for example,

time spent in the field or number of specimens available in herbaria) can vary by several orders of magnitude among grid cells. In such cases, the observed patterns in richness and diversity may depend more on the uneven distribution of field work effort than on any ecologically interesting phenomenon.

In addition to defining the limits of the dataset, it is important to define how abundance is measured. Many different ways are possible and justifiable. Possible measures include the number of individuals, the number of stems or ramets, biomass, basal area, volume or the percentage of surface cover. Obviously, if the size of individuals varies (as it generally does), the number of individuals and biomass can be expected to give different diversity values even when calculated for the same dataset, and diversity values based on different abundance measures are therefore not comparable across datasets.

Both richness and diversity also depend on how the individuals (or other observed entities) are classified into types. A lumper and a splitter will rarely report the same species richness for a given dataset, and family richness can change dramatically depending on which family circumscriptions are used. Consequently, interpretation of results for richness and diversity requires careful specification of the classification that was used. Comparing richness or diversity values among datasets is meaningful only if all of them are based on the same classification. For example, if one dataset has identified tropical trees to the species level and another only to the genus level, comparing richness and diversity values between them yields meaningless and misleading ecological conclusions.

IV. Diversity indices

A. Shannon index

Several indices of diversity have been popular in ecology (reviewed in Magurran 2004; Gauthier et al. 2010). Until recently, too little attention has been paid to the fundamental conceptual differences among those indices. Consequently, the term 'diversity' has been used in several conceptually different ways in the ecological literature, as indices of diversity have been equated with diversity itself (Jost 2006, 2007; Tuomisto 2010a, c).

The most popular diversity index by far in the ecological literature has been the Shannon entropy H' (Shannon 1948):

$$H' = - \sum_{i=1}^R p_i \log p_i = \log({}^1D) \quad [4]$$

This measure is also known as the Shannon index, the Shannon-Weaver index, the Shannon-Wiener index, the Shannon-Weiner index and (confusingly) the Shannon diversity. The Shannon entropy does not quantify species diversity (the effective number of species) but the

uncertainty in the species identity of an individual that is picked at random from the dataset. The measurement unit of Shannon entropy depends on the base of the logarithm, and can be, for example, bits, nats or decits when using log bases 2, e and 10, respectively. Taking the exponential of the Shannon entropy recovers true diversity, i.e. ${}^1D = \exp(H')$.

The Shannon entropy is a special case for $q = 1$ of the Rényi entropy (Rényi 1960; Tuomisto 2010a)

$${}^qH' = \log({}^qD) = \log(1/\bar{p}_i) \quad [5]$$

Here the value of q defines how the probability of choosing a particular individual is affected by the abundance of the species it belongs to. When $q = 1$, each individual has the same probability of being chosen, and hence the probability that the chosen individual represents species i equals p_i . When $q = 0$, each species has the same probability of being chosen no matter what its proportional abundance. The probability that the chosen individual belongs to the most abundant species of the dataset increases monotonically with increasing q . With sufficiently large q , the chosen individual almost certainly belongs to the single most abundant species in the dataset, which has proportional abundance p_{\max} . The uncertainty in the species identity of the randomly chosen individual (= Rényi entropy) hence decreases asymptotically towards $\log(1/p_{\max})$ as q increases towards infinity. This parallels the behaviour of the mean proportional species abundance, which increases towards the abundance of the most abundant species as q increases (see explanation after eqn [2]).

B. Simpson index and Gini-Simpson index

Equation [3] is usually written in the form

$${}^qD = \left(\sum_{i=1}^R p_i^q \right)^{1/(1-q)} \quad [6]$$

The term inside the parentheses ($= {}^qD^{1-q}$), known as the 'basic sum', has often been used as a diversity index. When $q = 2$, the basic sum equals the Simpson index $\sum_{i=1}^R p_i^2$ (Simpson 1949;

Hill 1973; Jost 2006, 2007; Tuomisto 2010a), which quantifies the probability that two individuals picked at random from the dataset (with replacement) represent the same species. Hurlbert (1971) called this the probability of interspecific encounter (*PIE*, which can also be calculated, in a slightly modified form, to account for not replacing the first individual before choosing the second). The Simpson index decreases as the species richness of a dataset increases, so in itself it is an impractical index of diversity. A popular transformation of the Simpson index that increases with increasing richness is the Gini-Simpson index, which is obtained by subtracting the Simpson index from unity. Therefore, the Gini-Simpson index quantifies the probability that two individuals picked at random from the dataset (with replacement) represent different species.

Another transformation of the Simpson index that also increases with increasing richness is its reciprocal. Although traditionally less popular as a diversity index than the Gini-Simpson index, the inverse Simpson index actually equals 2D , i.e. true diversity of order 2. In other words, the Simpson index itself represents the weighted arithmetic mean of the proportional species abundances, rather than an effective number of species.

The Gini-Simpson index is a special case of an entropy that is known in ecology as the Tsallis or HCDDT entropy (Jost 2006; Mendes et al. 2008). This entropy can be expressed as:

$${}^qT = (1 - {}^qD^{1-q})/(q - 1) \quad [7]$$

The HCDDT entropy is a simple transformation of the basic sum, and its value is related to the probability that q individuals chosen at random from the dataset (with replacement) represent at least two different species.

From eqns [3] and [6] it can be seen that the basic sum is just an intermediate step in calculating mean proportional species abundance and its reciprocal, true diversity. Therefore, using the basic sum to derive an index of diversity is unnecessary; diversity itself is just as easy to calculate as the diversity index, and the risk of misinterpreting the results is diminished if the variable that is reported represents the actual phenomenon of interest rather than a surrogate (or transformation) thereof (Tuomisto 2010c).

V. Partitioning diversity

A. *Multiplicative partitioning: richness × evenness*

Species diversity is traditionally thought to consist of two independent components, richness and evenness (e.g., Peet 1974; Alatalo 1981; Smith & Wilson 1996; Magurran 2004; Kindt et al. 2006; Mendes et al. 2008; Jost 2010). Components that represent different phenomena can be combined by multiplication, which in this case leads to the following partitioning of diversity (Jost 2010; Tuomisto 2012):

$$\text{diversity} = \text{richness} \times \text{evenness} \quad [8]$$

Since richness and diversity were already defined above, evenness can now be defined as (see Tuomisto 2012 for a review of which other phenomena have also been called 'evenness'):

$${}^qE = {}^qD/R \quad [9]$$

Because richness is based on presence-absence data and hence does not depend on the proportional species abundances p_i , it is not sensitive to the value of q . In contrast, diversity and evenness decrease when the value of q is increased, unless all species (or other types of interest) are equally abundant in the dataset of interest. Consequently, the value of q must always be specified to allow accurate communication about diversity and evenness.

In practice, qE quantifies how many effective types there are for each actual type in the dataset of interest. Species richness is based on presence-absence (binary) data, and is measured in units of actual species; species diversity is based on abundance (quantitative) data, and is measured in units of effective species. Consequently, the measurement unit of evenness is ${}^q\text{sp}_E/\text{sp}$. The highest possible species evenness ${}^qE_{\max}$ equals $1 {}^q\text{sp}_E/\text{sp}$. This is obtained when there are as many effective as actual species, i.e. when either all species are equally abundant or $q = 0$. When $q > 0$, diversity qD decreases towards $1 {}^q\text{sp}_E$ as the proportional abundance of the most abundant species increases towards unity. As a result, the smallest possible value of evenness ${}^qE_{\min}$ is $(1 {}^q\text{sp}_E)/R$.

