

Chun-Huo Chiu, Lou Jost, and Anne Chao. 2013. Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs*.

APPENDIX B. Diversity profile and strong Replication Principle for Hill numbers and phylogenetic Hill numbers

Diversity (Hill number) profile

A “diversity profile” is a single graph that plots the Hill number qD as a continuous function of the parameter q . This diversity profile completely characterizes the diversity of an assemblage with S species and relative abundances (p_1, p_2, \dots, p_S) . The diversity profile curve is a non-increasing function of q (Hill 1973). The mathematical proof of this monotonicity is given in Beck and Schlögl (1993, p. 53).

Consider the following four assemblage structures (50 species in each) with different degrees of evenness. The species relative abundances $(p_1, p_2, \dots, p_{50})$ are given below, $\sum_{i=1}^{50} p_i = 1$. In each case, we also give the CV (coefficient of variation), which is the ratio of standard deviation and mean of $(p_1, p_2, \dots, p_{50})$. The magnitude of CV quantifies the degree of heterogeneity (or unevenness) of the probabilities $(p_1, p_2, \dots, p_{50})$. When all probabilities are equal, CV = 0. A larger value of CV signifies higher degree of heterogeneity (or unevenness) among probabilities.

- (1) Completely even assemblage: $p_1 = p_2 = \dots = p_{50} = 0.02$ (CV = 0).
- (2) Slightly uneven assemblage: $p_1 = p_2 = \dots = p_{25} = 0.027, p_{26} = p_{27} = \dots = p_{50} = 0.013$ (CV = 0.34).
- (3) Moderately uneven assemblage: $p_1 = p_2 = \dots = p_{25} = 0.0364, p_{26} = p_{27} = \dots = p_{50} = 0.00363$ (CV = 0.83).
- (4) Highly uneven assemblage: $p_1 = 0.225, p_2 = 0.202, p_3 = 0.180, p_4 = 0.157, p_5 = 0.134, p_6 = p_7 = \dots = p_{50} = 0.0022$ (CV = 2.74).

The diversity profile for the four assemblages with different degree of unevenness are shown in Fig. B1. This profile is sensitive to the species abundance distribution. The slope of the curve reflects the unevenness of species relative abundances. The more uneven the distribution of relative abundances, the more steeply the curve declines. For completely even relative abundances, the curve is a constant at the level of species richness. The information of species richness and relative abundances (p_1, p_2, \dots, p_S) is completely characterized by a diversity profile which depicts Hill numbers as a function of order q .

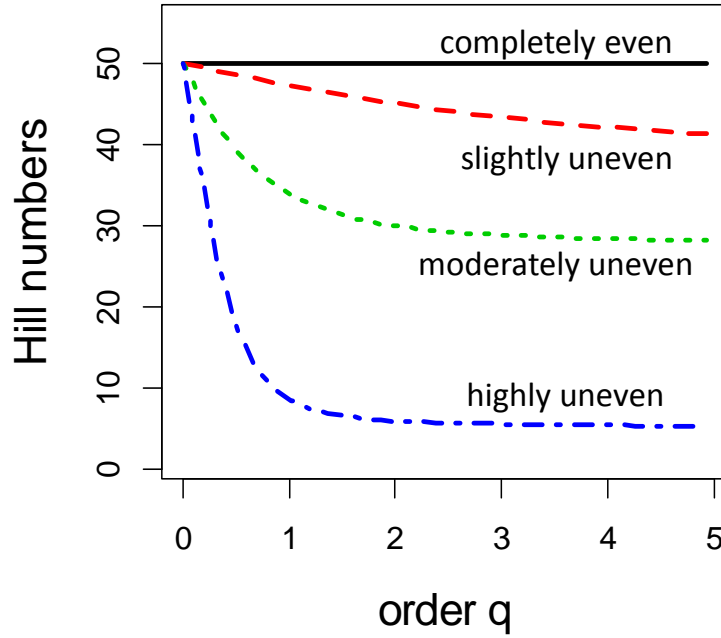


Fig. B1. Diversity profiles ($0 \leq q \leq 5$) of four assemblages with different degree of unevenness.

