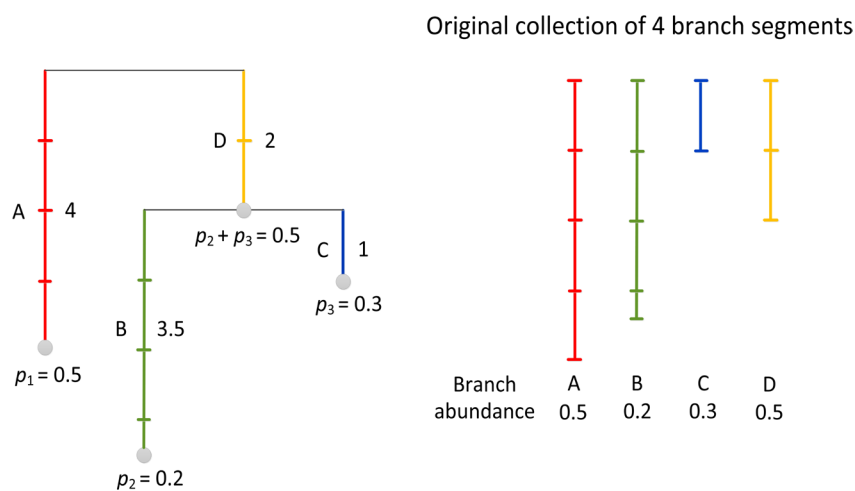## SUPPLEMENTAL APPENDIX 2: TWO SIMPLE ILLUSTRATIVE EXAMPLES

## 1. PHYLOGENETIC DIVERSITY

We use the following assemblage of 3 species and their phylogenetic tree to illustrate our unified

framework based on counting the effective number of phylogenetic entities, and to explain our

notation and terminology.



Original collection of 4 branch segments

| Branch abundance | A | B | C | D |
|---|---|---|---|---|
| | 0.5 | 0.2 | 0.3 | 0.5 |

_____

New collection of 10.5 entities, each entity is regarded as a "species" in a framework of Hill numbers
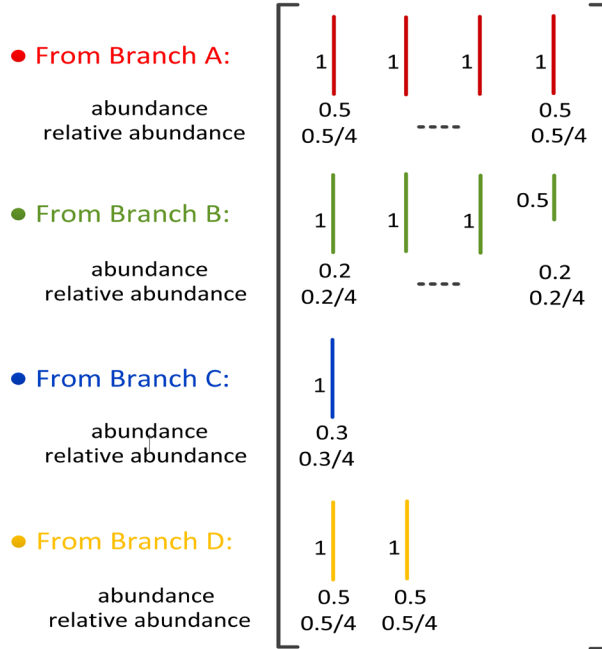


**1.1. The Original Collection *C* of Branch Segments**

Assume in the assemblage there are 3 species indexed by $\{1, 2, 3\}$ with relative abundances $p_1 = 0.5$, $p_2 = 0.2$, and $p_3 = 0.3$. The original collection $C$ consists of all branch segments in the phylogenetic tree. Here $C$ consists of 4 branch segments with branch length as shown by the number along each branch segment. The attribute value $v_u$ for each segment $u$ is its length: $v_1 = 4$, $v_2 = 3.5$, $v_3 = 1$, $v_4 = 2$. Each branch segment $i$ is weighted by $a_i$ the numerical importance of its descendants in the present-day assemblage (i.e., the summed relative abundance of all species descended from the branch). As in the main text, we refer to $a_i$ as branch abundance for simplicity. As shown in the above figure, we have $a_1 = p_1 = 0.5$, $a_2 = p_2 = 0.2$, $a_3 = p_3 = 0.3$, and $a_4 = p_2 + p_3 = 0.5$. Here the sum of all these $a_i$'s exceeds unity. The reason we consider this original collection in our derivation is for intuitive interpretation (see Section 1.4 below) and also for notational simplicity (i.e., we need an index running over this collection).