For any given dataset, small values of q lead to higher evenness than large values of q . This follows directly from the fact that diversity decreases when q increases, but richness does not change (see *Defining diversity*, above). Different kinds of species abundance distribution lead to different trajectories of diversity qD , and hence of evenness qE , when these are plotted as a function of q (Ricotta & Avena 2002; Kindt et al. 2006; Jost 2010; Tuomisto 2012). Graphs showing qE and qD against q provide complementary insights for comparing datasets that differ in R .

B. Multiplicative partitioning: $\alpha \times \beta$

1. The general principle

When Whittaker (1960) coined the terms alpha, beta and gamma diversity, his intention was to understand the species diversity in a landscape (gamma diversity) as the combined result of two different phenomena, namely the species diversity at a more local scale (alpha diversity), and the compositional heterogeneity among localities (beta diversity). Whittaker explored several approaches to quantifying compositional heterogeneity, and quite some confusion has resulted because he referred to all of them as 'beta diversity'. Other researchers have further expanded the circumscription of beta diversity, such that the term has been used to refer to more than 30 different phenomena in the ecological literature (e.g., Wilson and Shmida 1984; Vellend 2001; Koleff et al. 2003; Jurasinski et al. 2009; Tuomisto 2010a, b; Anderson et al. 2011).

Some of these "beta diversities" are not mathematically related to alpha and gamma diversity in any way, and the values of many are uncorrelated with each other. Although most of the resulting measures are useful in themselves, each one of them quantifies a different phenomenon and is therefore relevant for addressing different ecological questions (conceptual differences among the various measures have been reviewed in Tuomisto 2010a, b). Much of the confusion surrounding the beta diversity concept has emerged because the same term has been used to refer to very different things. This has also caused many unwarranted comparisons among studies and incorrect inferences from data.

When introducing the term 'beta diversity', Whittaker (1960) wrote that its simplest definition is $\beta = \gamma/\alpha$. This corresponds to what has since become known as the classical or multiplicative partitioning of gamma diversity: $\gamma = \alpha_d \times \beta_{Md}$. Subscript 'M' specifies that the components come from this multiplicative partitioning. It is important to be explicit about this, because also an additive partitioning is possible (to be covered in the next section). Subscript 'd' is related to the measurement units of diversity, and indicates that α is a measure of (mean) diversity concentration (e.g., effective number of species per compositional unit) and β_M is a measure of diversity (effective number of compositional units; notation as in Tuomisto 2010a). Alternatively, subscript 't' can be used to indicate that alpha is a (mean) diversity measured using the same unit as γ is (such as effective number of species). In this case, β_M is interpreted as a unitless ratio of two species diversities obtained at different scales of observation. The multiplicative partitioning of diversity has been one of the most widely used definitions of beta diversity in ecological studies, because it yields alpha and beta components that are conceptually and numerically independent of each other (e.g., Routledge 1977, 1979; Vellend 2001; Jost 2007, 2010; Baselga 2010a; de Bello et al. 2010; Tuomisto 2010a, c, 2011).

Gamma diversity is conceptually the simplest one of the three diversities: it is the total species diversity observed in the dataset of interest. The dataset may or may not represent a landscape, but whether it does is irrelevant for the definition of the concepts. To quantify gamma diversity for a dataset, the individuals of the dataset need to be classified in one way only, namely into species. This classification was already used above in *Defining diversity*, and it can be called the gamma classification (Tuomisto 2010a). Gamma diversity can hence be defined as the diversity in relation to the gamma classification, and identified by adding the subscript γ to eqn [1]:

$${}^q D_\gamma = 1 / {}^q \bar{p}_i \quad [10]$$

Species richness in relation to the gamma classification, i.e. the total number of actual species in the dataset of interest, can be called gamma richness R_γ . The symbols α , β and γ have been used to refer to both the diversity components and the corresponding richness components, and indeed it may be useful to continue using them in this broad sense, when the alpha, beta and gamma components are discussed in general terms without specifying whether the interest is in diversity or in richness. However, the distinction between diversity and richness becomes relevant as soon as their values are to be quantified. At this point, more specific notation should be used, such as ${}^q D_\alpha$, ${}^q D_\beta$ and ${}^q D_\gamma$ for the diversity components, and R_α , R_β and R_γ for the richness components.

Gamma diversity is the total number of effective species in a dataset, and gamma richness is the total number of actual species in the dataset. Both can be calculated whether the dataset has been subdivided into smaller units or not. In contrast, the alpha and beta components are defined only when the dataset consists of subunits. These subunits may or may not be the same sampling units that were used during the field inventory. It is possible to subdivide a

single dataset into subunits in many different ways, and each different subdivision corresponds to different alpha and beta values.

When quantifying alpha and beta diversity or richness, each individual is classified not only according to species but also according to subunit. The classification of individuals into subunits has been called the 'omega classification' (Tuomisto 2010a). The richness in relation to the omega classification (omega richness) is the actual number of subunits in the dataset. The diversity in relation to the omega classification (omega diversity) is the effective number of subunits in the dataset (the number of subunits the dataset would get divided into if each subunit received the same number of individuals as the actual subunits have, on average). Alpha and beta diversity (and alpha and beta richness) are obtained by taking into account both the gamma and the omega classifications simultaneously, i.e. by classifying each individual both into a species and into a subunit. This allows quantifying the number of compositionally non-overlapping subunits (the beta component) and the mean density of species per such subunit (the alpha component; Tuomisto 2010a, c, 2011).

Exactly the same idea has been expressed by Gregorius (2010) using a different notation. He characterized individuals by two attributes, type (T) and subcollection affiliation (S). The classification of individuals into types corresponds to the gamma classification, and the classification of individuals into subcollections corresponds to the omega classification. Gregorius (2010) used the notation v_T for gamma diversity (total species diversity in the dataset) and v_S for omega diversity (total subunit diversity in the dataset). The earliest definition of gamma diversity in these terms may be that of Routledge (1979), who used the notation $N_{a\gamma}$ (where a has the same meaning as q in ${}^qD_\gamma$).

2. Alpha, beta and gamma richness

The richness components are quantified without taking into account species proportional abundances. Alpha richness is simply calculated as the unweighted arithmetic mean of the actual species density values in the subunits

$$R_\alpha = \sum_{j=1}^N \frac{1}{N} R_{\alpha j} \quad [11]$$

Here N is the total number of subunits and $R_{\alpha j}$ is the actual species density (species richness per subunit) in subunit j . The multiplicative partitioning leads to defining beta richness as

$$R_\beta = R_\gamma / R_\alpha \quad [12]$$

This can be interpreted as the number of actual compositional units (CU) in the dataset. The actual (as opposed to effective; see the next section) compositional units are obtained by rearranging the R_γ actual species of the dataset evenly into new hypothetical subunits such that each subunit receives R_α species. In other words, each compositional unit has the same

species density as the original subunits do on average, but does not share species with any other compositional unit. Alpha richness equals the average number of actual species per actual compositional unit, which gives it the measurement unit sp/CU (Fig. 2; Tuomisto 2010c, 2011).