Phylogenetic diversity profiles

We consider four hypothetical rooted phylogenetic trees, each with four species. The four trees vary in structure from simple to highly subdivided; Tree (a) presents the simplest case of maximally distinct lineages. The branch length is shown along each branch. See Fig. B2.

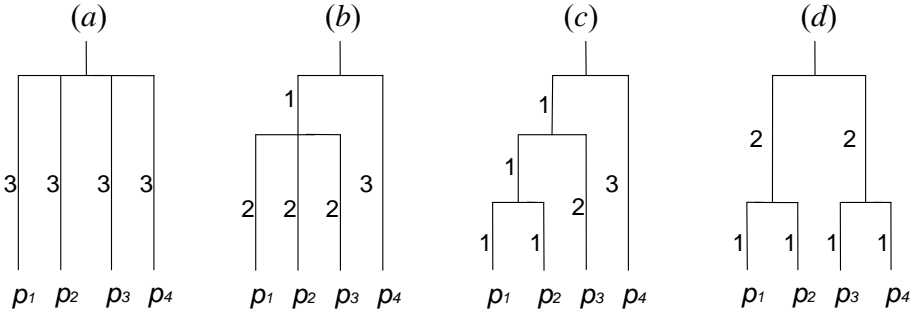
For each tree, we consider two species abundance distributions:

- (1) Species are equally common: $(p_1, p_2, p_3, p_4) = (0.25, 0.25, 0.25, 0.25)$.
- (2) Species abundances are uneven: $(p_1, p_2, p_3, p_4) = (0.1, 0.2, 0.3, 0.4)$.

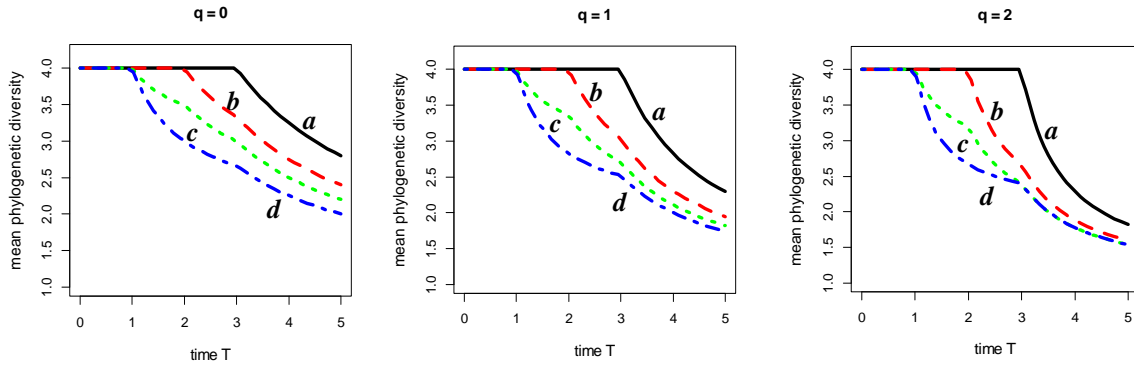
As discussed in the main text, the following two types of profiles completely characterize phylogenetic diversity:

- (1) The first type of diversity profile is obtained by plotting ${}^qPD(T)$ or ${}^q\overline{D}(T)$ as functions of T separately for $q = 0, 1$, and 2 ; see Fig. B2.
- (2) The other type of diversity profile is obtained by plotting ${}^qPD(T)$ or ${}^q\overline{D}(T)$ as a function of order q , for a selected value of temporal perspective T . See Fig. B3 for the profiles for $T = 1, 2, 3$.

Both tree structure and species abundances have effects on the patterns of the two types of profiles.



