Tom Leinster and Christina A. Cobbold. 2012. Measuring diversity: the importance of species similarity. *Ecology* **93:477–489.**

Appendix A. Proofs of all the mathematical results.

Measuring diversity:
the importance of species similarity

Tom Leinster and Christina A. Cobbold

Appendix A: Mathematical proofs

In this appendix, we write

$$\mathbb{P}_S = \{(p_1, \ldots, p_S) \in \mathbb{R}^S \mid p_i \geq 0, \sum p_i = 1\}$$

for the set of relative abundance vectors for S species. As usual, S denotes the *theoretical* number of species in the population. It may be that $p_i = 0$ for some values of i; we write $s \leq S$ for the number of values of i such that $p_i > 0$. A **similarity matrix** is an $S \times S$ matrix $\mathbf{Z}$ such that $0 \leq Z_{ij} \leq 1$ for all i and j, and $Z_{ii} = 1$ for all i.

In the definition

$$^qD^{\mathbf{Z}}(\mathbf{p}) = \left(\sum p_i (\mathbf{Zp})_i^{q-1}\right)^{\frac{1}{1-q}} \quad (q \neq 1, \infty),$$

the sum is over all $i = 1, \ldots, S$ such that $p_i \neq 0$. We now explain why the eventuality that some species are absent ($p_i = 0$) must be handled in this way.

We would like our measures of diversity to be continuous in $\mathbf{p}$, as far as possible. A small change in the abundance of a species should cause only a small change in the measured diversity. An exception is species richness (naive diversity of order 0): if the relative abundance of a species increases from 0 to 0.001, the species richness increases by 1. However, we will show that for all q with $0 < q < \infty$, diversity of order q is indeed continuous in $\mathbf{p}$.

When all S species are present, the sum in the definition of $^qD^{\mathbf{Z}}(\mathbf{p})$ is over all i from 1 to S. It follows easily that $^qD^{\mathbf{Z}}$ is continuous on the set

$$\mathbb{P}_S^\circ = \{\mathbf{p} \in \mathbb{P}_S \mid p_i > 0 \text{ for all } i\}.$$

Given a continuous function on $\mathbb{P}_S^\circ$, there is at most one way of extending it to a continuous function on the whole of $\mathbb{P}_S$. We show that $^qD^{\mathbf{Z}}$, defined as above, is indeed continuous on the whole of $\mathbb{P}_S$ (for $0 < q < \infty$). This implies that, once the definition has been decided for relative abundance vectors in which no p_i is zero, our formula is the *only* way of handling the case where one or more p_i is zero.

Proposition A1 *Let $0 < q < \infty$. Then the function $^qD^{\mathbf{Z}}$ on $\mathbb{P}_S$ is continuous.*

The delicacy of the proof arises from the possibility that $p_i = 0$ for some i: for then it may be that $(\mathbf{Zp})_i = 0$, and in that case $(\mathbf{Zp})_i^{q-1}$ is undefined for $q < 1$.

Proof When $1 < q < \infty$, the sum in the definition of $^qD^{\mathbf{Z}}(\mathbf{p})$ can equivalently be taken over *all* values of i from 1 to S. Then $^qD^{\mathbf{Z}}(\mathbf{p})$ is clearly continuous in $\mathbf{p}$.

Let $0 < q < 1$. Define functions $f_1, \ldots, f_S$ on $\mathbb{P}_S$ by

$$f_i(\mathbf{p}) = \begin{cases} p_i(\mathbf{Zp})_i^{q-1} & \text{if } p_i > 0 \\ 0 & \text{if } p_i = 0. \end{cases}$$

1

Since $^qD^{\mathbf{Z}}(\mathbf{p}) = (f_1(\mathbf{p})+\cdots+f_S(\mathbf{p}))^{1/(1-q)}$, it suffices to prove that each f_i is continuous.

Certainly f_i is continuous on $\mathbb{P}_S^{(i)} = \{\mathbf{p} \in \mathbb{P}_S \mid p_i > 0\}$, so all we have to prove is that if $\mathbf{p} \in \mathbb{P}_S$ with $p_i = 0$, then $f_i(\mathbf{r}) \to 0$ as $\mathbf{r} \to \mathbf{p}$ in $\mathbb{P}_S$. Since $f_i(\mathbf{r}) = 0$ whenever $r_i = 0$, we might as well constrain $\mathbf{r}$ to lie in $\mathbb{P}_S^{(i)}$. We have $r_i \leq (\mathbf{Z}\mathbf{r})_i \leq 1$, so

$$r_i^q = r_i \cdot r_i^{q-1} \geq r_i \cdot (\mathbf{Z}\mathbf{r})_i^{q-1} \geq r_i \cdot 1 = r_i. \tag{A.1}$$

Now as $\mathbf{r} \to \mathbf{p}$ we have $r_i \to p_i = 0$, so $r_i^q \to 0$; hence $r_i(\mathbf{Z}\mathbf{r})_i^{q-1} \to 0$ by (A.1), as required.

Finally, consider $q = 1$. It is enough to prove that $(\mathbf{Z}\mathbf{p})_i^{p_i}$ is continuous in $\mathbf{p}$, for each i. (One evaluates 0^0 as 1.) Certainly it is continuous on $\mathbb{P}_S^{(i)}$, so all we have to prove is that if $\mathbf{p} \in \mathbb{P}_S$ with $p_i = 0$, then $(\mathbf{Z}\mathbf{r})_i^{r_i} \to (\mathbf{Z}\mathbf{p})_i^{p_i} = 1$ as $\mathbf{r} \to \mathbf{p}$ in $\mathbb{P}_S$. We have $r_i \leq (\mathbf{Z}\mathbf{r})_i \leq 1$, so

$$r_i^{r_i} \leq (\mathbf{Z}\mathbf{r})_i^{r_i} \leq 1^{r_i} = 1. \tag{A.2}$$

Observe also that $\lim_{x\to 0+} x^x = 1$. Now as $\mathbf{r} \to \mathbf{p}$ we have $r_i \to p_i = 0$, so we also have $r_i^{r_i} \to 1$; hence $(\mathbf{Z}\mathbf{r})_i^{r_i} \to 1$ by (A.2), as required. $\qquad\square$

Diversity of order 0 need not be continuous in $\mathbf{p}$ (depending on the similarity matrix). The same goes for diversity of order ∞. However, $^0D^{\mathbf{Z}}$ and $^\infty D^{\mathbf{Z}}$ do fit naturally into the family $(^qD^{\mathbf{Z}})$ of diversity measures, in the sense made precise by the following proposition.

Proposition A2 *Let $\mathbf{p} \in \mathbb{P}_S$ and let $\mathbf{Z}$ be an $S \times S$ similarity matrix. Then:*

i. $^qD^{\mathbf{Z}}(\mathbf{p})$ is continuous in q, for $0 < q < \infty$

ii. $\lim_{q\to 0} {}^qD^{\mathbf{Z}}(\mathbf{p}) = {}^0D^{\mathbf{Z}}(\mathbf{p})$

iii. $\lim_{q\to\infty} {}^qD^{\mathbf{Z}}(\mathbf{p}) = {}^\infty D^{\mathbf{Z}}(\mathbf{p})$.