< Figure 2 near here >

3. Alpha, beta and gamma diversity

To quantify the diversity components, species proportional abundances need to be taken into account. The proportional abundance of species i that is conditional on the limits of subunit j is $p_{ij} = m_{ij}/m_j$, where m_{ij} is the number of individuals belonging to species i in subunit j and m_j is the total number of individuals in subunit j . The weighted generalised mean of all these within-subunit proportional abundances is

$${}^q\bar{p}_{ij} = {}^{q-1}\sqrt{\sum_{j=1}^N \sum_{i=1}^R p_{ij} p_{ij}^{q-1}} \quad [13]$$

Here N is the number of subunits. The proportion of data that individuals of species i in subunit j contribute to the entire dataset is p_{ij} . This is used as the nominal weight when calculating mean within-subunit proportional abundance (i.e., mean p_{ij}). The same mean can also be expressed (see Tuomisto 2010a for derivation)

$${}^q\bar{p}_{ij} = {}^{q-1}\sqrt{\sum_{j=1}^N w_j \sum_{i=1}^R p_{ij} p_{ij}^{q-1}} \quad [14]$$

Here the subunit weight $w_j = m_j/m$ equals the proportion of the individuals in the entire dataset contributed by subunit j . Alpha diversity is obtained as the inverse of this mean

$${}^qD_\alpha = 1/{}^q\bar{p}_{ij} \quad [15]$$

This can also be expressed (see Proof 2 in Tuomisto 2010a for derivation)

$${}^qD_\alpha = {}^{1-q}\sqrt{\sum_{j=1}^N w_j ({}^qD_{aj})^{1-q}} \quad [16]$$

This formula expresses the weighted generalised mean with exponent $1 - q$ of the mean effective species densities ${}^qD_{aj}$ that are first calculated separately for each subunit j .

The alpha diversity of Jost (2006, 2007) differs from the alpha diversity defined by eqn [16] in that Jost treated the gamma and the omega classifications symmetrically. This means that the subunit membership of an individual is considered ecologically as interesting as its species identity. In contrast, Routledge's definition (eqn [16]) considers the species identity of an individual to be of primary interest, and the subunit limits to be just a practical necessity

for the calculation of within-dataset heterogeneity. The latter corresponds better with what seemed to be Whittaker's original idea behind diversity partitioning, so it is followed throughout this article. The differences between the alpha diversity concepts of Routledge and Jost have been discussed in detail by Tuomisto (2010a) and Gregorius (2010). Although Jost's definition was favoured by Chao et al. (2012), it has the distinct disadvantage that under some conditions it produces alpha diversity values that exceed gamma diversity.

The multiplicative partitioning leads to defining beta diversity as

$${}^qD_\beta = {}^qD_\gamma / {}^qD_\alpha \quad [17]$$

This ratio can be interpreted as the number of effective compositional units (${}^q\text{CU}_E$). An effective compositional unit is a hypothetical subunit that has the same number (and density) of effective species as the original subunits do on average, but does not share effective species with the other effective compositional units. Alpha diversity equals the mean density of effective species as expressed on a per effective compositional unit basis, which gives it the measurement unit ${}^q\text{sp}_E/{}^q\text{CU}_E$ (Fig. 2; Tuomisto 2010a, c, 2011). In the notation of Routledge (1979), alpha diversity is $N_{a\alpha}$ and beta diversity is $N_{a\beta}$. In the notation of Gregorius (2010), alpha diversity is v_{TS} and beta diversity is N_{eS} .

4. Properties of beta richness and beta diversity

In any dataset, there must be at least one compositional unit (both actual and effective), so $R_\beta \geq 1$ CU and ${}^qD_\beta \geq 1$ ${}^q\text{CU}_E$. Similarly, the total number of species (both actual and effective) in the dataset must be at least as large as the mean number of species per subunit, so $R_\gamma \geq R_\alpha$ and ${}^qD_\gamma \geq {}^qD_\alpha$. Some researchers have complained that the latter condition does not always hold (Gadakar 1989; Lande 1996). However, this is an erroneous assertion based on averaging the effective species densities ${}^qD_{\alpha j}$ using the arithmetic mean, which corresponds to $q = 0$, even if some other value of q is used in the calculations otherwise. Using the generalized mean with exponent $1 - q$ (eqn [16]) correctly gives alpha diversity values that do not exceed gamma diversity for any value of q . Here it is worth recalling that Jost's definition of alpha diversity can exceed ${}^qD_\gamma$ when nominal subunit weights are unequal and q takes a value other than zero or unity. When this happens, Jost's beta diversity becomes smaller than unity. Nominal subunit weights are equal when, for example, sampling effort has been standardised to the same number of individuals in all subunits, or when abundances are recorded as percentages that sum to 100% in each subunit.

If all species occur in all subunits, beta richness (which is based on presence-absence data) necessarily equals one compositional unit. The more the original subunits differ in species composition, the more compositional units are needed (at a given species density) to

accumulate a given total number of species. Beta richness R_β obtains its maximum value of N when none of the actual subunits share any species.

Because beta diversity ${}^qD_\beta$ takes into account species abundances in addition to species composition, its behavior is more complex than that of beta richness. When the proportional abundances of the species vary among subunits, ${}^qD_\beta$ may exceed $1 {}^qCU_E$ even if all species occur in all subunits. In addition, when no species are shared among subunits, ${}^qD_\beta$ may be either smaller or larger than N . Beta diversity is constrained to the fixed range of $[1, N]$ only when either $q = 1$ or the nominal weights w_j in eqn [16] are equal for all subunits j (Gregorius 2010; Tuomisto 2010a). Again, this behaviour corresponds to Routledge's beta diversity; Jost's beta diversity is always constrained by the maximum value of N , but it can be smaller than unity if the nominal subunit weights are unequal when q has another value than zero or unity.

Increasing the value of q makes ${}^qD_\beta$ more sensitive to the variation among subunits in the proportional abundances of species and less sensitive to variation in species composition. Therefore, changing within-subunit species proportional abundances without changing presence-absence patterns or subunit weights has no effect on ${}^0D_\beta$, and changing species composition without changing the proportional abundances of the most abundant species has no effect on ${}^\infty D_\beta$ (Tuomisto 2010a).

Beta richness and beta diversity can be thought of as the richness and diversity in relation to the beta classification. Here the entities of interest are species (actual or effective, respectively) that are classified into compositional units (actual or effective, respectively; compare with the classification of individuals into actual or effective species in Fig. 1). Alpha richness and alpha diversity, in turn, correspond to the mean density of richness and diversity in relation to the gamma classification and at the same time conditional on the limits set by the beta classification. This combined classification can be called the alpha classification.

Whittaker (1977) referred to α and γ as inventory diversity and to β as differentiation diversity, which has since become a common practice (e.g., Magurran 2004, Jurasinski et al. 2009). Some researchers have even argued that beta diversity should not be called diversity at all (e.g. Lande 1996, Kiflawi and Spencer 2004, Gregorius and Gillet 2008). However, true beta diversity ${}^qD_\beta$ is the effective number of compositional units, which conforms with the general definition of diversity just as well as the effective number of species does.

