

A consistent terminology for quantifying species diversity? Yes, it does exist

Hanna Tuomisto

Received: 2 June 2010 / Accepted: 5 October 2010 / Published online: 27 October 2010
© Springer-Verlag 2010

Abstract The prevailing terminological confusion around the concept ‘diversity’ has hampered accurate communication and caused diversity issues to appear unnecessarily complicated. In fact, a consistent terminology for phenomena related to (species) diversity is already available. When this terminology is adhered to, diversity emerges as an easily understood concept. It is important to differentiate between diversity itself and a diversity index: an index of something is just a surrogate for the thing itself. The conceptual problem of defining diversity also has to be separated from the practical problem of deciding how to adequately quantify diversity for a community of interest. In practice, diversity can be quantified for any dataset where units of observation (such as individuals) have been classified into types (such as species). All that needs to be known is what proportion of the observed units belong to a type of mean abundance. Diversity equals the inverse of this mean, and it quantifies the effective number of the types of interest. In ecology, interest often (but not always) focuses on species diversity. If the dataset consists of (or gets divided into) subunits, then the total effective number of species (gamma diversity) can be partitioned into the effective number of compositionally distinct subunits (beta diversity) and the mean effective number of species per such subunit (alpha diversity). Species richness is related to species diversity, but they are not the same thing; richness does not take the proportional abundances into account and is therefore the actual—rather than the effective—number

of types. Most of the phenomena that have been called ‘beta diversity’ in the past do not quantify an effective number of types, so they should be referred to by names other than ‘diversity’ (for example, species turnover or differentiation).

Keywords Alpha diversity · Beta diversity · Gamma diversity · Species richness · Species turnover

Introduction

In a recent paper, Moreno and Rodríguez (2010; hereafter referred to as MR) voiced an opinion that is surely shared by many: there is a genuine need for consensus on a clear, standard terminology related to species diversity. Jurasinski et al. (2009; hereafter JRB) had attempted to provide such a terminology, but MR found their proposal unsatisfactory. MR especially criticised the manner in which JRB classified different kinds of diversities, which involved lumping alpha and gamma diversity under the same heading (inventory diversity) and dividing beta diversity into just two categories (differentiation diversity and proportional diversity; both were further divided into a few subcategories). As an alternative, MR proposed that alpha and gamma diversities be maintained separate and that beta diversity be divided into three categories.

Neither JRB nor MR seem to have recognised just how confused the situation has become. Numerous indices of diversity have been popular in ecology (reviewed in Magurran 2004), but little attention has been given to the conceptual differences among those indices. Indeed, the term ‘diversity’ has been used in at least four conceptually different ways in the ecological literature, primarily because indices of diversity have been equated with

Communicated by Scott Collins.

H. Tuomisto (✉)
Department of Biology, University of Turku,
20014 Turku, Finland
e-mail: hanna.tuomisto@utu.fi

diversity itself. Furthermore, an alpha component, or ‘alpha diversity’, has been separated from total or ‘gamma diversity’ in at least three different ways (Tuomisto 2010a). The situation is even worse with ‘beta diversity’, which has been defined in more than 30 different ways; some of these are not mathematically derived from alpha and gamma diversity in any way, and the values of many are uncorrelated with each other (reviewed in Tuomisto 2010a, b). Although most of the resulting measures are useful for addressing some ecological questions, each one of them quantifies a different phenomenon and emphasises different aspects of the data used to calculate them. Calling so many different phenomena ‘diversity’ or ‘beta diversity’ has caused a lot of confusion, unwarranted comparisons among studies and incorrect inferences from data.

In this paper, I discuss the conceptual definition of diversity and show that a logical terminology for diversity studies is already available. I will also comment on a number of other issues related to diversity that were touched upon by JRB and MR.