Proof All of this follows from standard results on generalized means (also called power means), which can be found in Hardy et al. (1952). Writing $x_i = (\mathbf{Z}\mathbf{p})_i$, we have

$$1/{}^qD^{\mathbf{Z}}(\mathbf{p}) = \left(\sum_{i:\, p_i > 0} p_i x_i^{q-1} \right)^{1/(q-1)}$$

$(q \neq 1)$, which is the mean of order $q - 1$ of the family $(x_i)_{i:\, p_i > 0}$, weighted by the p_is. Similarly, $1/{}^1D^{\mathbf{Z}}(\mathbf{p})$ is the mean of order 0.

Parts (i) and (ii) are immediate except for continuity at $q = 1$, which follows from Theorem 3 of Hardy et al. Part (iii) follows from Theorem 4 of Hardy et al. $\qquad\square$

Now consider the diversity $^qD^{\mathbf{Z}}(\mathbf{p})$ when q is an integer greater than 1. Let μ_q be the expected value of

$$Z_{i_1,i_2}Z_{i_1,i_3}\cdots Z_{i_1,i_q}$$

over all samples with replacement of q individuals from the community, whose respective species have been written as $i_1, i_2, \ldots, i_q$. Thus,

$$\mu_q = \sum_{i_1,i_2,\ldots,i_q} p_{i_1} p_{i_2} p_{i_3} \cdots p_{i_q} Z_{i_1,i_2} Z_{i_1,i_3} \cdots Z_{i_1,i_q}.$$

Proposition A3 *Let $q \geq 2$ be an integer. Then $^qD^{\mathbf{Z}}(\mathbf{p}) = \mu_q^{1/(1-q)}$.*

2

Proof Since $q > 1$, the sum in the definition of ${}^qD^{\mathbf{Z}}(\mathbf{p})$ might as well be over all $i = 1, \ldots, S$ (including those for which $p_i = 0$). We have

$$
{}^qD^{\mathbf{Z}}(\mathbf{p})^{1-q} = \sum_{i=1}^{S} p_i(\mathbf{Zp})_i^{q-1} = \sum_{i=1}^{S} p_i\left(\sum_{j=1}^{S} Z_{ij}p_j\right)^{q-1}
$$

$$
= \sum_{i,j_1,\ldots,j_{q-1}} p_i Z_{i,j_1} p_{j_1} Z_{i,j_2} p_{j_2} \cdots Z_{i,j_{q-1}} p_{j_{q-1}} = \mu_q,
$$

as required. $\qquad\square$

The next result states that our formula for ${}^qH^{\mathbf{Z}}(\mathbf{p})$ agrees with that of Ricotta and Szeidl (2006), except that they did not specify how their formula was to be interpreted in the case that some of the relative abundances p_i are 0. Recall from Section 3 of the main text that although Ricotta and Szeidl referred to their inter-species differences as 'distances' and denoted them by d_{ij}, we call them 'dissimilarities' and denote them by Δ_{ij}, since they are measured on a scale of 0 to 1.

Let $\mathbf{Z}$ be an $S \times S$ similarity matrix. Let $\mathbf{\Delta}$ be the corresponding dissimilarity matrix, defined by $\Delta_{ij} = 1 - Z_{ij}$.

Proposition A4 *For $0 \le q < \infty$ and $\mathbf{p} \in \mathbb{P}_S$,*

$$
{}^qH^{\mathbf{Z}}(\mathbf{p}) = \begin{cases} \frac{1}{q-1}\left(1 - \sum_{i:\ p_i>0} p_i(1 - \sum_{j\neq i} \Delta_{ij}p_j)^{q-1}\right) & \text{if } q \neq 1 \\ -\sum_{i:\ p_i>0} p_i \ln(1 - \sum_{j\neq i} \Delta_{ij}p_j) & \text{if } q = 1. \end{cases}
$$

Proof For $i = 1, \ldots, S$ we have

$$
(\mathbf{Zp})_i = \sum_{j=1}^{S} Z_{ij}p_j = \sum_{j=1}^{S}(1 - \Delta_{ij})p_j = 1 - \sum_{j=1}^{S}\Delta_{ij}p_j = 1 - \sum_{j\neq i}\Delta_{ij}p_j
$$

since $\sum_{j=1}^{S} p_j = 1$ and $\Delta_{ii} = 1 - Z_{ii} = 0$. The result follows. $\qquad\square$

Many authors have found it convenient to assume that the measure of dissimilarity or distance between species is a metric in the mathematical sense. For distances (d_{ij}) to define a **metric** means that (i) $d_{ij} = 0$ if and only if $i = j$; (ii) $d_{ij} = d_{ji}$; and (iii) the triangle inequality holds: $d_{ij} + d_{jk} \ge d_{ik}$.

Now, it may be that we started with inter-species distances d_{ij} measured on a scale of 0 to ∞ and converted them to similarities Z_{ij} by the formula $Z_{ij} = e^{-ud_{ij}}$, where u is a positive constant. These in turn correspond to dissimilarities $\Delta_{ij} = 1 - Z_{ij}$, measured on a scale of 0 to 1. Whether we are using (d_{ij}) or $\mathbf{\Delta} = (\Delta_{ij})$ matters: asking that (d_{ij}) defines a metric is not the same as asking that (Δ_{ij}) defines a metric, as the following proposition shows.

Proposition A5 *If (d_{ij}) is a metric then (Δ_{ij}) is a metric, but the converse implication fails.*

Proof The quantities d_{ij} and Δ_{ij} are related by $\Delta_{ij} = 1 - e^{-ud_{ij}}$, or equivalently, $d_{ij} = -(1/u)\ln(1 - \Delta_{ij})$. It is easy to see that axioms (i) and (ii) hold for d if and only if they hold for $\mathbf{\Delta}$. For the triangle inequality (iii),

$$
\begin{aligned}
d_{ij} + d_{jk} \ge d_{ik} &\iff -\ln(1 - \Delta_{ij}) - \ln(1 - \Delta_{jk}) \ge -\ln(1 - \Delta_{ik}) \\
&\iff (1 - \Delta_{ij})(1 - \Delta_{jk}) \le 1 - \Delta_{ik} \\
&\iff \Delta_{ij} + \Delta_{jk} - \Delta_{ij}\Delta_{jk} \ge \Delta_{ik}.
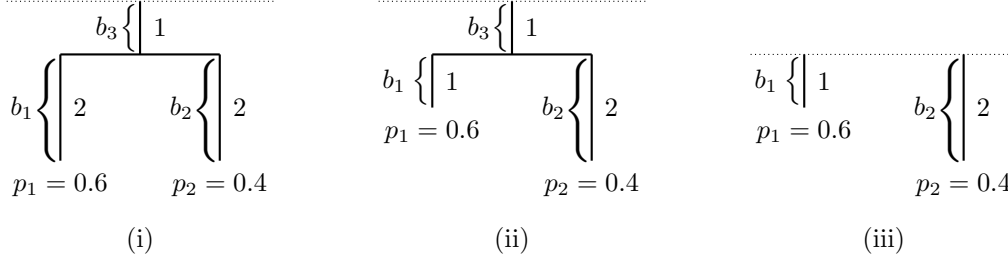\end{aligned}
$$

Figure A1: Three simple phylogenetic trees, each with two present-day species. The dotted horizontal lines show the beginning of the time period under consideration. In (i), the tree has three branches, b_1, b_2, b_3, with $L(b_1) = L(b_2) = 2$ and $L(b_3) = 1$. Tree (i) is ultrametric; trees (ii) and (iii) are not. In trees (i) and (ii), the present-day species have a common ancestor in the time period considered; in tree (iii), they do not.