Alpha, beta and gamma diversity are conceptually similar in that each quantifies the effective number of types in a dataset according to a specific kind of classification of the entities that constitute that dataset. The three differ in two ways. Firstly, they classify different entities: individuals in alpha and gamma diversity vs. effective species in beta diversity. Secondly, they quantify the relevant diversity at a different level: as a total diversity in the entire dataset in beta and gamma diversity vs. as a mean density of diversity as expressed per effective

compositional unit in alpha diversity. This conceptual similarity to ${}^qD_\alpha$ and especially ${}^qD_\gamma$ justifies singling ${}^qD_\beta$ out as the sole measure of true beta diversity, and recommending that all other things that have been referred to as "beta diversity" in the past be called something else in the future (Jost 2006, 2007; Tuomisto 2010a, b, c, 2011).

C. Additive partitioning: alpha + beta

1. Absolute species turnover

Lande (1996) proposed defining beta diversity as the difference rather than the ratio of gamma and alpha diversity, which leads to the equation

$$\beta_{At} = \gamma - \alpha_t \quad [18]$$

This approach corresponds to partitioning gamma diversity into additive components $\gamma = \alpha_t + \beta_{At}$. The subscript 'A' specifies that the beta component arises from additive partitioning, which is important because β_{At} differs from the beta component of the multiplicative partitioning β_{Md} both conceptually and in its numerical value. The subscript 't' specifies that the components relate to species turnover, which affects their measurement units: all components in the additive partitioning have the same measurement unit, namely species (actual species in the case of richness and effective species in the case of diversity).

Lande (1996) borrowed the additive approach from MacArthur (1964, 1965), who had used it with Shannon entropies. However, MacArthur applied the exponential function to the entropies before interpreting the results, and therefore he actually followed Whittaker's multiplicative approach with true diversities at $q = 1$. In contrast, Lande used and interpreted the raw values of different diversity indices, and therefore the components obtained from his partitioning correspond to conceptually different phenomena for each different diversity index (see Tuomisto 2010a for more detailed interpretations and discussion). The interpretations provided in the present chapter are all based on the assumption that any diversity indices are converted to the corresponding effective numbers of types (true diversity) before being subjected to additive partitioning.

The additive alpha component expresses how many species there are, on average, in one subunit. The additive beta component expresses how many more species the dataset as a whole has than the average subunit does. In other words, both components quantify a count of species (i.e. the types based on the gamma classification), but they do so for different subsets of the data (an average subunit vs. the rest). In contrast, both multiplicative components (α_d and β_{Md}) give a measure that describes the entire dataset, but they differ in that each is based on a different classification (the alpha or the beta classification, respectively; Tuomisto 2010a, c). In addition, α_d is conceptually a density measure (density

of types per subunit), which differs from β_{Md} , β_{At} , α_t and γ , all of which are simply counts of types.

If all observed species are not present in all subunits, there will be change in species composition, or species turnover, among subunits. Therefore, β_{At} can be interpreted as a measure of total species turnover in the dataset: if one were to start from one (actual or effective) compositional unit and walk through all the other compositional units, β_{At} would indicate how many (actual or effective) species are swapped for a new species along the way. Conceptually, every new species one encounters when moving to a new compositional unit equals half a species turned over, as does every old species that drops out. The additive beta component has recently become widely used in ecological studies (e.g., Veech et al. 2002; Kiflawi and Spencer 2004; Ricotta 2008; de Bello et al. 2010). Unfortunately, many studies refer to species turnover as ‘beta diversity’, even though the former is a count of species and the latter is a count of compositional units, and these are both conceptually and numerically very different things.

Equation [18] does not differentiate between presence-absence data (richness) and abundance data (diversity), and using the notation β for both species turnover and true beta diversity easily leads to confusing them. When the interest is in the turnover of actual species (actual species turnover), a more explicit notation is

$$R_{diff} = R_{tot} - R_{mean} \quad [19]$$

Here the subscript 'tot' refers to total richness in the dataset, 'mean' to mean richness in the subunits, and 'diff' to the difference between the dataset total and the subunit mean. The relevant mean here is the unweighted arithmetic mean. All components of the additive partitioning are based on the gamma classification, so all of them could have subscript γ . R_{tot} hence equals R_γ , but in the additive partitioning the components differ in which part of the dataset they refer to, rather than in which classification they are based on, so the subscript ' γ ' can be omitted to simplify the notation. When the interest is in the turnover of effective species (effective species turnover), the equation can be written

$${}^qD_{diff} = {}^qD_{tot} - {}^qD_{mean} \quad [20]$$

The relevant mean here is the weighted generalized mean with exponent $1 - q$ (calculated as in eqn [16]).

As with the multiplicative partitioning, the components of the additive partitioning behave in different ways depending on whether they are components of richness or of diversity. When all species are present in all subunits, R_{diff} invariably takes the value zero. In contrast, ${}^qD_{diff}$ does so only when either each species also has a constant abundance in all subunits, or when q is sufficiently large and the most abundant species is the same and has the same proportional abundance in all subunits.

The upper limit of absolute species turnover in a dataset depends both on the number of subunits N and on the mean species count within the subunits. If no subunits share any

species, actual species turnover R_{diff} necessarily takes the value $(N - 1)R_{\text{mean}}$. In the same situation, effective species turnover ${}^qD_{\text{diff}}$ takes the value $(N - 1){}^qD_{\text{mean}}$ only if either all nominal subunit weights w_j are equal (i.e. all subunits contain the same abundance either by design or after conversion of absolute abundances to proportions of site totals) or if $q = 1$ (see eqn [16]). Otherwise, the value that ${}^qD_{\text{diff}}$ takes when no subunits share any species cannot be predicted in advance, because it depends on dataset properties (Tuomisto 2010a, c). In any case, eqn [16], which is based on Routledge's concepts, guarantees that species turnover calculated with either eqn [18] or eqn [20] is always non-negative. This is not the case if Jost's equation for alpha diversity is used, because then alpha may exceed gamma. When this happens, species turnover (the beta component) becomes negative.

2. Whittaker's species turnover

In many cases, it is useful to quantify species turnover in relative terms, such that its value does not depend on the absolute number of species involved. Whittaker (1972) developed such a species turnover measure by subtracting unity from the value of multiplicative beta diversity. The same measure, Whittaker's species turnover, can be obtained by dividing the additive beta component by the additive alpha component (notation from Tuomisto 2010a; see also Kiflawi and Spencer 2004, but note that β_M refers to both β_{Mt} and β_{Mt-1} in their text)

$$\beta_{Mt-1} = \beta_{At}/\alpha_t = (\gamma - \alpha_t)/\alpha_t = \gamma/\alpha_t - 1 = \beta_{Mt} - 1 \quad [21]$$

Here β_{Mt} has the same numerical value as true beta diversity β_{Md} does, but the two differ in that β_{Mt} is unitless (since γ and α_t have the same measurement unit).