Defining diversity

Much of the discussion on how to define diversity has actually focused on how to accurately estimate diversity for a community of interest. For example, it has often been recommended that diversity is best measured using diversity indices whose values stabilise at small sample sizes (e.g. Lande 1996; Magurran 2004; Beck and Schwanghart 2010). According to this criterion, species richness is a very bad index of diversity, the Shannon index is clearly better and the Gini–Simpson index is superior to both. However, for an index of diversity to be useful, we need to know what it is an index of, which in turn requires a definition of diversity itself. Such a conceptual definition of a scientific term should have two properties. Firstly, accuracy of scientific communication requires that terms be unambiguous, so only one phenomenon should be referred to as diversity, and other phenomena should be referred to by other names. The three indices mentioned above quantify conceptually different phenomena: species richness is a count of species, the Shannon index is an entropy, and the Gini–Simpson index is a probability. Therefore, all three measures cannot equal diversity itself, although each can be used as an index of diversity. Secondly, 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. Nevertheless, few diversity studies seem to have made an explicit distinction between the conceptual definition of a phenomenon and its practical measurement (but see Shmida and Wilson 1985).

Imagine we were discussing volume instead of diversity. Volume has an unambiguous definition, and any two persons discussing volume are likely to think about the same phenomenon. Nevertheless, two persons measuring the volume of a flock of birds might obtain very different results depending on when they choose to make the measurement and how they choose to define the physical limits of the flock. The number of birds in the flock would be much easier to quantify accurately, and the bird count would also be a more relevant variable for many ecological questions. However, this is no justification to expand the definition of volume so as to include the bird count. Combining such different phenomena under the single term volume would just create confusion and hence diminish the value of the term.

It can be difficult to measure the physical limits of a flock of birds, but once this has been done, calculating the volume of the flock is a simple mathematical operation. The same principle applies when the variable of interest is species diversity. It can be difficult to decide which individuals belong to the flock and which species each individual belongs to, but once this has been done, calculating the species diversity of the flock is a simple mathematical operation. The accuracy of the result is in both cases affected by sampling problems, but the definition of the variable to be measured (whether volume or diversity) is not.

Volume can be measured in, for example, litres or gallons, but what are the units of measurement of diversity? Imagine an inventory dataset consisting of a total abundance of m observed individuals (some other unit of abundance could also be used, such as grams of biomass or square metres of cover). If species diversity is of interest, each individual that belongs to the dataset needs to be identified to a species. Let us assume there is a wall with pigeonholes and that each individual is placed into the pigeonhole labeled with its species name. Species richness (S) of the dataset can then be quantified simply by counting the number of pigeonholes that were needed to give each observed species its own pigeonhole. Each pigeonhole represents one richness unit, so the measurement unit of species richness is species (sp). If the classification of individuals into pigeonholes had been based on genus identity or growth form identity rather than species identity, the richness unit would be genus or growth form, respectively.

In most real datasets, species differ in abundance, so some pigeonholes become more crowded with individuals than others. The concept of richness does not take this into account, but the concept of diversity does. To quantify species diversity, we therefore need to know what proportion of all individuals is contained in each pigeonhole. The proportional abundance of species i is quantified as $p_i = m_i/m$, where m_i is the number of individuals classified

into species i , and m is the total number of individuals. The weighted mean of these values equals $\bar{p}_i = \bar{m}_i/m$ (what the weights are and how they affect the mean will be explained in a moment). The species diversity of the dataset is the inverse of this mean, $D = 1/\bar{p}_i$. This equals the effective number of species in the dataset, which can be visualised as follows.

Imagine a new set of pigeonholes on a second wall and move all of the individuals from the pigeonholes on the first wall into those on the second wall such that each new pigeonhole receives exactly \bar{m}_i individuals. Each one of the new pigeonholes then contains the same proportion of the individuals as the original pigeonholes did on average (\bar{p}_i). Species diversity of the dataset can now be quantified simply by counting the number of new pigeonholes that were needed to accommodate all m individuals (the last pigeonhole may only be needed in part). The measurement unit of species diversity is effective species (sp_E) rather than actual species (sp) as in the case of species richness. This is because the new pigeonholes do not correspond to the actual named species that were observed during the inventory, but to hypothetical equally abundant species, each of which has the same proportional abundance as the actual species did on average. If all actual species are equally abundant, then there are as many effective species as actual species. Otherwise, there are fewer effective than actual species.