**1.2. The New Collection of Phylogenetic Entities or "Species"**

As discussed in Section 3 of the main text, there are two essential assumptions in the ordinary species-based Hill numbers. First, all species are assumed to be equally distinct. Second, only species relative abundances are involved in the formula of ordinary Hill numbers; thus species abundances or weights must be properly normalized to calculate Hill numbers.

Note that in the original collection of branch segments, the 4 branches are not phylogenetically equally distinct because they have different values of branch lengths. The basic idea in our unified approach is to first consider a new collection of phylogenetically equally distinct entities so that they can be regarded as equally distinct "species" in the framework of Hill numbers. As defined in the main text, a unit-length branch is regarded as a phylogenetic entity in our unified approach. In other words, all phylogenetic entities are equally distinct so that they play the role of "species" (Faith 1992, 2013) in ordinary species-based diversity.

The branch with length of 4 (Branch A in the figure) is counted as 4 entities (red little segments in the above figure); the branch with length of 3.5 (Branch B) is counted as 3.5 entities (green little segments in the above figure); and analogously for the other two branches. For this example, the 4 branch segments with attribute values 4, 3.5, 1 and 2 are respectively counted as 4, 3.5, 1 and 2 entities in the new collection. Therefore, we have a new collection of 10.5 phylogenetic entities as shown in the above figure. All entities are phylogenetically equally distinct. The branch abundance $a_i$ is also indicated under each entity. As indicated in the main text, the number of entities in our unified framework is allowed to be any real number, not necessarily an integer.

**1.3. The Total Abundance of Entities and the Relative Abundance of Each Entity**

Next we determine the relative abundance of each entity in the collection by proper normalization. The total abundance over all the entities in the new collection is

$$\overline{V} = \sum_{u \in C} v_u a_u \, ,$$

which is also the mean attribute value defined in the main text. For the specific example, we have $\overline{V} = 4 \times 0.5 + 3.5 \times 0.2 + 1 \times 0.3 + 2 \times 0.5 = 4$.

We then divide each abundance $a_i$ by the total abundance to obtain the relative abundance of each entity. Each of the four 4 entities (corresponding to the branch of length 4 with abundance 0.5) has relative abundance $0.5 / \overline{V} = 0.125$; each of the 3.5 entities (corresponding to the branch of length 3.5 with abundance 0.2) has relative abundance $0.2 / \overline{V} = 0.05$; 1 entity (corresponding to the branch of length 1 with abundance 0.3) with relative abundance $0.3 / \overline{V} = 0.075$, and each of the 2 entities (corresponding to the branch of length 2 with abundance 0.5) with relative abundance $0.5 / \overline{V} = 0.125$. Hence here the sum of the relative abundances of all entities is unity, so that we can apply ordinary Hill numbers to this assemblage of 10.5 "species". This type of normalization for phylogenetic diversity was first proposed in Chao et al. (2010) with an illustrative example in figure 1 of their paper.