When the interest is in presence-absence data, Whittaker's actual species turnover can be expressed more explicitly

$$R'_W = (R_{\text{tot}} - R_{\text{mean}})/R_{\text{mean}} = R_{\text{tot}}/R_{\text{mean}} - 1 \quad [22]$$

The prime is added to specify that this is a relativised measure of richness, i.e. one richness is expressed in multiples of another richness, which yields a unitless value. Subscript 'W' refers to Whittaker's approach. When the interest is in abundance data, Whittaker's effective species turnover can be expressed

$${}^qD'_W = ({}^qD_{\text{tot}} - {}^qD_{\text{mean}})/{}^qD_{\text{mean}} = {}^qD_{\text{tot}}/{}^qD_{\text{mean}} - 1 \quad [23]$$

Whittaker's turnover measures express species turnover among the compositional units of the dataset in multiples of the within-unit species count (actual or effective). In other words, these measures quantify how many times the entire species composition turns over among the compositional units, rather than the number of species that turn over (as absolute species turnover does).

Whittaker's species turnover and absolute species turnover obtain their minimum and maximum values under the same conditions. However, the numerical values of the actual species turnover measures differ by factor R_{mean} and those of effective species turnover differ

by factor ${}^qD_{\text{mean}}$. The maximum value of Whittaker's actual species turnover (which is based on presence-absence data) is obtained when none of the N subunits share any species, and it is $N - 1$ [rather than $(N - 1)R_{\text{mean}}$ as in the case of absolute species turnover]. The maximum value of Whittaker's effective species turnover (which takes species abundances into account) is also $N - 1$, provided that the nominal subunit weights are equal or $q = 1$. In any case, the values are independent of the number of species involved: replicating each species (so that it gives rise to r new species of the same absolute abundance as the original) has no influence on Whittaker's species turnover. This allows comparing compositional heterogeneity among datasets in such a way that the comparisons are not confounded by differences in species richness or diversity. Since the absolute species turnover measures (R_{diff} and ${}^qD_{\text{diff}}$) do depend on the absolute number of species (actual or effective, respectively), they are not monotonically related to the corresponding Whittaker's species turnover values, and each measure can hence lead to a different ranking of datasets.

3. Proportional species turnover

To obtain Whittaker's species turnover, absolute species turnover (β_{At}) is divided by the mean species count in the subunits (α_t). Equally well, β_{At} can be divided by the total species count in the dataset (γ). Doing so leads to a new relative species turnover measure, *proportional species turnover* (Tuomisto 2010a):

$$\beta_{\text{Pt}} = (\gamma - \alpha_t)/\gamma = 1 - \alpha_t/\gamma \quad [24]$$

When the interest is in presence-absence data, proportional actual species turnover can be expressed more explicitly

$$R'_{\text{P}} = (R_{\text{tot}} - R_{\text{mean}})/R_{\text{tot}} = 1 - R_{\text{mean}}/R_{\text{tot}} \quad [25]$$

The prime is added to specify that this is a relativised measure of richness, i.e. one richness is expressed as a proportion of another richness, which yields a unitless value. Subscript 'P' refers to proportion. When the interest is in abundance data, proportional effective species turnover can be expressed

$${}^qD'_{\text{P}} = ({}^qD_{\text{tot}} - {}^qD_{\text{mean}})/{}^qD_{\text{tot}} = 1 - {}^qD_{\text{mean}}/{}^qD_{\text{tot}} \quad [26]$$

The proportional species turnover measures express what proportion of the species in the entire dataset is not contained in a single compositional unit (actual or effective). The term $R_{\text{mean}}/R_{\text{tot}}$ also indicates the proportion of subunits in which the average actual species occurs (mean species frequency; Whittaker 1972, Routledge 1977).

Proportional species turnover obtains its maximum and minimum values in the same situations as Whittaker's and absolute species turnovers do. Its minimum value is zero, and its maximum value is $1 - 1/N$ (for effective species turnover, the upper limit is fixed only if all nominal subunit weights are equal or $q = 1$). Just like Whittaker's species turnover, proportional species turnover is independent of how many species the dataset contains, and it

can be used to compare the compositional heterogeneities of datasets that differ in the number of species. Proportional species turnover is monotonically (but curvilinearly) related with Whittaker's species turnover, so the two measures will always rank datasets in the same way (assuming, of course, that both are either based on richness R or on diversity qD of the same order q). Proportional species turnover has been derived several times in the ecological literature (e.g., Roschewitz et al. 2005; Ricotta 2008; de Bello et al. 2010). Unfortunately, it has usually been called beta diversity, even though it is both conceptually and numerically very different from true beta diversity (the effective number of compositional units). The numerical difference is obvious from the fact that unity is the maximum value for proportional species turnover but minimum value for true beta diversity.

VI. Resemblance indices related to diversity

Several familiar similarity and dissimilarity indices can be derived as transformations of beta richness, beta diversity or actual or effective species turnover in a dataset consisting of two subunits (Jost 2006, 2007; Tuomisto 2010a, b, c). For example, the Manhattan metric, as calculated using presence-absence data, is a simple transformation of absolute actual species turnover (Tuomisto 2010a)

$$M = 2R_{\text{diff}} = 2(R_{\text{tot}} - R_{\text{mean}}) = b + c \quad [27]$$

Here b is the number of actual species unique to the first subunit and c is the number of actual species unique to the second. The Manhattan metric is not bound to any fixed interval, but can increase indefinitely as the number of species in the dataset increases. In some applications, this can be a desirable property, but in most situations where compositional resemblance indices are used, it is not. Therefore, in most cases it is more appropriate to use indices whose values are constrained to a fixed interval. Typically, compositional similarity indices take the value zero when the two subunits that are being compared share no species, and the value unity when both subunits have identical species compositions.

The Sørensen index equals the one-complement of Whittaker's actual species turnover (see Tuomisto 2010a for derivation)

$$C_S = 1 - R'_W = 2 - R_{\text{tot}}/R_{\text{mean}} = 2a/(2a + b + c) \quad [28]$$

Here a is the number of species found in both subunits, and b and c the numbers of species found in one or the other but not both subunits. The Sørensen index expresses the number of shared species as a proportion of the average number of species in the two subunits $R_{\text{mean}} = ([a + b] + [a + c])/2$.

The Sørensen index can easily be expanded to abundance data by replacing actual species turnover with effective species turnover

$${}^qC_S = 1 - {}^qD'_W = 2 - {}^qD_{\text{tot}}/{}^qD_{\text{mean}} \quad [29]$$

The addition of superscript q specifies that the calculations are based on diversity of order q . The classical (richness-based) Sørensen index is invariably constrained to the interval $[0, 1]$, which is often a desirable property in a similarity index. With the diversity-based Sørensen index, the minimum value is fixed at zero only when the maximum value of ${}^qD'_w$ is fixed at unity, which happens when both subunits have the same nominal weight (w_j) or $q = 1$. Using qC_S as a similarity index therefore needs to be done with care to avoid misinterpreting the results. If different subunits have different total abundances, and it is for some reason not justified or desirable to convert them to within-subunit percentages or rarefy them to the same number of individuals, then only $q = 1$ should be used when correct interpretation of the results depends on qC_S being restricted to the interval $[0, 1]$.