The same principle can be applied to quantifying the diversity of any types into which units of observation may be classified. For example, people attending a meeting can be classified by native language, and atoms in a container can be classified into chemical elements. The measurement unit of the obtained diversity will then change accordingly, such that linguistic diversity equals the effective number of languages and chemical diversity the effective number of chemical elements. This concept of diversity was apparently first used in ecology by MacArthur (1964, 1965), but it has been more thoroughly discussed and advocated by others (e.g. Hill 1973; Routledge 1977, 1979; Jost 2006, 2007, 2009; Tuomisto 2010a, b). Since Hill’s influential paper on the topic, the diversity values have been known as ‘Hill numbers’.

Hill (1973) provided the first thorough discussion on the concept of species diversity and explicitly described diversity as the inverse of mean species proportional abundance. Most diversity indices are very abstract, and their practical meaning is difficult to understand, but Hill’s (1973) diversity is simply what was visualised with pigeonholes above. The general definition of diversity can be written

$${}^qD = 1/{}^q\bar{p}_i \tag{1}$$

Because a mean can be calculated in different ways, it is always necessary to specify which mean is being used.

Hill’s definition of diversity is based on the weighted generalised mean with exponent $q - 1$. The order of the diversity equals q , and this is made explicit in Eq. (1) by showing q as a superscript on both sides of the equation. The generalised mean is calculated as follows:

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

For example, at $q = 0$, the mean becomes the harmonic mean, and its inverse gives diversity of order zero; at $q = 1$, the geometric mean yields diversity of order one; and at $q = 2$, the arithmetic mean yields diversity of order two. In all cases, the p_i^{q-1} term is nominally weighted by p_i itself. In this way, each species contributes to the value of the mean in proportion to the amount of data it contributed to the dataset. The effect of the parameter q is to modify the weights given to abundant versus rare species. Small values ($q < 1$) de-emphasise the abundance differences among species (and hence give rare species more weight than implied by their p_i), whereas large values ($q > 1$) exaggerate the abundance differences (and hence give abundant species more weight). At $q = 0$, all species are treated as if they were equally abundant no matter what their actual proportional abundances, and hence ${}^0\bar{p}_i$ equals $1/S$. Unless all p_i values are identical, increasing q causes ${}^q\bar{p}_i$ to asymptotically approach the p_i value of the most abundant species, which causes mean p_i to increase—and hence diversity qD to decrease. Negative values of q give the rarest species most weight, which causes the number of effective species to exceed the number of actual species; in practice, q can therefore be limited to non-negative values. The interpretation of qD does not change with q , however; it always quantifies the effective number of species (or other types of interest) in the dataset.

By combining Eqs. (1) and (2), we obtain

$${}^qD = 1/{}^q\bar{p}_i = 1 / \sqrt[q-1]{\sum_{i=1}^S p_i p_i^{q-1}} \tag{3}$$

which is usually written

$${}^qD = \left(\sum_{i=1}^S p_i^q \right)^{1/(1-q)} \tag{4}$$

The term inside the parentheses is known as the basic sum (Jost 2006, 2007). When $q = 0$, the basic sum equals 0D and has the same numeric value as species richness S . When $q = 2$, the basic sum equals the original Simpson index, and 2D is obtained as its inverse; the Gini-Simpson index is the one-complement of the basic sum. When $q = 1$, the basic sum itself cannot be used as a diversity index because it equals unity by definition, but $\log({}^1D)$ is

the Shannon index (more details on how qD relates to diversity indices can be found in Hill 1973; Jost 2006, 2007; Tuomisto 2010a). From Eqs. (3) and (4), it can be seen that the basic sum is just an intermediate step in calculating mean proportional species abundance and hence diversity itself. So there is, in fact, no need to use the basic sum to derive an index of diversity; 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 thereof.