(1) Phylogenetic diversity profiles as a function of time for species abundances $(p_1, p_2, p_3, p_4) = (0.25, 0.25, 0.25, 0.25)$



(2) Phylogenetic diversity profiles as a function of time for species abundances $(p_1, p_2, p_3, p_4) = (0.1, 0.2, 0.3, 0.4)$

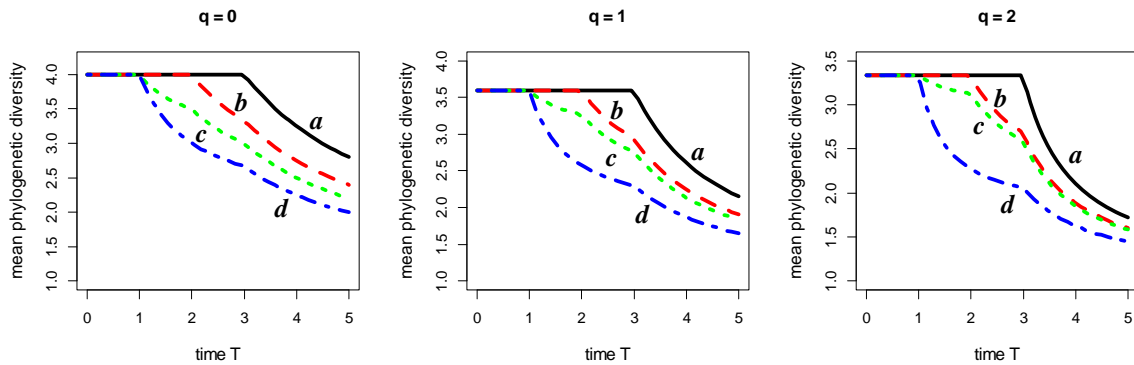
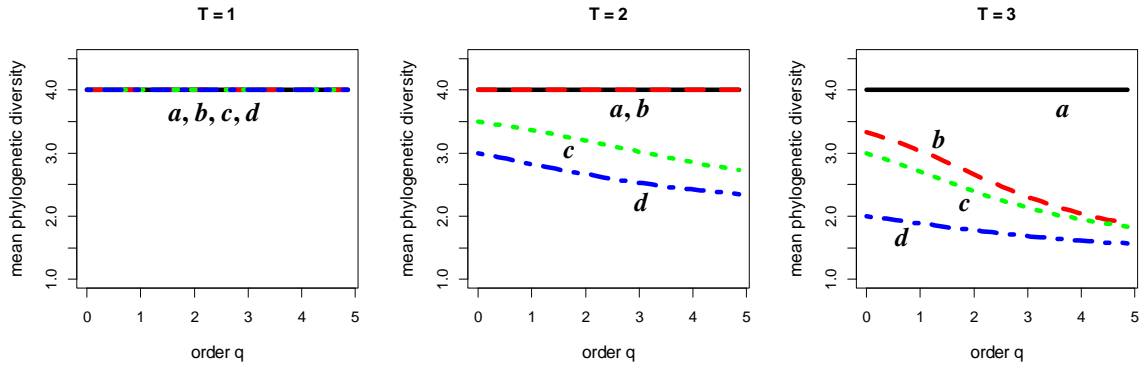


Fig. B2. Phylogenetic diversity profile as a function of time based on the mean diversity ${}^q\overline{D}(T)$ for $q = 0, 1$ and 2 and $0 < T < 5$ for the four phylogenetic trees and two species abundance distributions. The plots show that the effect of tree structures and the diversity ordering follows (a) \geq (b) \geq (c) \geq (d) for the two species abundance distributions.

- (1) Phylogenetic diversity profiles as a function of the order q for species abundances $(p_1, p_2, p_3, p_4) = (0.25, 0.25, 0.25, 0.25)$



- (2) Phylogenetic diversity profiles as a function of the order q for species abundances $(p_1, p_2, p_3, p_4) = (0.1, 0.2, 0.3, 0.4)$