**1.4 The Attribute Diversity and Its Interpretation**

Our unified approach is to treat the collection of 10.5 entities as a usual assemblage with 10.5 equally distinct "species" (each is weighted by its relative abundance) in ordinary species-based diversity. The attribute diversity is the effective number of entities or the effective total branch length, because each entity has a unit of length. This attribute diversity is identical to the phylogenetic diversity developed in Chao et al. (2010).
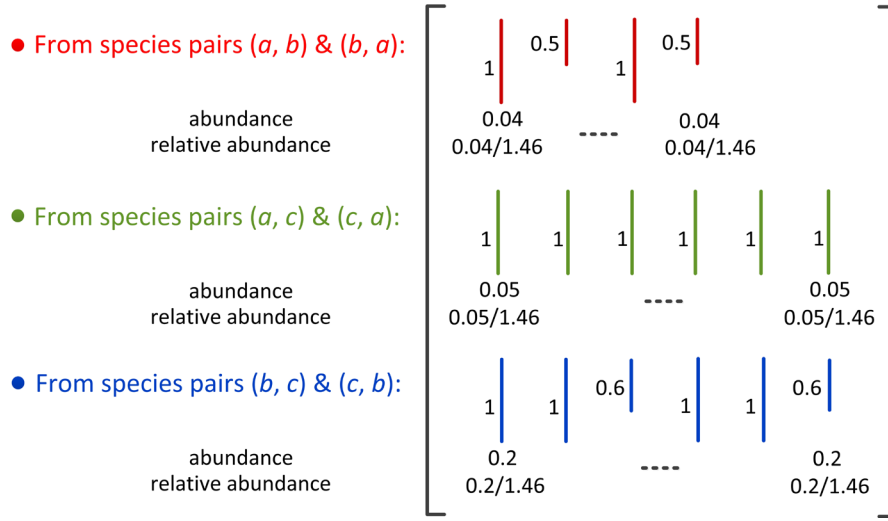
The attribute diversity (Equations 5b and 6a of the main text) or Hill numbers based on the collection of 10.5 entities with their relative abundances $a_i / \overline{V}$ are: $^0 AD(\overline{V}) = 10.5$, $^1 AD(\overline{V}) = 9.76$, and $^2 AD(\overline{V}) = 9.26$. We can interpret these values based on the collection of entities: Shannon diversity is the same as that of a collection with 9.76 equally abundant entities; and Simpson diversity is the same as that of a collection with 9.26 equally abundant entities. However, it is more intuitive to give an interpretation based on the original collection. That is, the phylogenetic diversity of order 1 of the actual assemblage is the same as that of an idealized assemblage with $^1 AD(\overline{V}) / \overline{V} = 9.76 / 4 = 2.44$ equally common and phylogenetically equally distinct species or lineages all with lineage length of 4; The phylogenetic diversity of order 2 of the actual assemblage is the same as that of an idealized assemblage with $^2 AD(\overline{V}) / \overline{V} = 9.25 / 4 = 2.31$ equally common and phylogenetically equally distinct species or lineages all with lineage length of 4.

## 2. FUNCTIONAL DIVERSITY

We use the following distance matrix for an assemblage of 3 species to illustrate our unified framework based on counting the effective number of functional entities.

$$
\begin{array}{c|ccc}
 & a & b & c \\
\hline
\begin{matrix} a \\ (0.1) \end{matrix} & 0 & 1.5 & 3 \\
\begin{matrix} b \\ (0.4) \end{matrix} & 1.5 & 0 & 2.6 \\
\begin{matrix} c \\ (0.5) \end{matrix} & 3 & 2.6 & 0
\end{array}
$$

Original collection of 9 species pairs

$(a, b)$ & $(b, a)$    $d_{ab} = d_{ba} = 1.5$

$(a, c)$ & $(c, a)$    $d_{ac} = d_{ca} = 3$

$(b, c)$ & $(c, b)$    $d_{bc} = d_{cb} = 2.6$

$(a, a), (b, b), (c, c)$   $d_{aa} = d_{bb} = d_{cc} = 0$

New collection of 14.2 entities, each entity is regarded
as a "species" in a framework of Hill numbers

● From species pairs (*a*, *b*) & (*b*, *a*):

|   | 0.5 |   | 0.5 |
| 1 |     | 1 |     |

abundance      0.04    ----    0.04
relative abundance    0.04/1.46      0.04/1.46

● From species pairs (*a*, *c*) & (*c*, *a*):

| 1 | 1 | 1 | 1 | 1 | 1 |

abundance      0.05    ----    0.05
relative abundance    0.05/1.46      0.05/1.46

● From species pairs (*b*, *c*) & (*c*, *b*):

| 1 | 1 | 0.6 | 1 | 1 | 0.6 |

abundance      0.2    ----    0.2
relative abundance    0.2/1.46      0.2/1.46

## 2.1. The Original Collection *C* of Species Pairs

Assume the 3 species in an assemblage are indexed by {*a*, *b*, *c*} with relative abundances $p_a = 0.1$, $p_b = 0.4$, and $p_c = 0.5$. The distance between any two species are shown in the above distance matrix. Here *C* consists of 9 species pairs: 6 species pairs with non-zero distances and 3 species pairs with zero distance. The 6 non-zero distances in the collection *C* have attribute values: $d_{ab} = d_{ba} = 1.5$, $d_{ac} = d_{ca} = 3$, and $d_{bc} = d_{cb} = 2.6$. Each species pair is weighted by the product of the relative abundances of the two species involved ($p_a p_b$, $p_a p_c$, and $p_b p_c$) as shown in the above figure.