Indeed, each one of the traditional diversity indices represents a conceptually different phenomenon (Tuomisto 2010a). The Shannon index $\log({}^1D)$ quantifies the uncertainty in the species identity of an individual that is picked at random from the dataset. The Gini–Simpson index $1 - 1/{}^2D$ quantifies the probability that two individuals picked at random from the dataset do not represent the same species. In contrast, qD quantifies the effective number of species present in the dataset, which is something quite different. Out of the three measures, qD corresponds most closely to an intuitive sense of what diversity is; also mathematically, this measure behaves as one would expect diversity to behave. Consequently, Jost (2006, 2007) proposed that qD be called ‘true diversity’ to distinguish it from the other phenomena that have also been called “diversity” in the ecological literature, and I have adopted this convention (Tuomisto 2010a, b).

Of course, for some ecological questions it may be more relevant to quantify entropy (as represented by the Shannon index) or probability (as represented by the Gini–Simpson index) rather than true diversity. It should then be stated explicitly what the phenomenon of interest is, rather than hiding this behind the vague umbrella term ‘index of diversity’. Analogously, both the radius and the circumference of a sphere can be used as indices of its volume, but doing so may be more confusing than helpful (Jost 2006). Incorrect interpretations are hard to avoid in this case, especially if the term ‘index of volume’ is used interchangeably with ‘volume’. If radius and circumference had already been called “volume” in the literature, one might start to use the term ‘true volume’ in order to specify that one is interested in volume itself, not radius or circumference.

Even though Hill (1973) has been widely cited (1,007 citations in ISI Web of Science as of 04 October 2010), true diversity seems to have been much less used in the ecological literature than the diversity indices. Jost (2006, 2007, 2009) has made a great contribution to the field by evaluating the properties of true diversity and pointing out how easily diversity index values are misinterpreted. Jost argued that traditional diversity indices are superfluous: to be interpreted correctly, they need to be converted to their

numbers equivalents, and since the latter equal true diversity, the diversity indices themselves are unnecessary. I would take the argument one step further: because diversity indices prevent us from capitalising on the conceptual simplicity of Hill’s definition of diversity itself, they are actually counter-productive. In order to understand the concept of diversity, all we really need to understand is what the inverse of a proportional abundance stands for. It is unnecessary to complicate the matter by invoking entropies, probabilities and their numbers equivalents.

Alpha, beta and gamma diversity

JRB followed Whittaker (1977) in applying the term inventory diversity to both alpha and gamma diversity. In contrast to Whittaker, JRB argued that the separation between alpha and gamma is unnecessary because both quantify the same phenomenon, only at different spatial scales. This was criticised by MR, who agreed with Whittaker that the differences in scale and the associated degree of heterogeneity are so important that alpha and gamma diversity should be maintained separate. I agree with this conclusion, but believe that the reasons are more profound than those presented by Whittaker (1977) and MR. It can even be argued that diversity issues have become so confused in part because the criteria that have been used to distinguish between alpha and gamma diversity cannot be applied in a consistent manner. There is no objective and universally superior way of fixing the boundary between alpha and gamma diversity to a specific scale of observation and/or to a specific amount of heterogeneity.

Whittaker (1960) realised that the total species diversity observed in a landscape can be considered to consist of conceptually different components, and introduced the terms alpha, beta and gamma diversity to refer to these components. Gamma diversity is the simplest of the three: it is the total species diversity observed in the dataset (or landscape) of interest. Using the diversity concept reviewed above, gamma diversity can be notated with ${}^qD_\gamma$. To quantify gamma diversity, we need to make sampling decisions to specify two things:

1. The limits of the dataset (how individuals that belong to the dataset are chosen).
2. The limits among the species (how individuals are classified into species).