This implies the triangle inequality for $\boldsymbol{\Delta}$, but not conversely: e.g. $S = 3$, $\Delta_{12} = \Delta_{23} = 1/2$, $\Delta_{13} = 4/5$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

The next few results concern the relationships between our measures $^q D^{\mathbf{Z}}$ and other indices of diversity.

First we compare our measures with some measures of phylogenetic diversity. All of the latter are based on phylogenetic trees, some very simple examples of which are shown in Fig. A1. From a phylogenetic tree we extract the following information:

- the number S of present-day species (which we put in order: $1, 2, \ldots, S$)

- the set of branches

- for each branch b, its length $L(b)$

- for each branch b, the set $I_b \subseteq \{1, 2, \ldots, S\}$ of present-day species descended from b.

We make the convention that the variable i always ranges over the present-day species $1, 2, \ldots, S$, and the variable b always ranges over the set of branches.

Before making the comparison, we review the phylogenetic measures concerned. The simplest is that of Faith (1992), which is just the total branch length,

$$\sum_b L(b).$$

Suppose now that we have a relative abundance vector $\mathbf{p} = (p_1, \ldots, p_S)$ for the present-day species. For each branch b, write

$$p(b) = \sum_{i:\ i \in I_b} p_i,$$

which is the total relative abundance of present-day species descended from b. For each present-day species i, write

$$L_i = \sum_{b:\ i \in I_b} L(b),$$

which is the total evolutionary change undergone by the ith species over the time-span considered. For the tree to be **ultrametric** means that $L_1 = L_2 = \cdots = L_S$.

4

Chao et al. (2010) write

$$\overline{T} = \sum_i p_i L_i = \sum_{i,b:\ i \in I_b} p_i L(b) = \sum_b p(b) L(b) \qquad (A.3)$$

for the mean evolutionary change per present-day species. This reflects the time-span under consideration. When the tree is ultrametric, $\overline{T} = L_1 = \cdots = L_S$. For $0 \leq q < \infty$, they define the **mean phylogenetic diversity** of order q as

$$
{}^q\overline{D}(\overline{T}) =
\begin{cases}
\left(\sum_b \dfrac{L(b)}{\overline{T}} p(b)^q \right)^{1/(1-q)} & \text{if } q \neq 1 \\[2ex]
\prod_b p(b)^{-(L(b)/\overline{T})p(b)} & \text{if } q = 1.
\end{cases}
$$

The expression at $q = 1$ is the limit as $q \to 1$ of ${}^q\overline{D}(\overline{T})$. Although they do not mention it, there is also a limit as $q \to \infty$, namely

$$ {}^\infty\overline{D}(\overline{T}) = 1/\max_b p(b). $$

When the present-day species have a common ancestor in the time-span considered, the tree has a root b; then $p(b) = 1$ and so ${}^\infty\overline{D}(\overline{T}) = 1$.

The phylogenetic entropy of Allen, Kon, and Bar-Yam (2009) is

$$ H^{\text{AKB}}(\mathbf{p}) = -\sum_b L(b)p(b)\ln p(b). $$

They also implicitly propose a phylogenetic entropy of each order $q \geq 0$,

$$ {}^q H^{\text{AKB}}(\mathbf{p}) = \sum_b L(b)p(b)\sigma_q(p(b)) $$

where σ_q is the surprise function of Section 3. For example, ${}^1 H^{\text{AKB}} = H^{\text{AKB}}$.

Chao et al. (2010) showed that Faith's measure and H^{AKB} can be recovered from their measures. We now connect our measures to Chao et al.'s, Faith's, and ${}^q H^{\text{AKB}}$ (for any q). To do this, we begin by showing how, from the data given (a phylogenetic tree and relative abundances for the present-day species), we can extract a matrix $\mathbf{Z}$ and a relative abundance vector $\boldsymbol{\pi}$.

Since we are considering the evolution of species through time, our basic biological units (which would usually be called 'species') are not present-day species, but species *in a particular period of their evolutionary history*. That is, a unit is a pair (i, b) where b is a branch and $i \in I_b$. We call such a pair a **historical species**. Its relative abundance $\pi_{(i,b)}$ is weighted according to how great a portion of evolutionary time it occupies:

$$ \pi_{(i,b)} = \frac{L(b)}{\overline{T}} p_i. \qquad (A.4) $$

(Equation (A.3) implies that $\sum_{i,b:\ i \in I_b} \pi_{(i,b)} = 1$.) The matrix $\mathbf{Z}$ is defined by

$$
Z_{(i,b),\,(j,c)} =
\begin{cases}
\overline{T}/L_j & \text{if } j \in I_b \\
0 & \text{otherwise.}
\end{cases}
\qquad (A.5)
$$

We offer no intuitive interpretation of this last formula. But intuitive or not, this formula provides a strong logical connection (Proposition A7) between our diversity measures and the measures of phylogenetic diversity mentioned above.

If the tree is ultrametric then $\overline{T}/L_j$ is always 1, so the matrix $\mathbf{Z}$ consists entirely of 0s and 1s. It is not symmetric, but *is* a similarity matrix in the precise sense defined in the main text and at the beginning of this appendix. This demonstrates that non-symmetric similarity matrices can serve a useful purpose.

If the tree is not ultrametric then matters are more delicate. First, the matrix $\mathbf{Z}$ depends on $\overline{T} = \sum_i p_i L_i$, which in turn depends on the relative abundances $\mathbf{p}$ of the present-day species. (There is no such dependence when the tree is ultrametric.) When the entries of $\mathbf{Z}$ depend on species abundances, it can no longer be thought of as a 'similarity' matrix in the same way. Second, $\mathbf{Z}$ is *not* a similarity matrix in the precise sense, since some of its entries are strictly greater than 1. But this turns out to cause no mathematical difficulty. The situation is clarified by introducing a new piece of terminology, as follows.