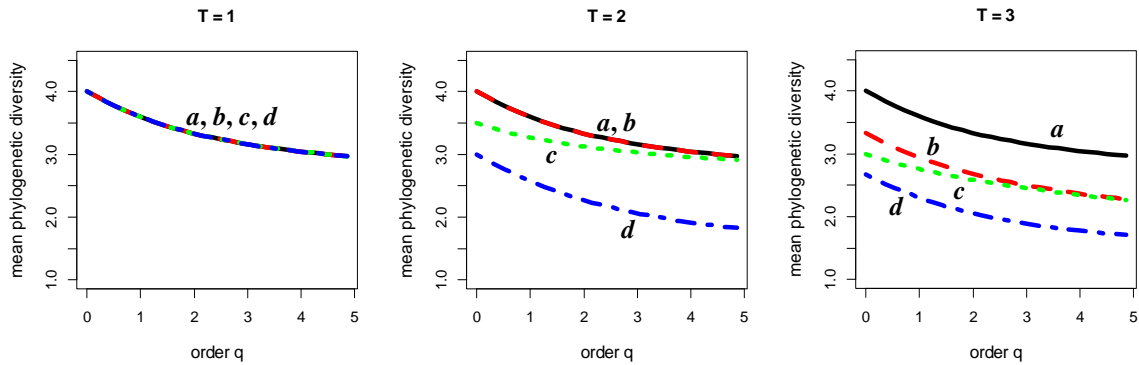


Fig. B3. Phylogenetic diversity profile as a function of order of q for a fixed perspective time $T = 1, 2$, and 3 based on the mean diversity ${}^q\overline{D}(T)$ for $0 \leq q \leq 5$ for four phylogenetic trees and two species abundance distributions.

Strong Replication Principle for Hill numbers

Suppose there are S_k species in the k th assemblage, $k = 1, 2, \dots, N$. Let the relative abundance of the i th species in the k th assemblage be p_{ik} . Assume that for a given value of q , each of the N assemblages have the same Hill number X , implying $\sum_{i=1}^{S_k} p_{ik}^q = X^{1-q}$ for each $k=1, 2, \dots, N$. When the N completely distinct assemblages are pooled with equal weight for each assemblage, each species relative abundance in the pooled tree becomes p_{ik}/N . Then the Hill number for the pooled assemblage becomes

$$\left\{ \sum_{k=1}^N \sum_{i=1}^{S_k} (p_{ik}/N)^q \right\}^{1/(1-q)} = (N^{1-q} \times X^{1-q})^{1/(1-q)} = N \times X.$$

Therefore, the strong Replication Principle holds for Hill numbers.

Strong Replication Principle for phylogenetic Hill numbers (mean phylogenetic diversity)

Suppose we have N completely distinct assemblages (no lineages shared between assemblages in the interval $[-T, 0]$, though lineages within an assemblage may be shared; see Fig. 1 of the main text), all with the same mean diversity ${}^q\overline{D}(T) = X$ in the time interval $[-T, 0]$, based on an ultrametric tree. Then we can prove the following strong replication principle: If these assemblages are pooled in equal proportions, the mean diversity of the pooled assemblages is $N \times X$. Suppose that in the tree for the k th assemblage, the branch set is $\mathbf{B}_{T,k}$ (we omit T in the subscript and just use \mathbf{B}_k in the following proof for notational simplicity) with branch lengths $\{L_{ik}; i \in \mathbf{B}_k\}$ and corresponding node abundances $\{a_{ik}; i \in \mathbf{B}_k\}$, $k = 1, 2, \dots, N$. The N assemblages have the same mean diversity X over the time interval $[-T, 0]$, implying $\sum_{i \in \mathbf{B}_k} (L_{ik}/T) a_{ik}^q = X^{1-q}$ for all $k = 1, 2, \dots, N$. When the N assemblages are pooled with equal weight for each tree, each node abundance a_{ik} in the pooled tree becomes a_{ik}/N . Then the measure ${}^q\overline{D}(T)$ for the pooled tree becomes

$$\left\{ \sum_{k=1}^N \sum_{i \in \mathbf{B}_k} \frac{L_{ik}}{T} \left(\frac{a_{ik}}{N} \right)^q \right\}^{1/(1-q)} = \{N^{1-q} \times X^{1-q}\}^{1/(1-q)} = N \times X.$$

The above proof can be directly extended to non-ultrametric trees (Chao et al. 2010). In our proof of this replication principle, the N assemblages must have the same time T (or for non-ultrametric trees, the same mean quantity \overline{T}), but may have different numbers of species, and the tree structures of the N assemblages can be totally different. See Fig. 1 of the main text for an example.

LITERATURE CITED

- Beck, C., and F. Schlögl. 1993. Thermodynamics of chaotic systems: an introduction, Cambridge University Press, Cambridge, UK.
- 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.
- Hill, M. 1973. Diversity and evenness: A unifying notation and its consequences. Ecology 54:427–432.