Both of these limits were specified in the preceding discussion on diversity, so we can write ${}^qD_\gamma = 1/{}^q\bar{p}_i$. Depending on the questions at hand, we might choose just a single kind of mean to calculate qD , or we might use several kinds of means in parallel (by choosing different

values for the parameter q) to explore how the effective number of species varies when differences in species abundances are emphasised in different ways.

If we are interested in its internal heterogeneity, the dataset needs to be divided into subunits. These subunits may or may not have been used as sampling units during the field inventory; this is irrelevant for the definition of the concepts. So, in addition to the limits of the dataset and the limits among the species, we also need to specify

3. The limits among the subunits (how individuals are classified into subunits).

Alpha and beta diversity are obtained when gamma diversity is partitioned into two components that take the subunit limits into account. Alpha diversity can be quantified directly from the species inventory data, following the same procedure of allocating individuals to pigeonholes as with gamma diversity. The difference is that now the proportional species abundances are calculated for each subunit separately; therefore, instead of placing all individuals of species i into the same pigeonhole, the individuals observed in different subunits are placed into different pigeonholes. 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

$${}^q\bar{p}_{ij} = \sqrt[q-1]{\sum_{j=1}^N w_j \sum_{i=1}^S p_{ij} p_{ij}^{q-1}} \tag{5}$$

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

$${}^qD_\alpha = 1/{}^q\bar{p}_{ij} \tag{6}$$

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

$${}^qD_\alpha = \sqrt[1-q]{\sum_{j=1}^N w_j ({}^qD_{\alpha j})^{1-q}} \tag{7}$$

This is the weighted generalised mean with exponent $1 - q$ of the single-subunit alpha diversities ${}^qD_{\alpha j}$ (i.e. the alpha diversity that would result if the dataset of interest consisted of subunit j only). Just like gamma diversity, alpha diversity can be visualised as the number of pigeonholes needed on a wall to accommodate all m individuals, such that each pigeonhole receives a proportion equal to ${}^q\bar{p}_{ij}$ of them.

Before going into details on what beta diversity is, let us consider how a total (gamma) can be partitioned into two

components (alpha and beta) using the volume analog. There are two conceptually different approaches: additive and multiplicative. If gamma (the total volume of an object) is partitioned additively, the object is simply cut into two parts, and the volume of the first part is called alpha and the volume of the second part is called beta. Each component then quantifies the same phenomenon as the total does, but for a different part of the original dataset. If gamma is partitioned multiplicatively, then each component quantifies a different phenomenon, but for the entire original dataset. For example, if the object of interest is a cube, alpha could represent the surface area of the top face of the cube, and beta could represent the height of the cube. For objects of other shapes, the volume can be calculated as the product of mean cross section area (alpha) and mean height (beta), provided that the dimensions in which alpha and beta are measured are orthogonal to each other (if they are not, a third term is needed in the equation to correct for this).

Whittaker’s original idea was that alpha and beta diversity represent different aspects of gamma diversity, which is achieved by the multiplicative partitioning. So we obtain:

$${}^qD_\gamma = {}^qD_\alpha {}^qD_\beta \tag{8}$$

Consequently, beta diversity is obtained as

$${}^qD_\beta = {}^qD_\gamma / {}^qD_\alpha \tag{9}$$

To visualise this, we need one more wall of pigeonholes. We move all ${}^qD_\gamma$ effective species from the pigeonholes on the gamma diversity wall to new pigeonholes on the beta diversity wall such that each new pigeonhole receives exactly as many effective species as there were pigeonholes on the alpha diversity wall. Each pigeonhole on the beta diversity wall now represents an effective compositional unit (CU_E), i.e., a virtual subunit that has the same number of effective species as the actual subunits have on average, but does not share effective species with any other effective compositional unit. If all species occur in all subunits, and the proportional abundance of any given species is constant across the subunits, then beta diversity will equal one effective compositional unit. The more the original subunits differ in species composition and abundance, the smaller the proportion of the total diversity that is contained in one subunit, and the more effective compositional units there will be. In some cases, it is even possible that the number of effective compositional units exceeds the number of actual subunits (Tuomisto 2010a).