Let us say that a **relatedness matrix** is a real square matrix $\mathbf{Y}$ such that $Y_{ij} \geq 0$ for all i, j and $Y_{ii} > 0$ for all i. Certainly every similarity matrix is a relatedness matrix, but not vice versa. Observe that for any relatedness matrix $\mathbf{Y}$ and relative abundance vector $\mathbf{r}$, if $r_1 > 0$ then $(\mathbf{Yr})_1 > 0$: for

$$(\mathbf{Yr})_1 = Y_{11}r_1 + \sum_{j=2}^{S} Y_{1j}r_j \geq Y_{11}r_1 > 0.$$

Similarly, for any $i \in \{1, \ldots, n\}$, if $r_i > 0$ then $(\mathbf{Yr})_i > 0$; so $(\mathbf{Yr})_i^{q-1}$ is a well-defined real number (even if $q < 1$). The definitions of the diversities ${}^qD^{\mathbf{Y}}(\mathbf{r})$ and entropies ${}^qH^{\mathbf{Y}}(\mathbf{r})$ therefore make mathematical sense for an arbitrary relatedness matrix $\mathbf{Y}$. The phylogenetic matrix $\mathbf{Z}$ in equation (A.5) is always a relatedness matrix. It only satisfies the stronger condition of being a similarity matrix if the tree is ultrametric. But it is a relatedness matrix in any case, so ${}^qD^{\mathbf{Z}}(\boldsymbol{\pi})$ and ${}^qH^{\mathbf{Z}}(\boldsymbol{\pi})$ are always mathematically well-defined quantities.

Section 4 of the main text contains a list of important properties satisfied by the diversity measures ${}^qD^{\mathbf{Z}}$. Most of them hold for an arbitrary relatedness matrix $\mathbf{Z}$. The stronger assumption that $\mathbf{Z}$ is a similarity matrix is only needed for the naive model and range properties, as shown in Propositions A9–A17. So the notion of relatedness matrix widens the scope of our results. It is also useful because it allows us to prove new results about the phylogenetic measures of Chao et al. (2010) (Corollary A12).

Example A6 In the ultrametric tree of Fig. A1(i), there are four historical species: $(1, b_1)$, $(1, b_3)$, $(2, b_2)$ and $(2, b_3)$. We have $\overline{T} = 2 + 1 = 3$, and

$$\mathbf{Z} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \qquad \boldsymbol{\pi} = \begin{pmatrix} \frac{2}{3} \times 0.6 \\ \frac{1}{3} \times 0.6 \\ \frac{2}{3} \times 0.4 \\ \frac{1}{3} \times 0.4 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.2 \\ 0.267 \\ 0.133 \end{pmatrix}.$$

This $\mathbf{Z}$ is a similarity matrix, depending only on the structure of the phylogenetic tree (and independent of the species abundances). As observed above, these properties of $\mathbf{Z}$ are guaranteed by the fact that the tree is ultrametric.

In the non-ultrametric tree of Fig. A1(ii), there are the same four historical species (i, b). We have $L_1 = 1 + 1 = 2$, $L_2 = 2 + 1 = 3$, and $\overline{T} = 0.6 \times L_1 + 0.4 \times L_2 = 2.4$, giving

$$\mathbf{Z} = \begin{pmatrix} 1.2 & 1.2 & 0 & 0 \\ 1.2 & 1.2 & 0.8 & 0.8 \\ 0 & 0 & 0.8 & 0.8 \\ 1.2 & 1.2 & 0.8 & 0.8 \end{pmatrix}, \qquad \boldsymbol{\pi} = \begin{pmatrix} \frac{1}{2.4} \times 0.6 \\ \frac{1}{2.4} \times 0.6 \\ \frac{2}{2.4} \times 0.4 \\ \frac{1}{2.4} \times 0.4 \end{pmatrix} = \begin{pmatrix} 0.25 \\ 0.25 \\ 0.333 \\ 0.167 \end{pmatrix}.$$

This $\mathbf{Z}$ is a relatedness matrix but not a similarity matrix, and it does depend on the abundances of the present-day species.

The following result makes the connection between our measures and the phylogenetic measures of Chao et al. (2010), of Faith (1992), and of Allen et al. (2009).

Proposition A7 *Take a phylogenetic tree and a relative abundance vector for the present-day species. Then, defining $\mathbf{Z}$ and $\boldsymbol{\pi}$ as in equations (A.4) and (A.5) above,*

i. ${}^q D^{\mathbf{Z}}(\boldsymbol{\pi}) = {}^q \overline{D}(\overline{T})$, the mean phylogenetic diversity of Chao et al. (2010), for all $0 \le q \le \infty$. In particular, ${}^0 D^{\mathbf{Z}}(\boldsymbol{\pi})$ is $\frac{1}{\overline{T}}$ times Faith's phylogenetic diversity.

ii. ${}^q H^{\mathbf{Z}}(\boldsymbol{\pi}) = \frac{1}{\overline{T}} \times {}^q H^{AKB}(\mathbf{p})$, the phylogenetic entropy of Allen et al. (2009), for all $0 \le q < \infty$. In particular, ${}^1 H^{\mathbf{Z}}(\boldsymbol{\pi}) = \frac{1}{\overline{T}} \times H^{AKB}(\mathbf{p})$.

So where Chao et al.'s measures are diversities, Allen et al.'s are (up to a multiplicative factor) the accompanying entropies.

Proof First we compute $\mathbf{Z}\boldsymbol{\pi}$. For each historical species (i, b),

$$(\mathbf{Z}\boldsymbol{\pi})_{(i,b)} = \sum_{j,c: \ j \in I_c} Z_{(i,b),(j,c)} \pi_{(j,c)} = \sum_{j,c: \ j \in I_b \cap I_c} \frac{\overline{T}}{L_j} \frac{L(c)}{\overline{T}} p_j$$

$$= \sum_{j: \ j \in I_b} \frac{p_j}{L_j} \sum_{c: \ j \in I_c} L(c) = \sum_{j: \ j \in I_b} p_j = p(b).$$

It suffices to prove (i) and (ii) when $q \ne 1, \infty$, by Proposition A2 (which holds for arbitrary relatedness matrices, by exactly the same proof). For (i),

$$
{}^q D^{\mathbf{Z}}(\boldsymbol{\pi}) = \left(\sum_{\substack{i,b: \ i \in I_b, \\ \pi_{(i,b)} > 0}} \pi_{(i,b)} (\mathbf{Z}\boldsymbol{\pi})_{(i,b)}^{q-1} \right)^{1/(1-q)} = \left(\sum_{\substack{i,b: \ i \in I_b, \\ p_i > 0}} \frac{L(b)}{\overline{T}} p_i \cdot p(b)^{q-1} \right)^{1/(1-q)}
$$

$$
= \left(\sum_b \frac{L(b)}{\overline{T}} p(b)^q \right)^{1/(1-q)} = {}^q \overline{D}(\overline{T}).
$$

In particular, ${}^0 D^{\mathbf{Z}}(\boldsymbol{\pi}) = (1/\overline{T}) \sum_b L(b)$.