The Jaccard index equals

$$C_J = 1 - 2R'_p = 2R_{\text{mean}}/R_{\text{tot}} - 1 = a/(a + b + c) \quad [30]$$

The Jaccard index expresses the number of shared species as a proportion of the total number of species in the two subunits $R_{\text{tot}} = a + b + c$. Just as the Sørensen index, the Jaccard index can be adapted to abundance data

$${}^qC_J = 1 - 2{}^qD'_p = 2{}^qD_{\text{mean}}/{}^qD_{\text{tot}} - 1 \quad [31]$$

As with the diversity-based Sørensen index, the values of the diversity-based Jaccard index lie in the exact interval $[0, 1]$ only if both subunits contain the same abundance or $q = 1$.

In some cases, it may be desirable to partition a similarity index (or the corresponding dissimilarity index) into components that correspond to different sources of species turnover, such as richness difference, one-to-one species replacement or nestedness. At least two different approaches have been proposed to partition the Jaccard and Sørensen dissimilarities (Baselga 2010b; Podani & Schmera 2011). This has been followed by active discussion on which of the partitioning methods is more appropriate, and what the obtained components mean in practice (Schmera & Podani 2011; Podani et al. 2013; Legendre 2014; Baselga & Leprieur 2015).

Since both richness and diversity can be calculated for any dataset, irrespective of the number of subunits it contains, both the Sørensen index and the Jaccard index can be generalized to more than two subunits. The Jaccard index equals the one-complement of proportional species turnover between two subunits as ranged to the interval $[0, 1]$ (obtained by dividing with the maximum value). Its values therefore naturally stay in this range even when $N > 2$. The multiple-site Jaccard index for presence-absence data can be written (Tuomisto 2010a)

$$C_{JN} = 1 - \frac{R_p}{1 - 1/N} = \frac{R_{\text{mean}}/R_{\text{tot}} - 1/N}{1 - 1/N} \quad [32]$$

When $N = 2$, the Sørensen index equals the one-complement of Whittaker's species turnover, but the value of the latter can exceed unity when $N > 2$ so ranging is needed to keep its value within the interval $[0, 1]$. The multiple-site Sørensen index for presence-absence data then becomes (Diserud and Ødegaard 2007; Tuomisto 2010a)

$$C_{SN} = 1 - \frac{R_w}{N-1} = \frac{N - R_{\text{tot}} / R_{\text{mean}}}{N-1} \quad [33]$$

The one-complements of both C_{JN} and C_{SN} quantify how much the actual species turnover of the corresponding kind exceeds its minimum possible value, expressed as a proportion of the total possible range of values (given N). The one-complement of C_{SN} was proposed by Harrison et al. (1992) under the name 'beta-1' (although their original equation gives values in percentages rather than proportions).

For effective species turnover, the maximum values are definable in terms of N only when all nominal subunit weights (w_j) are equal or $q = 1$, and ranging is therefore possible only when at least one of these conditions holds. Then the multiple-site Sørensen index for abundance data is (Tuomisto 2010a)

$${}^q C_{SN} = 1 - \frac{{}^q D_w}{N-1} = \frac{N - {}^q D_{\text{tot}} / {}^q D_{\text{mean}}}{N-1} \quad [34]$$

Similarly, the multiple-site Jaccard index for abundance data is (Jost 2006, Tuomisto 2010a)

$${}^q C_{JN} = 1 - \frac{{}^q D_p}{1 - 1/N} = \frac{{}^q D_{\text{mean}} / {}^q D_{\text{tot}} - 1/N}{1 - 1/N} \quad [35]$$

The one-complements of both indices quantify how much the effective species turnover of the corresponding kind exceeds its minimum possible value, expressed as a proportion of the total possible range of values (given N).

Binary measures can be generalized to abundance data in different ways, and a particular non-linear (but monotonic) transformation of diversity gives the following overlap index (Chao et al. 2008):

$$C_{qN} = \frac{({}^q D_{\text{mean}} / {}^q D_{\text{tot}})^{q-1} - (1/N)^{q-1}}{1 - (1/N)^{q-1}} \quad [36]$$

At $q < 1$, the exponents become negative and hence cause all terms to be inverted. How this affects the calculation can be seen by rewriting the equation in the mathematically equivalent form that has positive exponents when $q < 1$

$$C_{qN} = \frac{N^{1-q} - ({}^q D_{\text{tot}} / {}^q D_{\text{mean}})^{1-q}}{N^{1-q} - 1} \quad [37]$$

Note that the parameter q in C_{qN} is specified in a subscript rather than a superscript. This distinction is important, because even though C_{qN} is expressed as a function of true diversities of order q , it is not linearly related with ${}^q D$ but with the basic sum ${}^q D^{1-q}$ (see the explanation in connection with eqn 6). The C_{qN} index is therefore a measure of compositional overlap (Chao et al. 2008), not of relative species turnover in the same way as ${}^q C_{SN}$ and ${}^q C_{JN}$ are. As a result, C_{qN} behaves in a different way than ${}^q C_{SN}$ and ${}^q C_{JN}$ do when the value of q is changed.

In the special case of $q = 0$, the C_{qN} index converges on the ${}^qC_{SN}$ index (compare eqns [37] and [34]), and both thus converge on the classical Sørensen index when $N = 2$ (see eqns [29] and [28]). In the special case of $q = 2$, the C_{qN} index converges on the ${}^qC_{JN}$ index instead (compare eqns [36] and [35]), and both converge on the Morisita-Horn index when $N = 2$. The Morisita-Horn index generalized to N subunits can be written (see Jost 2006 for derivation)

$$C_{2N} = \frac{{}^2D_{\text{mean}} / {}^2D_{\text{tot}} - 1/N}{1 - 1/N} \quad [38]$$

In the special case of $q = 1$ and $N = 2$, the C_{qN} index converges on the Horn index (Chao et al. 2008). The Horn index generalized to N subunits can be written (see Tuomisto 2010a for derivation)

$$C_{1N} = \frac{\log(N) - \log({}^1D_{\text{tot}} / {}^1D_{\text{mean}})}{\log(N)} \quad [39]$$

The C_{qN} index provides a valid and useful measure of overall compositional overlap among all sampling units in a dataset of interest at all values of q . At some values of q , it can also be interpreted in terms of effective species turnover. At $q = 0$, C_{qN} equals Whittaker's species turnover expressed as a proportion of the maximum value obtainable, given N . At $q = 2$, C_{qN} equals proportional species turnover expressed on a relative scale between the minimum and maximum possible, given N . However, C_{qN} is not primarily a species turnover measure, and at other values of q interpreting the value of C_{qN} in terms of species turnover is misleading. For example, at $q = 1$, C_{qN} quantifies the amount of relative Shannon entropy, not relative species turnover. Conversely, the value of ${}^qC_{SN}$ can be interpreted in terms of relative Whittaker's species turnover at any value of q , but in terms of compositional overlap only at $q = 0$; and the value of ${}^qC_{JN}$ can be interpreted in terms of relative proportional species turnover at any value of q , but in terms of compositional overlap only at $q = 2$.

In order to choose an appropriate index for a particular study, it is necessary to understand which aspect of the data each index quantifies. Ecological hypotheses invariably concern some specific aspect of data behavior, and choosing an index that corresponds to the wrong aspect can therefore lead to incorrect interpretations of the results and unjustified conclusions about the tested hypotheses. The mathematical properties of the indices also determine what needs to be taken into account when comparing their values among datasets, and indeed whether the index values can meaningfully be compared at all. How an index is calculated is not a mere technical detail, because it defines the variable of interest and hence the ecological meaning of the results.

