**Anne Chao, Nicholas J. Gotelli, T. C. Hsieh, Elizabeth L. Sander, K. H. Ma, Robert K. Colwell, and Aaron M. Ellison. 2013. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecology Monographs*.**

## Appendix J: An example: sample-size- and coverage-based Shannon diversity curves may exhibit inconsistent patterns

Here we provide an example to show that the sample-size- and coverage-based Shannon diversity curves may exhibit inconsistent patters. Consider two assemblages (Assemblage A and Assemblage B), with frequency counts given in Table J1.

For Assemblage A, there are 201 species represented by 2100 individuals. It consists of one dominant species (i.e., one species with 1000 individuals i.e., $f_{1000} = 1$), 100 moderately abundant species (each with 10 individuals, i.e., $f_{10} = 100$) and 100 very rare species (each with 1 individual, i.e., $f_1 = 100$). The true (asymptotic) Hill numbers for this data set are ${}^0D = 201$, ${}^1D = 26.15$ and ${}^2D = 4.37$.

For Assemblage B, there are 1120 species represented by 9550 individuals. It consists of 20 very abundant species, 100 rare species (each with 5 individuals, i.e., $f_5 = 100$) and 1000 very rare species (each with 1 individual, i.e., $f_1 = 1000$). The true (asymptotic) Hill numbers for this data set are ${}^0D = 1120$, ${}^1D = 37.72$ and ${}^2D = 15.05$.

If we only compare the asymptotic Hill numbers, Assemblage B is clearly more diverse for species richness, Shannon diversity and Simpson diversity. The expected sample-size and coverage-based diversity accumulation curves in Fig. J1 facilitate the comparison for any finite sample sizes and any coverage value less than unity; see Chao and Jost (2012) and Colwell et al. (2012). For species richness, the two types of curves exhibit consistent patterns, as theoretically proved by Chao and Jost (2012). When sample sizes (or expected coverages) are relatively small or large, Assemblage B has higher diversity, but when sample sizes (or expected coverages) are in between, Assemblage A is more diverse. This can be understood because Assemblage A has more moderately abundant species, whereas Assemblage B has more extremely abundant and rare species.

However, the two types of curves for Shannon diversity show inconsistent patterns. The sample-size-based Shannon diversity curves for the two assemblages do not cross (the curve of Assemblage B is higher), but the coverage-based curves cross twice. When the expected coverage is less than 55% or higher than 80%, Assemblage B has higher Shannon diversity; when coverage is between 55% and 80%, the ordering is reversed. For the Simpson diversity, the two types of curves give the same ordering for this example, but Example 1 discussed in the main text (Figs. 3b and 5b of the main text) exhibits inconsistent patterns.

Table J1. The species frequency counts and percentages of individuals for two assemblages. Assemblage A (201 species, 2100 individuals). The frequency counts (from abundant to rare species) are: $f_{1000} = 1$, $f_{10} = 100$, $f_1 = 100$;
Assemblage B (1120 species, 9550 individuals). The frequency counts (from abundant to rare species) are: $f_{1000} = f_{500} = f_{100} = f_{10} = 5$, $f_5 = 100$, $f_1 = 1000$.

| | Assemblage A | | Assemblage B | |
|---|---|---|---|---|
| $i$ | Frequency count, $f_i$ | Percentage of total individuals | Frequency count, $f_i$ | Percentage of total individuals |
| 1000 | 1 | 47.6 % | 5 | 52.4% |
| 500 | | | 5 | 26.2% |
| 100 | | | 5 | 5.2% |
| 10 | 100 | 47.6% | 5 | 0.5% |
| 5 | | | 100 | 5.2% |
| 1 | 100 | 4.8% | 1000 | 10.5% |