For (ii),

$$
{}^q H^{\mathbf{Z}}(\boldsymbol{\pi}) = \sum_{\substack{i,b: \ i \in I_b, \\ \pi_{(i,b)} > 0}} \pi_{(i,b)} \sigma_q\big((\mathbf{Z}\boldsymbol{\pi})_{(i,b)}\big) = \sum_{\substack{i,b: \ i \in I_b, \\ p_i > 0}} \frac{L(b)}{\overline{T}} p_i \cdot \sigma_q(p(b))
$$

$$
= \sum_b \frac{L(b)}{\overline{T}} p(b) \cdot \sigma_q(p(b)) = \frac{1}{\overline{T}} \times {}^q H^{AKB}(\mathbf{p}),
$$

as required. $\qquad \square$

Next we turn to the indices studied by Hurlbert (1971) and Smith and Grassle (1977), showing that they can be derived from the naive diversities (Hill numbers) ^qD. For each $m \geq 2$, let $H_m^{\mathrm{HSG}}(\mathbf{p})$ be the expected number of species represented in a random sample of m individuals. As Hurlbert observed,

$$H_m^{\mathrm{HSG}}(\mathbf{p}) = \sum_{i=1}^{S}(1 - (1 - p_i)^m).$$

Proposition A8 *For each $m \geq 2$, the Hurlbert–Smith–Grassle index is given by*

$$H_m^{HSG}(\mathbf{p}) = m - \sum_{q=2}^{m}(-1)^q\binom{m}{q}{}^qD(\mathbf{p})^{1-q},$$

where $\binom{m}{q}$ is the binomial coefficient $m!/q!(m-q)!$.

Proof We have

$$\sum_{i=1}^{S}(1 - (1 - p_i)^m) = S - \sum_{i=1}^{S}\sum_{q=0}^{m}\binom{m}{q}(-p_i)^q = S - \sum_{q=0}^{m}(-1)^q\binom{m}{q}\sum_{i=1}^{S}p_i^q$$

$$= S - \left\{\binom{m}{0}S - \binom{m}{1}1 + \sum_{q=2}^{m}(-1)^q\binom{m}{q}{}^qD(\mathbf{p})^{1-q}\right\}$$

$$= m - \sum_{q=2}^{m}(-1)^q\binom{m}{q}{}^qD(\mathbf{p})^{1-q},$$

as required. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

We now prove the properties of our diversity measures stated in Section 4.

> *Our standing assumption for the rest of this appendix is that matrices called $\mathbf{Z}$, $\mathbf{Z}(i)$, etc. are relatedness matrices. They are not required to be similarity matrices except where this is stated explicitly.*

We observe that although Proposition A2 is stated for similarity matrices, it holds for arbitrary relatedness matrices, by exactly the same proof.

Proposition A9 (Effective number) *Let $0 \leq q \leq \infty$. Then diversity of order q is an effective number; that is, if $p_1 = \cdots = p_S = 1/S$ then $^qD^{\mathbf{I}}(\mathbf{p}) = S$.*

Proof This follows immediately from the definition of $^qD^{\mathbf{Z}}(\mathbf{p})$, substituting $\mathbf{Z} = \mathbf{I}$ and $\mathbf{p} = (1/S, \ldots, 1/S)$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Suppose now that the community is divided into m subcommunities. No species appears in more than one subcommunity, and species in different subcommunities are totally dissimilar.

Write $w_1, \ldots, w_m$ for the relative sizes of the subcommunities (so that $\sum w_i = 1$). Within the ith subcommunity, write S_i for the number of species, $\mathbf{r}_i = (r_{i1}, \ldots, r_{iS_i})$ for its relative abundance vector (so that $\sum_{j=1}^{S_i} r_{ij} = 1$), and $\mathbf{Z}(i)$ for its matrix (an $S_i \times S_i$ matrix).

Write S for the number of species in the whole community, $\mathbf{p}$ for the overall relative abundance vector, and $\mathbf{Z}$ for the overall, $S \times S$, matrix.

Proposition A10 (Modularity) *For $0 \leq q \leq \infty$,*

$$
{}^{q}D^{\mathbf{Z}}(\mathbf{p}) = \begin{cases} \left(\sum_{i:\, w_i>0} w_i^q d_i^{1-q}\right)^{\frac{1}{1-q}} & \text{if } q \neq 1, \infty \\ {}^{1}D(\mathbf{w}) d_1^{w_1} d_2^{w_2} \cdots d_m^{w_m} & \text{if } q = 1 \\ \min_{i:\, w_i>0}(d_i/w_i) & \text{if } q = \infty \end{cases}
$$

where $d_i = {}^{q}D^{\mathbf{Z}(i)}(\mathbf{r}_i)$.

Proof We have

$$
S = S_1 + \cdots + S_m,
$$
$$
\mathbf{p} = (w_1 r_{11}, \ldots, w_1 r_{1S_1}, \ \ldots \ , w_m r_{m1}, \ldots, w_m r_{mS_m}),
$$
$$
\mathbf{Z} = \begin{pmatrix} \mathbf{Z}(1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}(2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{Z}(m) \end{pmatrix},
$$

the last expression being a block sum of matrices. So for $1 \leq i \leq m$ and $1 \leq j \leq S_i$,

$$
(\mathbf{Z}\mathbf{p})_{S_1 + \cdots + S_{i-1} + j} = w_i(\mathbf{Z}(i)\mathbf{r}_i)_j.
$$

Hence for $q \neq 1, \infty$, letting i range over $1, \ldots, m$ and j range over $1, \ldots, S_i$,

$$
{}^{q}D^{\mathbf{Z}}(\mathbf{p})^{1-q} = \sum_{i,j:\, w_i r_{ij}>0} w_i r_{ij} \big(w_i(\mathbf{Z}(i)\mathbf{r}_i)_j\big)^{q-1}
$$
$$
= \sum_{i:\, w_i>0} w_i^q \sum_{j:\, r_{ij}>0} r_{ij}(\mathbf{Z}(i)\mathbf{r}_i)_j^{q-1} = \sum_{i:\, w_i>0} w_i^q d_i^{1-q},
$$

as required. For $q = 1$,

$$
{}^{1}D^{\mathbf{Z}}(\mathbf{p}) = \prod_{i,j} \big(w_i(\mathbf{Z}(i)\mathbf{r}_i)_j\big)^{-w_i r_{ij}}
$$
$$
= \left(\prod_{i,j} w_i^{-w_i r_{ij}}\right)\left(\prod_{i,j}(\mathbf{Z}(i)\mathbf{r}_i)_j^{-w_i r_{ij}}\right) = {}^{1}D(\mathbf{w}) \prod_i d_i^{w_i}
$$

since $\sum_j r_{ij} = 1$ for each i. Finally, for $q = \infty$,

$$
{}^{\infty}D^{\mathbf{Z}}(\mathbf{p}) = 1 \Big/ \max_{i,j:\, w_i r_{ij}>0} w_i(\mathbf{Z}(i)\mathbf{r}_i)_j = 1 \Big/ \max_{i:\, w_i>0}\left(w_i \max_{j:\, r_{ij}>0}(\mathbf{Z}(i)\mathbf{r}_i)_j\right)
$$
$$
= 1 \Big/ \max_{i:\, w_i>0}(w_i/d_i) = \min_{i:\, w_i>0}(d_i/w_i),
$$

as required. $\qquad\square$

Proposition A11 (Replication) *If $w_1 = \cdots = w_m$ and $d_1 = \cdots = d_m = d$ then ${}^{q}D^{\mathbf{Z}}(\mathbf{p}) = md$ for all $0 \leq q \leq \infty$.*

Proof Substitute $w_i = 1/m$ and $d_i = d$ into Proposition A10. $\qquad\square$

We can deduce some facts about the mean phylogenetic diversity of Chao et al. (2010), extending the replication principle that they proved.