True diversity is the effective number of types in a classification of interest, and this criterion is fulfilled equally well by alpha, beta and gamma diversity. Each component quantifies a different kind of diversity, however, because they correspond to pigeonholes on different

walls and hence to different measurement units: sp_E for gamma diversity, CU_E for beta diversity and sp_E/CU_E for alpha diversity. This conceptual difference is why alpha and gamma diversity need to be maintained separate—not because of differences in the scale of observation or the degree of internal heterogeneity (as proposed by Whittaker 1977 and MR).

Presence–absence data or abundance data?

MR criticised JRB for “mixing methods that use presence/absence data with methods that also use species abundance distribution”. MR argued that this “may mask the factors that drive species diversity” and “encourages an explosion of new measures and indices of diversity”. I agree with MR in that we need to make a distinction between measures based on presence–absence data and those based on abundance data. However, I believe that MR reached the right conclusion for the wrong reasons.

Firstly, the factors that drive species diversity must be considered separately from the definition of what diversity is. Diversity is simply a variable that can be measured for a dataset, and the factors that determine the amount of diversity need to be unravelled, just as in the case of any other variable, by observing which external factors can explain the variation in diversity among datasets. This is no different from how we treat volume: it is a variable that can be measured for an object, and the factors that determine the amount of volume need to be unravelled by observing which external factors can explain the variation in volume among objects. The factors that drive the volume of a flock of birds are likely to be different from the factors that drive the volume of a tree; therefore, to obtain meaningful results on volume, comparisons have to be limited to similar objects. For species diversity, the driving factors may be dependent on the organism group observed, the geographical region considered, the sampling scheme used and a number of other external factors. To obtain meaningful results on diversity, comparisons have to be limited to similar datasets.

Secondly, it is indeed possible to invent a larger number of different indices for abundance data than for presence–absence data. This is because quantitative variables are more flexible than binary variables, such that several different indices that use abundance data can be thought of as extensions of the same binary index (i.e., they give the same result if applied to presence–absence data). However, this is no reason to discard the abundance information altogether. Instead, we need to critically evaluate how each of the proposed indices uses the abundance information, so as to understand which phenomenon each index actually quantifies. Then we must be careful to only compare results obtained with comparable indices.

What, then, is the important difference between methods based on presence–absence versus abundance data? In the present context, the answer boils down to the difference between richness and diversity. It has been stated several times that diversity consists of two components, namely, richness and evenness (e.g., Magurran 2004). Logically, if richness is not the only component of diversity, it cannot be the same thing as diversity. Above, we visualised species richness as the number of pigeonholes needed to give each observed species a pigeonhole of its own. Species diversity, in contrast, was the number of new pigeonholes needed to accommodate all individuals such that each pigeonhole has the same proportion of the individuals as the original pigeonholes did on average. 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.

As we saw above, the total species diversity of a dataset (gamma diversity ${}^qD_\gamma$) can be partitioned into two independent components: the number of effective species per effective compositional unit (alpha diversity, ${}^qD_\alpha$) and the number of effective compositional units (beta diversity, ${}^qD_\beta$). I propose that the total species richness of a dataset be called gamma richness (S_γ) and that when it is partitioned into two independent components, these be called alpha richness (S_α , the mean number of actual species per compositional unit) and beta richness (S_β , the number of compositional units). A compositional unit (CU) is a virtual subunit that has the same number of actual species as the actual subunits do on average, but does not share species with any other compositional unit. Alpha richness is simply obtained as the unweighted arithmetic mean of the actual species numbers observed in the subunits:

$$S_\alpha = \frac{1}{N} \sum_{j=1}^N S_{xj} \quad (10)$$

Here, S_{xj} is the number of species per compositional unit that would be obtained if the dataset only consisted of subunit j . Beta richness is obtained as

$$S_\beta = S_\gamma / S_\alpha, \quad (11)$$

and it quantifies the number of compositionally different subunits needed to accommodate all actual species such that each compositional unit receives as many species as the actual subunits had on average.