## 2.2. The New Collection of Functional Entities or "Species"

In the original collection of species pairs, the 9 species pairs are not functionally equally distinct because they have different values of species pairwise distances. Our idea is to first consider a new collection of functionally equally distinct entities. As defined in the main text, a species pair

with a unit of pairwise distance is regarded as a functional entity. In other words, all entities are functionally equally distinct, so they play the role of "species" in the framework of Hill numbers.

A species pair (*a*, *b*) with distance of 1.5 is counted as 1.5 entities, and the same set of entities for the species pair (*b*, *a*). The species pair (*a*, *c*) with distance of 3 is counted as 3 entities, and the same set of entities for the species pair (*c*, *a*). The species pair (*b*, *c*) with distance of 2.6 is counted as 2.6 entities, and the same set of entities for the species pair (*c*, *b*). All pairs of same-species do not contribute any entity in the collection because the distance for any pair of same-species is 0. (If intraspecific variability exists such that the distance of a same-species pair is nonzero, then any pair of same-species will also contribute to entities in the same manner as described.) For the example, we have a collection of 14.2 functional entities or "species" as shown in the above figure. The corresponding abundance is also indicated under each entity. Here the total abundance over the 14.2 entities is not unity.


### 2.3. The Total Abundance of Entities and the Relative Abundance of Each Entity
Next we determine the relative abundance of each entity in the new collection by proper normalization. The total abundance over all 14.2 entities is identical to Rao's quadratic entropy:

$$\overline{V} = \sum_{u \in C} v_u a_u = \sum_{i,j} d_{ij} p_i p_j ,$$

which is also the mean attribute value defined in the main text. For the specific example, we have $\overline{V} = 1.46$.

We then divide each abundance ($p_a p_b$, $p_a p_c$, and $p_b p_c$) by the total abundance to obtain the relative abundance for each entity. In the collection of entities, each of the 1.5 entities (corresponding to the species pair (*a*, *b*) with a distance of 1.5) has relative abundance $p_a p_b / \overline{V} =$ 0.04/1.46 = 0.027; each of the 3 entities (corresponding to the species pair (*a*, *c*) with a distance

of 3) has relative abundance $p_a p_c / \overline{V} = 0.05/1.46 = 0.034$, ..., etc. Hence here the sum of the relative abundances of all 14.2 entities is unity, so that we can apply Hill numbers to this assemblage of 14.2 entities.

## 2.4. The Attribute Diversity and Its Interpretation

Our unified approach is to treat the collection of 14.2 entities as a usual assemblage with 14.2 equally distinct "species" (each is weighted by its relative abundance) in ordinary species-based diversity. The attribute diversity is the effective number of entities or the effective total sum of functional distances between any two species, because the attribute value of each entity is unity. This attribute diversity is identical to the functional diversity developed in Chiu and Chao (2014).

The attribute diversity (Equations 5b and 7a of the main text) or Hill numbers based on the collection of 14.2 entities with their relative abundances $p_i p_j / \overline{V}$ are: $^0 AD(\overline{V}) = 14.2$, $^1 AD(\overline{V}) = 11.08$, and $^2 AD(\overline{V}) = 9.36$. We can interpret these values based on the collection of entities: Shannon diversity is the same as that of a collection with 11.08 equally abundant entities; and Simpson diversity is the same as that of a collection with 9.36 equally abundant entities. It is more intuitive to give an interpretation based on the original collection. That is, the functional diversity of order 1 of the actual assemblage is the same as that of an idealized assemblage with $^1 AD(\overline{V}) / \overline{V} = 11.08 / 1.46 = 7.59$ equally common and functionally equally distinct species pairs with a constant species pairwise distance of 1.46. Also, the functional diversity of order 2 of the actual assemblage is the same as that of an idealized assemblage with $^2 AD(\overline{V}) / \overline{V} = 9.36/1.46 = 6.41$ equally common and functionally equally distinct species pairs with a constant species pairwise distance of 1.46.

**LITERATURE CITED**

Chao A, Chiu C-H, Jost L. 2010. Phylogenetic diversity measures based on Hill numbers. *Philos. Trans. R. Soc. B* 365:3599–609

Chiu C-H, Chao A. 2014. Distance-based functional diversity measures and their decomposition: a framework based on Hill numbers. *PLOS ONE* 9(7):e100014

Faith DP. 1992. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* 61:1–10

Faith DP. 2013. Biodiversity and evolutionary history: useful extensions of the PD phylogenetic diversity assessment framework. *Ann. N.Y. Acad. Sci.* 1289:69–89