Corollary A12 *Let $0 \leq q \leq \infty$. Take m completely distinct phylogenetic assemblages, with relative sizes $w_1, \ldots, w_m$. Write $\overline{T}_i$ for the mean evolutionary change per species in the ith assemblage, $d_i = {}^q\overline{D}(\overline{T}_i)$ for its mean phylogenetic diversity of order q, and $\overline{T}$ for the mean evolutionary change per species in the pooled assemblage.*

    i. If $\overline{T}_1 = \cdots = \overline{T}_m$ then the mean phylogenetic diversity ${}^q\overline{D}(\overline{T})$ of the pooled assemblage is determined by $w_1, \ldots, w_m$ and $d_1, \ldots, d_m$ via the formula in Proposition A10.

    ii. If $w_1 = \cdots = w_m = 1/m$ and $d_1 = \cdots = d_m = d$ then the mean phylogenetic diversity ${}^q\overline{D}(\overline{T})$ of the pooled assemblage is md.

Chao et al. (2010) proved (ii) under the further assumption that the m assemblages all have the same mean evolutionary change per species; but this assumption is unnecessary.

Proof We begin the proof without assuming either the hypothesis in (i) or that in (ii). We use the notation set up after Proposition A5.

Write $\mathbf{Z}(i)$ for the matrix (A.5) of the ith assemblage, and $\mathbf{Z}$ for the matrix (A.5) of the pooled assemblage. Then $\mathbf{Z}$ is the block sum

$$
\mathbf{Z} = \begin{pmatrix} (\overline{T}/\overline{T}_1)\mathbf{Z}(1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\overline{T}/\overline{T}_2)\mathbf{Z}(2) & & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & (\overline{T}/\overline{T}_m)\mathbf{Z}(m) \end{pmatrix}.
$$

Write $\boldsymbol{\sigma}(i)$ for the relative abundance vector (A.4) of the ith assemblage, and $\boldsymbol{\pi}$ for the relative abundance vector (A.4) of the pooled assemblage. Then whenever a present-day species j belongs to the ith assemblage and is descended from a branch b,

$$
\pi_{(j,b)} = \frac{\overline{T}_i}{\overline{T}} w_i \cdot \sigma(i)_{(j,b)}.
$$

Note also that $\overline{T} = \sum_{i=1}^{m} w_i \overline{T}_i$.

We now apply Proposition A10. For $q \neq 1, \infty$, it gives

$$
\begin{aligned}
{}^q D^{\mathbf{Z}}(\boldsymbol{\pi}) &= \left(\sum_{i:\, w_i > 0} \left(\frac{\overline{T}_i}{\overline{T}} w_i \right)^q \left({}^q D^{(\overline{T}/\overline{T}_i)\mathbf{Z}(i)}(\boldsymbol{\sigma}(i)) \right)^{1-q} \right)^{1/(1-q)} \\
&= \left(\sum_{i:\, w_i > 0} \left(\frac{\overline{T}_i}{\overline{T}} w_i \right)^q \left(\frac{\overline{T}_i}{\overline{T}} {}^q D^{\mathbf{Z}(i)}(\boldsymbol{\sigma}(i)) \right)^{1-q} \right)^{1/(1-q)} \\
&= \left(\sum_{i:\, w_i > 0} \frac{\overline{T}_i}{\overline{T}} w_i^q \left({}^q D^{\mathbf{Z}(i)}(\boldsymbol{\sigma}(i)) \right)^{1-q} \right)^{1/(1-q)}
\end{aligned}
$$

where in the second step, we have used the easily verified fact that when a matrix $\mathbf{Z}$ is multiplied by a constant $\lambda > 0$, the resulting diversity (of any order) is divided by λ. So by Proposition A7,

$$
{}^q\overline{D}(\overline{T}) = \left(\sum_{i:\, w_i > 0} \frac{\overline{T}_i}{\overline{T}} w_i^q d_i^{1-q} \right)^{1/(1-q)}. \tag{A.6}
$$

To prove (i), suppose that $\overline{T}_1 = \cdots = \overline{T}_m$: then $\overline{T} = \overline{T}_i$ for all i, so (A.6) gives

$$^q\overline{D}(\overline{T}) = \left(\sum_{i:\, w_i > 0} w_i^q d_i^{1-q} \right)^{1/(1-q)},$$

and similarly for $q = 1, \infty$. To prove (ii), suppose instead that $w_1 = \cdots = w_m = 1/m$ and $d_1 = \cdots = d_m = d$: then $\overline{T} = (1/m)\sum \overline{T}_i$, so (A.6) gives

$$^q\overline{D}(\overline{T}) = \left(\sum_{i=1}^m \frac{\overline{T}_i}{\overline{\overline{T}}} \left(\frac{1}{m} \right)^q d^{1-q} \right)^{1/(1-q)} = md,$$

and similarly for $q = 1, \infty$, as required. $\qquad\square$

The 'elementary properties' all follow from a single general lemma. For this we need some notation. Take $S, T \geq 1$ and a function $\theta \colon \{1, \ldots, S\} \to \{1, \ldots, T\}$. For each $\mathbf{r} \in \mathbb{P}_S$, define $\theta \cdot \mathbf{r} \in \mathbb{P}_T$ by

$$(\theta \cdot \mathbf{r})_j = \sum_{i:\, \theta(i)=j} r_i$$

($j \in \{1, \ldots, T\}$), where the sum is over all $i \in \{1, \ldots, S\}$ such that $\theta(i) = j$. For each $T \times T$ matrix $\mathbf{Z}$, define an $S \times S$ matrix $\mathbf{Z} \cdot \theta$ by

$$(\mathbf{Z} \cdot \theta)_{ii'} = Z_{\theta(i),\theta(i')}$$

($i, i' \in \{1, \ldots, S\}$).

Lemma A13 $^qD^{\mathbf{Z}\cdot\theta}(\mathbf{r}) = {}^qD^{\mathbf{Z}}(\theta \cdot \mathbf{r})$, *for all* $0 \leq q \leq \infty$.

Proof By continuity of $^qD^{\mathbf{Z}}$ in q (Proposition A2), it is enough to prove this when $q \neq 1, \infty$. We will use the convention that indices i, i' range over $1, \ldots, S$ and indices j, j' range over $1, \ldots, T$.