The measurement units are: sp for gamma richness, CU for beta richness and sp/CU for alpha richness. The measurement units of the diversity components parallel these, but are based on sp_E and CU_E . Taking the abundance data into account causes the focus to shift from actual types to effective types, i.e., to a conceptually different phenomenon. It may be noted that in my recent review on beta diversity

(Tuomisto 2010a, b), I used the measurement unit CU for beta diversity. I now realise that the logical consistency of the scheme necessitates that CU be the unit of beta richness and that the unit of beta diversity be changed to CU_E .

As we saw above, choosing $q = 0$ causes species diversity to be calculated such that the number of effective species becomes equal to the number of actual species. Furthermore, at $q = 0$, alpha diversity becomes the arithmetic mean of the within-subunit values, just as in the case of alpha richness. Does this mean that richness and diversity always obtain the same value at $q = 0$? No, it does not. Alpha diversity is a weighted mean, whereby each subunit is weighted by the proportion of data it contributed to the dataset (Eq. 7). In contrast, alpha richness is an unweighted mean (Eq. 10). Therefore, if the subunits differ in the number of individuals, alpha diversity and alpha richness (and, consequently, beta diversity and beta richness) will obtain different numerical values. This is the case in many (if not most) real datasets because subunits are often defined on the basis of equal surface area or observation period, and then the observed abundance can vary among subunits. For this reason, confusing richness and diversity leads to inaccurate conclusions on the alpha and beta components.

It can now also be seen that the popular notation α , β and γ for alpha, beta and gamma diversity, respectively, is problematic for two reasons. Firstly, the same notation has been used to refer to both species diversity and species richness, which is confusing. Secondly, any notation used to refer to diversity should make the value of q explicit, because q specifies the kind of mean used when quantifying mean proportional species abundance and hence needs to be known to interpret the obtained value correctly. Therefore, it is preferable to use the explicit notation S_α , S_β and S_γ when referring to the richness components, and the notation ${}^qD_\alpha$, ${}^qD_\beta$ and ${}^qD_\gamma$ when referring to the diversity components. If the notation α , β and γ is maintained, its use should be limited to presence–absence data (richness). Tuomisto (2010a, b) introduced the notation ${}^q\beta$ to refer to beta diversity, and the corresponding ${}^q\alpha$ and ${}^q\gamma$ could be used for alpha and gamma diversity, respectively. Unfortunately, in those papers, I treated β and ${}^q\beta$ as synonyms, which created unnecessary ambiguity.

Beta diversity and species turnover

Although both JRB and MR attempted to clarify the concept of beta diversity, they still treated it very broadly. Whittaker (1960) himself had lumped several different phenomena (including turnover and differentiation) under this newly introduced term, and the majority of ecological literature continues the practice today. Such a broad usage of the term beta diversity should be avoided, however,

because it unnecessarily muddles an otherwise clearly definable concept.

Recently, Jost (2006, 2007) showed how some familiar similarity indices can be derived as transformations of true beta diversity, and I explored the practical meaning of these and many other indices that researchers have used when addressing questions related to beta diversity or turnover (Tuomisto 2010a, b). However, neither one of us made a conceptual distinction between indices based on presence–absence data and those based on abundance data, which I now believe should be done to clarify the concepts further. Because binary indices such as the Jaccard and Sørensen indices are based on presence–absence data, they are actually transformations of beta richness, not beta diversity. The Jaccard index equals