References

Alatalo, R. V. (1981). Problems in the measurement of evenness in ecology. *Oikos* 37, 199–204.

- Anderson, M. J., Crist, T. O., Chase, J. M. et al. (2011). Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters* 14, 19–28.
- Baselga, A. (2010a) Multiplicative partition of true diversity yields independent alpha and beta components; additive partition does not. *Ecology*, **91**, 1974–1981.
- Baselga, A. (2010b) Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, **19**, 134–143.
- Baselga, A. & Leprieux, F. (2015) Comparing methods to separate components of beta diversity. *Methods in Ecology and Evolution*, 6: 1069–1079.
- Beck, J. and Schwanghart, W. (2010). Comparing measures of species diversity from incomplete inventories: an update. *Methods in Ecology and Evolution* 1, 38–44.
- Chao, A., Chazdon, R. L., Colwell, R. K. and Shen, T.-J. (2005). A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters* 8, 148–159.
- Chao, A., Chiu, C.-H. & Hsieh, T.C. (2012) Proposing a resolution to debates on diversity partitioning. *Ecology*, **93**, 2037–2051.
- Chao, A., Chiu, C.-H. & Jost, L. (2014) 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., Gotelli, N.J., Hsieh, T.C., Sander, E.L., Ma, K.H., Colwell, R.K. & Ellison, A.M. (2013) 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., Chiang, S. C., Jiang, Y.-H. and Chazdon, R. L. (2008). A two-stage probabilistic approach to multiple-community similarity indices. *Biometrics* 64, 1178–1186.
- Chao, A. & Jost, L. (2015) Estimating diversity and entropy profiles via discovery rates of new species. *Methods in Ecology and Evolution*, **6**, 873–882.
- Colwell, RK and Coddington JA (1994) Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions: Biological Sciences* 345: 101–118.
- de Bello, F., Lavergne, S., Meynard, C. N., Lepš, J. and Thuiller, W. (2010). The partitioning of diversity: showing Theseus a way out of the labyrinth. *Journal of Vegetation Science* 21: 992–1000.
- Diserud, O. and Ødegaard, F. (2007). A multiple-site similarity measure. *Biology Letters* 3, 20–22.
- Gadagkar, R. (1989). An undesirable property of Hill's diversity index N_2 . *Oecologia* 80, 140–141.
- Gauthier, O., Sarrazin, J. and Desbruyères, D. (2010). Measure and mis-measure of species diversity in deep-sea chemosynthetic communities. *Marine Ecology Progress Series* 402, 285–302.
- Good, I. J. and Toulmin, G. H. (1956). The number of new species, and the increase in population coverage, when a sample is increased. *Biometrika* 43, 45–63.

- Gorelick, R. (2011). Commentary: Do we have a consistent terminology for species diversity? The fallacy of true diversity. *Oecologia* 167: 885–888.
- Gotelli, N. J. and Colwell, R. K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* 4, 379–391.
- Gregorius, H.-R. (2010). Linking diversity and differentiation. *Diversity* 2, 370–394.
- Gregorius, H.-R. and Gillet, E. M. (2008). Generalized Simpson-diversity. *Ecological Modelling* 211, 90–96.
- Harrison, S., Ross, S. J. and Lawton, J. H. (1992). Beta diversity on geographic gradients in Britain. *Journal of Animal Ecology* 61, 151–158.
- Hill, M.O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology* 54, 427–432.
- Hoffmann, S., Hoffmann, A. (2008). Is there a "true" diversity? *Ecological Economics* 65, 213–215.
- Hurlbert, S. H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 52, 577–586.
- 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.
- Jost, L. (2009). Mismeasuring biological diversity: Response to Hoffmann and Hoffmann (2008). *Ecological Economics* 68, 925–928,
- Jost, L. (2010). The relation between evenness and diversity. *Diversity* 2, 207–232.
- Jost, L. (2010) Independence of alpha and beta diversities. *Ecology*, **91**, 1969–1974.
- Jurasinski, G., Retzer, V. and Beierkuhnlein, C. (2009). Inventory, differentiation, and proportional diversity: a consistent terminology for quantifying species diversity. *Oecologia* 159, 15–26.
- Jurasinski, G. and Koch, M. (2011). Commentary: do we have a consistent terminology for species diversity? We are on the way. *Oecologia* 167: 893–902.
- Kiflawi, M. and Spencer, M. (2004). Confidence intervals and hypothesis testing for beta diversity. *Ecology* 85, 2895–2900.
- Kindt, R., van Damme, P. and Simons, A. J. (2006). Tree diversity in western Kenya: using profiles to characterise richness and evenness. *Biodiversity and Conservation* 15, 1253–1270.
- Koleff, P., Gaston, K. E. and Lennon, J. J. (2003). Measuring beta diversity for presence–absence data. *Journal of Animal Ecology* 72, 367–382.
- Lande, R. (1996). Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* 76, 5–13.
- Legendre, P. (2014) Interpreting the replacement and richness difference components of beta diversity. *Global Ecology and Biogeography*, **23**, 1324–1334.
- MacArthur, R. H. (1964). Environmental factors affecting bird species diversity. *American Naturalist* 98, 387–397.
- MacArthur, R. H. (1965). Patterns of species diversity. *Biological Reviews* 40, 510–533.
- Magurran, A. E. (2004). *Measuring biological diversity*. Oxford, Blackwell Publishing.

- Mendes, R.S., Evangelista, L.R., Thomaz, S.M., Agostinho, A.A. and Gomez, L.C. (2008). A unified index to measure ecological diversity and species rarity. *Ecography* 31, 450–456.
- Morante-Filho, J.C., Arroyo-Rodríguez, V. & Faria, D. (2016) Patterns and predictors of β -diversity in the fragmented Brazilian Atlantic forest: a multiscale analysis of forest specialist and generalist birds. *Journal of Animal Ecology*, **85**, 240–250.
- Moreno, C.E. & Rodríguez, P. (2010) A consistent terminology for quantifying species diversity? *Oecologia*, **163**, 279–282.
- Moreno, C. and Rodríguez, P. (2011). Commentary: Do we have a consistent terminology for species diversity? Back to basics and toward a unifying framework. *Oecologia* 167: 889–892.
- Peet, R. K. (1974). The measurement of species diversity. *Annual Review of Ecology and Systematics* 5, 285–307.
- Plotkin, J. B. and Muller-Landau, H. C. (2002). Sampling the species composition of a landscape. *Ecology* 83, 3344–3356.
- Podani, J., Ricotta, C. & Schmera, D. (2013) A general framework for analyzing beta diversity, nestedness and related community-level phenomena based on abundance data. *Ecological Complexity*, **15**, 52–61.
- Podani, J. & Schmera, D. (2011) A new conceptual and methodological framework for exploring and explaining pattern in presence - absence data. *Oikos*, **120**, 1625–1638.
- Rényi, A. (1960). On measures of entropy and information. *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability* 1960, 547–561.
- Ricotta, C. (2008). Computing additive β -diversity from presence and absence scores: A critique and alternative parameters. *Theoretical Population Biology* 73, 244–249.
- Ricotta, C. and Avena, C. G. (2002). On the information-theoretical meaning of Hill's parametric evenness. *Acta Biotheoretica* 50, 63–71.
- Roschewitz I, Gabriel D, Tscharnkte T, and Thies C (2005) The effects of landscape complexity on arable weed species diversity in organic and conventional farming. *Journal of Applied Ecology* 42: 873–882.
- Routledge, R. D. (1977). On Whittaker's components of diversity. *Ecology* 58, 1120–1127.
- Routledge, R. D. (1979). Diversity indices: which ones are admissible? *J. Theor. Biol.* 76, 503–515.
- Schmera, D. & Podani, J. (2011) Comments on separating components of beta diversity. *Community Ecology*, **12**, 153–160.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal* 27, 379–423.
- Simpson, E. H. (1949). Measurement of diversity. *Nature* 163, 688.
- Smith, B. and J. B. Wilson. (1996). A consumer's guide to evenness indices. *Oikos* 76, 70–82.
- 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.