We have

$$((\mathbf{Z} \cdot \theta)\mathbf{r})_i = \sum_{i'} (\mathbf{Z} \cdot \theta)_{ii'} r_{i'} = \sum_{i'} Z_{\theta(i),\theta(i')} r_{i'},$$

$$(\mathbf{Z}(\theta \cdot \mathbf{r}))_j = \sum_{j'} Z_{jj'} (\theta \cdot \mathbf{r})_{j'} = \sum_{j'} \sum_{i':\, \theta(i')=j'} Z_{jj'} r_{i'} = \sum_{i'} Z_{j,\theta(i')} r_{i'},$$

so

$$((\mathbf{Z} \cdot \theta)\mathbf{r})_i = (\mathbf{Z}(\theta \cdot \mathbf{r}))_{\theta(i)}.$$

Hence

$$^qD^{\mathbf{Z}\cdot\theta}(\mathbf{r})^{1-q} = \sum_{i:\, r_i > 0} r_i ((\mathbf{Z} \cdot \theta)\mathbf{r})_i^{q-1} = \sum_{i:\, r_i > 0} r_i (\mathbf{Z}(\theta \cdot \mathbf{r}))_{\theta(i)}^{q-1}$$

$$= \sum_{j:\, (\theta\cdot\mathbf{r})_j > 0} \sum_{i:\, \theta(i)=j} r_i (\mathbf{Z}(\theta \cdot \mathbf{r}))_j^{q-1}$$

$$= \sum_{j:\, (\theta\cdot\mathbf{r})_j > 0} (\theta \cdot \mathbf{r})_j (\mathbf{Z}(\theta \cdot \mathbf{r}))_j^{q-1} = {}^qD^{\mathbf{Z}}(\theta \cdot \mathbf{r})^{1-q}.$$

The result follows. $\qquad\square$

The three elementary properties of diversity can be deduced. In each, q may take any value in the range $0 \leq q \leq \infty$.

Proposition A14 (Symmetry) *Let θ be a permutation of $\{1, \ldots, S\}$, let $\mathbf{Z}$ be an $S \times S$ matrix, and let $\mathbf{p} \in \mathbb{P}_S$. Define $\mathbf{Z}'$ and $\mathbf{p}'$ by $Z'_{ij} = Z_{\theta(i),\theta(j)}$ and $p'_i = p_{\theta(i)}$. Then $^qD^{\mathbf{Z}'}(\mathbf{p}') = {}^qD^{\mathbf{Z}}(\mathbf{p})$.*

Proof In the notation above, we have $\mathbf{Z}' = \mathbf{Z} \cdot \theta$ and $\mathbf{p} = \theta \cdot \mathbf{p}'$, since

$$(\theta \cdot \mathbf{p}')_j = \sum_{i: \, \theta(i)=j} p'_i = \sum_{i: \, \theta(i)=j} p_{\theta(i)} = p_j.$$

Hence by Lemma A13,

$$^qD^{\mathbf{Z}'}(\mathbf{p}') = {}^qD^{\mathbf{Z} \cdot \theta}(\mathbf{p}') = {}^qD^{\mathbf{Z}}(\theta \cdot \mathbf{p}') = {}^qD^{\mathbf{Z}}(\mathbf{p}),$$

as required. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proposition A15 (Absent species) *Let $\mathbf{Z}$ be an $S \times S$ matrix and let $\mathbf{p} \in \mathbb{P}_S$ with $p_S = 0$. Write $\mathbf{Z}'$ for the restriction of $\mathbf{Z}$ to the first $(S-1)$ species; that is, $\mathbf{Z}$ is the $(S-1) \times (S-1)$ matrix given by $Z'_{ij} = Z_{ij}$. Write $\mathbf{p}' = (p_1, \ldots, p_{S-1}) \in \mathbb{P}_{S-1}$. Then $^qD^{\mathbf{Z}'}(\mathbf{p}') = {}^qD^{\mathbf{Z}}(\mathbf{p})$.*

Proof Let $\theta \colon \{1, \ldots, S-1\} \to \{1, \ldots, S\}$ be the embedding $\theta(i) = i$. Then $\mathbf{Z}' = \mathbf{Z} \cdot \theta$ and $\mathbf{p} = \theta \cdot \mathbf{p}'$, so the result follows as in the previous proof. $\qquad\square$

Proposition A16 (Identical species) *Let $\mathbf{Z}$ be an $S \times S$ matrix such that $Z_{i,S} = Z_{i,S-1}$ and $Z_{S,i} = Z_{S-1,i}$ for all i. Let $\mathbf{p} \in \mathbb{P}_S$. Write $\mathbf{Z}'$ for the restriction of $\mathbf{Z}$ to the first $(S-1)$ species, and define $\mathbf{p}' \in \mathbb{P}_{S-1}$ by*

$$p'_j = \begin{cases} p_j & \text{if } j < S-1 \\ p_{S-1} + p_S & \text{if } j = S-1. \end{cases}$$

Then $^qD^{\mathbf{Z}'}(\mathbf{p}') = {}^qD^{\mathbf{Z}}(\mathbf{p})$.

Proof Define a function $\theta \colon \{1, \ldots, S\} \to \{1, \ldots, S-1\}$ by

$$\theta(i) = \begin{cases} i & \text{if } i \leq S-1 \\ S-1 & \text{if } i = S. \end{cases}$$

Then $\mathbf{Z} = \mathbf{Z}' \cdot \theta$ and $\mathbf{p}' = \theta \cdot \mathbf{p}$. The result follows from Lemma A13. $\qquad\square$

The final properties to be proved are those from group 3: 'effect of species similarity on diversity'.

Proposition A17 (Monotonicity) *Let $\mathbf{Z}$ and $\mathbf{Z}'$ be $S \times S$ matrices with $Z_{ij} \leq Z'_{ij}$ for all i, j. Then $^qD^{\mathbf{Z}}(\mathbf{p}) \geq {}^qD^{\mathbf{Z}'}(\mathbf{p})$, for all $0 \leq q \leq \infty$ and $\mathbf{p} \in \mathbb{P}_S$.*

Proof By continuity in q (Proposition A2), it is enough to prove this when $q \neq 1, \infty$. We have $(\mathbf{Z}\mathbf{p})_i \leq (\mathbf{Z}'\mathbf{p})_i$ for all i.

If $0 \leq q < 1$ then x^{q-1} is decreasing in $x > 0$ and $y^{1/(1-q)}$ is increasing in $y > 0$. Hence

$$(\mathbf{Zp})_i \leq (\mathbf{Z'p})_i \text{ for all } i \implies (\mathbf{Zp})_i^{q-1} \geq (\mathbf{Z'p})_i^{q-1} \text{ for all } i \text{ such that } p_i > 0$$
$$\implies \sum_{i:\, p_i > 0} p_i (\mathbf{Zp})_i^{q-1} \geq \sum_{i:\, p_i > 0} p_i (\mathbf{Z'p})_i^{q-1}$$
$$\implies {}^q D^{\mathbf{Z}}(\mathbf{p}) \geq {}^q D^{\mathbf{Z'}}(\mathbf{p'}).$$

If $q > 1$ then x^{q-1} is increasing in x and $y^{1/(1-q)}$ is decreasing in y, and a similar argument applies. $\qquad\square$

Proposition A18 (Naive model) *Let $\mathbf{Z}$ be a similarity matrix, $\mathbf{p} \in \mathbb{P}_S$, and $0 \leq q \leq \infty$. Then ${}^q D(\mathbf{p}) \geq {}^q D^{\mathbf{Z}}(\mathbf{p})$.*