$$C_J = 2S_\alpha/S_\gamma - 1 = 2/S_\beta - 1, \quad (12)$$

and the Sørensen index equals

$$C_S = S_\gamma/S_\alpha - 1 = S_\beta - 1 \quad (13)$$

Their abundance-based versions (qC_J and qC_S) are obtained by replacing the richness components (S_α , S_γ and S_β) with the corresponding diversity components (${}^qD_\alpha$, ${}^qD_\gamma$ and ${}^qD_\beta$). The one-complements of both C_J and C_S have been popular measures of species turnover. The abundance-based indices qC_J and qC_S were introduced relatively recently (Jost 2006, 2007; Tuomisto 2010a), but it can be anticipated that they will be found useful as well. When these indices are used, it is worth remembering that $1 - C_J$ and $1 - C_S$ quantify actual species turnover (the turnover of actual species between compositional units), whereas $1 - {}^qC_J$ and $1 - {}^qC_S$ quantify effective species turnover (the turnover of effective species between effective compositional units). The diversity measures treat all species as if they were equally abundant when $q = 0$, but alpha diversity is still a weighted mean, whereas alpha richness is an unweighted mean (Eqs. 7 and 10). So if different subunits contain different numbers of individuals, the presence–absence indices give numerically different results than the corresponding abundance-based indices. It is only in the special case that all subunits contain the same number of individuals (or other units of abundance) that $q = 0$ leads to $C_J = {}^qC_J$ and $C_S = {}^qC_S$.

Neither turnover nor differentiation is unambiguous as such; both terms still need to be specified to indicate what kind of turnover or what kind of differentiation is meant. The Sørensen and Jaccard indices correspond to different kinds of turnover (Whittaker's turnover and proportional turnover, respectively; Tuomisto 2010a), and both can be adapted to quantify either actual or effective species turnover. Other indices quantify other things, and some of them have very little to do with each other either conceptually or numerically.

Making an informed choice of index for a particular study necessitates an understanding of which aspect of the data the indices quantify, and which of them corresponds to what is needed to answer the ecological questions of interest. The mathematical properties of the indices also determine what needs to be taken into account when comparing their values for different datasets and, indeed, if the index values can meaningfully be compared at all. How an index is calculated is not a mere technical detail, it essentially defines the variable of interest and hence the ecological meaning of the obtained results.

Acknowledgments I thank many colleagues for inspiring discussions on diversity, especially Kalle Ruokolainen, Lou Jost, Robert K. Colwell, Mirkka Jones, Jari Oksanen and Teemu Tahvanainen. Kalle Ruokolainen, Lou Jost, Mark V. Wilson and an anonymous reviewer gave useful comments on the manuscript. All thought experiments performed in this paper comply with the current laws of the countries visited while contemplating diversity.

References

- Beck J, Schwanghart W (2010) Comparing measures of species diversity from incomplete inventories: an update. *Methods Ecol Evol* 1:38–44
- Hill MO (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology* 54:427–432
- 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). *Ecol Econ* 68:925–992
- Jurasinski G, Retzer V, Beierkuhnlein C (2009) Inventory, differentiation, and proportional diversity: a consistent terminology for quantifying species diversity. *Oecologia* 159:15–26
- Lande R (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* 76:5–13
- MacArthur RH (1964) Environmental factors affecting bird species diversity. *Am Nat* 98:387–397
- MacArthur RH (1965) Patterns of species diversity. *Biol Rev* 40:510–533
- Magurran AE (2004) *Measuring biological diversity*. Blackwell, Oxford
- Moreno CE, Rodríguez P (2010) A consistent terminology for quantifying species diversity? *Oecologia* 163:279–282
- Routledge RD (1977) On Whittaker's components of diversity. *Ecology* 58:1120–1127
- Routledge RD (1979) Diversity indices: which ones are admissible? *J Theor Biol* 76:503–515
- Shmida A, Wilson MV (1985) Biological determinants of species diversity. *J Biogeogr* 12:1–20
- 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 RH (1960) Vegetation of the Siskiyou Mountains, Oregon and California. *Ecol Monogr* 30:279–338
- Whittaker RH (1977) Evolution of species diversity in land communities. *Evol Biol* 10:1–67