- Tuomisto, H. (2010c). A consistent terminology for quantifying species diversity? Yes, it does exist. *Oecologia* 164, 853–860.
- Tuomisto, H. (2011). Commentary: do we have a consistent terminology for species diversity? Yes, if we choose to use it. *Oecologia* 167, 903–911.
- Tuomisto H (2012) An updated consumer's guide to evenness and related indices. *Oikos* 121: 1203–1218.
- Tuomisto, H., Zuquim, G. & Cárdenas, G. (2014) Species richness and diversity along edaphic and climatic gradients in Amazonia. *Ecography*, **37**, 1034–1046.
- Veech, J. A., Summerville, K. S., Crist, T. O. and Gering, J. C. (2002). The additive partitioning of species diversity: recent revival of an old idea. *Oikos* 99, 3–9.
- Vellend, M. (2001). Do commonly used indices of β -diversity measure species turnover? *Journal of Vegetation Science* 12, 545–552.
- Whittaker, R. H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30, 279–338.
- Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon* 21, 213–251.
- Whittaker, R. H. (1977). Evolution of species diversity in land communities. *Evolutionary Biology* 10, 1–67.
- Wilson, M. V. and Shmida, A. (1984). Measuring beta diversity with presence-absence data. *Journal of Ecology* 72, 1055–1064.

Figures

Fig. 1. The calculation of species richness and species diversity in a small dataset consisting of a single tree plot. Species richness is the number of boxes (the set at left) needed to place all individuals into a box with an appropriate species name. Species diversity is the number of new boxes needed (the sets at right) to place all individuals into a box such that each box receives as many individuals (or as large a proportion of all individuals) as the named species have, on average (only a part of the last new box may be needed). The measure of 'average' here is the weighted generalized mean with exponent $q - 1$. This can be calculated by first taking the weighted mean of the absolute species abundances (= 6, 7.7 and 9.1 individuals for $q = 0, 1$ and 2 , respectively) and then dividing this by total abundance (18 ind.). Equally well, one can calculate the weighted mean of the proportional abundances directly (as shown in the equation of the generalized mean; R is the number of actual named species and p_i is the proportional abundance of the i th species). Species diversity qD equals the inverse of mean p_i , and it is the effective number of species (= the number of equally-abundant species that would give the observed mean species abundance). The measurement unit is hence actual species (sp) in the case of richness and effective species (${}^q\text{sp}_E$) in the case of diversity. With kind permission from Springer Science+Business Media: fig. 1 from Tuomisto 2011.

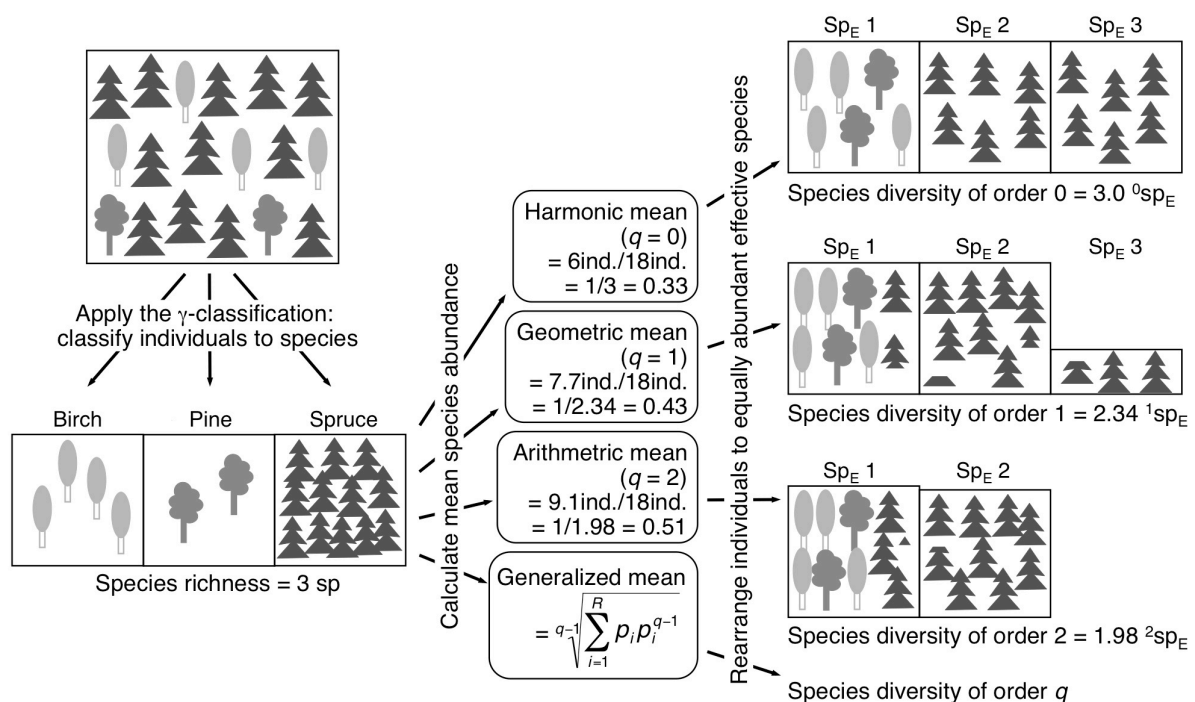


Fig. 2. Two different ways of partitioning the total species diversity of a dataset (true gamma diversity) into two components, given three subunits. The dataset is the same as in Fig. 1, but only true diversities of order one (which are based on the geometric mean of species proportional abundances) are shown for simplicity. Species diversity in each subunit j (1D_j) is calculated in the same way as species diversity in the entire dataset (${}^1D_j = {}^1D_{\text{tot}}$; see Fig. 1). Multiplicative partitioning gives two conceptually and mathematically independent components, namely the effective number of compositionally distinct units (true beta diversity; measurement unit ${}^1\text{CU}_E$), and the mean effective species density in those units (true alpha diversity). Additive partitioning gives two components that represent the same concept as true gamma diversity does, namely species diversity. One of these components quantifies the mean within-subunit species diversity (${}^1D_{\text{mean}}$), and the other (${}^1D_{\text{diff}}$) quantifies the total number of effective species that differ (turn over) among all subunits. With kind permission from Springer Science+Business Media: fig. 2 from Tuomisto 2011.