Proof We have $I_{ij} = 0 \leq Z_{ij}$ for all $i \neq j$, and $I_{ii} = 1 = Z_{ii}$ for all i, so ${}^q D(\mathbf{p}) = {}^q D^{\mathbf{I}}(\mathbf{p}) \geq {}^q D^{\mathbf{Z}}(\mathbf{p})$ by Proposition A17. $\qquad\square$

Proposition A19 (Range) *Let $\mathbf{Z}$ be a similarity matrix, $\mathbf{p} \in \mathbb{P}_S$, and $0 \leq q \leq \infty$. Then $1 \leq {}^q D^{\mathbf{Z}}(\mathbf{p}) \leq S$.*

Proof Let $\mathbf{Y}$ be the similarity matrix with $Y_{ij} = 1$ for all i, j. Then $Z_{ij} \leq Y_{ij}$ for all i, j, so ${}^q D^{\mathbf{Z}}(\mathbf{p}) \geq {}^q D^{\mathbf{Y}}(\mathbf{p})$ by Proposition A17. But ${}^q D^{\mathbf{Y}}(\mathbf{p}) = 1$ by repeated application of Proposition A16, or by direct calculation. Hence ${}^q D^{\mathbf{Z}}(\mathbf{p}) \geq 1$.

By Proposition A18, we have ${}^q D^{\mathbf{Z}}(\mathbf{p}) \leq {}^q D(\mathbf{p})$. But the Hill number ${}^q D(\mathbf{p})$ is maximized at the uniform distribution $\mathbf{p} = (1/S, \ldots, 1/S)$, where, being an effective number, it takes the value S. Hence ${}^q D^{\mathbf{Z}}(\mathbf{p}) \leq S$. $\qquad\square$

In the last two properties (naive model and range), we explicitly assumed that $\mathbf{Z}$ was a similarity matrix, not merely a relatedness matrix. The following example shows that they can fail without that assumption.

Example A20 In Proposition A7, the mean phylogenetic diversity of Chao et al. (2010) was expressed as ${}^q D^{\mathbf{Z}}(\boldsymbol{\pi})$ for a suitable matrix $\mathbf{Z}$ and vector $\boldsymbol{\pi}$ (equations (A.4) and (A.5)). When the phylogenetic tree is ultrametric, $\mathbf{Z}$ is a similarity matrix; hence, mean phylogenetic diversity for ultrametric trees satisfies all the properties above. But when the tree is not ultrametric, the mean phylogenetic diversity can be greater than the number of species, contrary to the naive model and range properties. An example is given in the supplement to Chao et al. (2010). Another example is the tree of Fig. A1(iii). There,

$$\overline{T} = 0.6 \times 1 + 0.4 \times 2 = 1.4,$$

and the mean phylogenetic diversity of order 0 is

$$\frac{1}{\overline{T}} \times (\text{total branch length}) = \frac{1}{1.4} \times 3 = 2.142\ldots,$$

which is greater than 2, the number of species. Similarly, short calculations show that the mean phylogenetic diversities of orders 1 and 2 are also greater than the number of species.

We finish with some facts about diversity profiles.

13

Proposition A21 *Let $\mathbf{Z}$ be an $S \times S$ matrix and $\mathbf{p} \in \mathbb{P}_S$. Then ${}^q D^{\mathbf{Z}}(\mathbf{p})$ is decreasing in q. That is, whenever $0 \leq q \leq q' \leq \infty$, we have ${}^q D^{\mathbf{Z}}(\mathbf{p}) \geq {}^{q'} D^{\mathbf{Z}}(\mathbf{p})$.*

It may be that the diversity profile is constant: e.g. if $\mathbf{Z} = \mathbf{I}$ and $\mathbf{p} = (1/S, \ldots, 1/S)$ then ${}^q D^{\mathbf{Z}}(\mathbf{p}) = S$ for all q.

Proof This follows from standard results on generalized means: Theorems 5 and 16 of Hardy et al. (1952). $\square$

The final result states that if two communities have the same naive diversity profiles then their relative abundance vectors are the same, except perhaps for the order in which the relative abundances p_i are listed.

Proposition A22 *Let $\mathbf{p}, \mathbf{p}' \in \mathbb{P}_S$ and suppose that ${}^q D(\mathbf{p}) = {}^q D(\mathbf{p}')$ for all $0 < q < \infty$. Then $(p'_1, \ldots, p'_S)$ is a permutation of $(p_1, \ldots, p_S)$.*

Proof We prove by induction on S that if $\mathbf{p}, \mathbf{p}' \in \mathbb{P}_S$ with $p_1 \leq \cdots \leq p_S$, $p'_1 \leq \cdots \leq p'_S$ and ${}^q D(\mathbf{p}) = {}^q D(\mathbf{p}')$ for all $q > 1$, then $\mathbf{p} = \mathbf{p}'$. Clearly this holds for $S = 1$. Now let $S \geq 2$, and write $f(q) = {}^q D(\mathbf{p}) = {}^q D(\mathbf{p}')$.

We have $p_S = 1/\lim_{q \to \infty} f(q) = p'_S$. If $p_S = p'_S = 1$ then $\mathbf{p} = \mathbf{p}' = (0, \ldots, 0, 1)$. Otherwise, we may define $\mathbf{r}, \mathbf{r}' \in \mathbb{P}_{S-1}$ by

$$\mathbf{r} = \left(\frac{p_1}{1 - p_S}, \ldots, \frac{p_{S-1}}{1 - p_S} \right)$$

and similarly $\mathbf{r}'$. Then for all $q > 1$,

$$ {}^q D(\mathbf{r}) = (1 - p_S)^{\frac{q}{q-1}} \left(\sum_{i=1}^{S-1} p_i^q \right)^{\frac{1}{1-q}} = (1 - p_S)^{\frac{q}{q-1}} (f(q)^{1-q} - p_S^q)^{\frac{1}{1-q}}. $$

But since $p_S = p'_S$, this expression is also equal to ${}^q D(\mathbf{r}')$. Hence by inductive hypothesis, $\mathbf{r} = \mathbf{r}'$; that is, $p_i = p'_i$ for all $i < S$. $\square$

References

Allen, B., M. Kon, and Y. Bar-Yam, 2009. A new phylogenetic diversity measure generalizing the Shannon index and its application to phyllostomid bats. *The American Naturalist* **174**:236–243.

Chao, A., C.-H. Chiu, and L. Jost, 2010. Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society B* **365**:3599–3609.

Faith, D. P., 1992. Conservation evaluation and phylogenetic diversity. *Biological Conservation* **61**:1–10.

Hardy, G., J. E. Littlewood, and G. Pólya, 1952. Inequalities. Cambridge University Press, Cambridge, second edition.

Hurlbert, S. H., 1971. The nonconcept of species diversity: A critique and alternative parameters. *Ecology* **52**:577–586.

Ricotta, C. and L. Szeidl, 2006. Towards a unifying approach to diversity measures: Bridging the gap between the Shannon entropy and Rao's quadratic index. *Theoretical Population Biology* **70**:237–243.

Smith, W. and J. F. Grassle, 1977. Sampling properties of a family of diversity measures. *Biometrics* **33**:283–292.