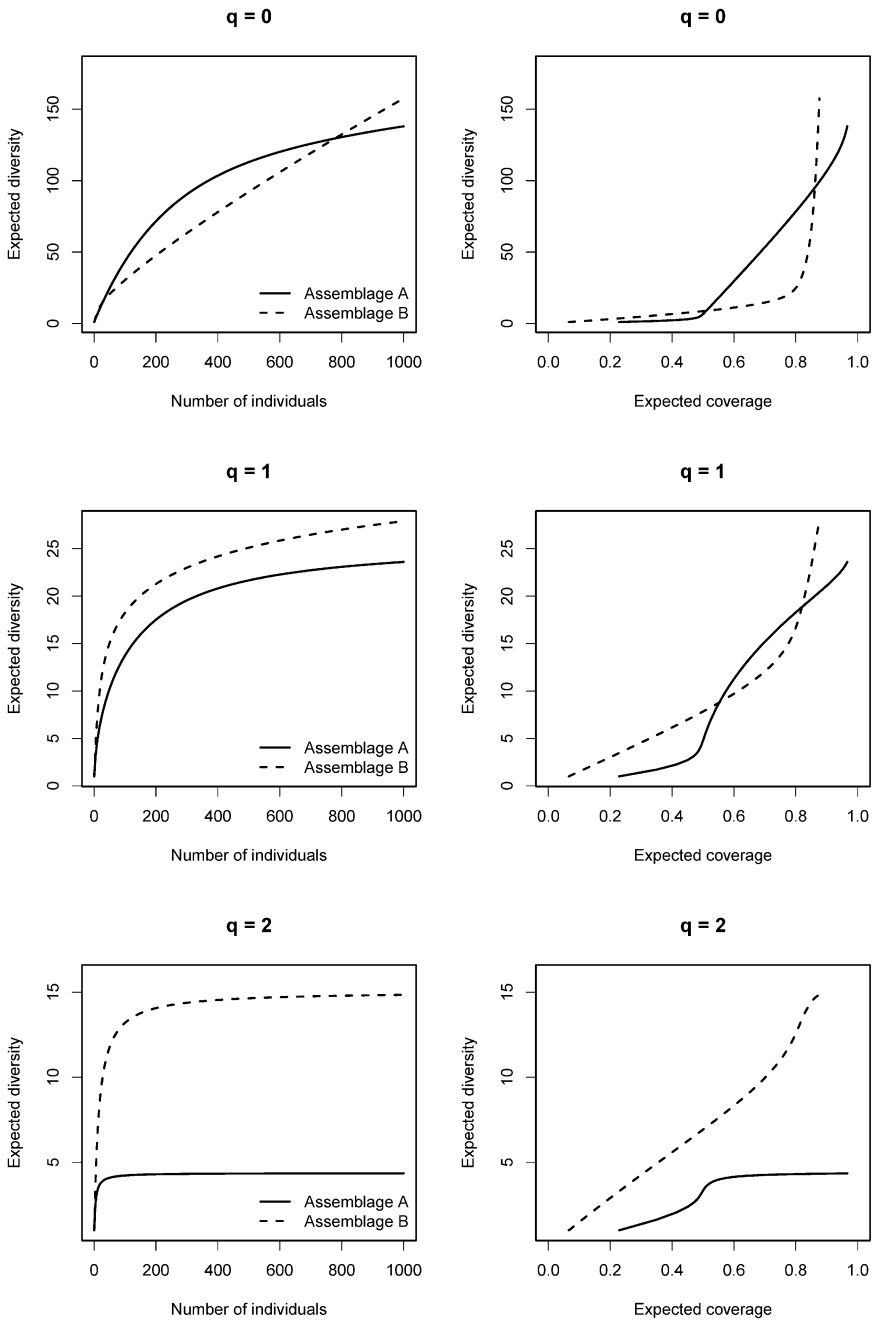


Fig J1. (Left panel) The sample-size-based expected diversity accumulation curves for species richness (the first row), Shannon diversity (the second row) and Simpson diversity (the third row) for the two assemblages described in Table J1. (Right panel) The corresponding coverage-based diversity accumulation curves.

LITERATURE CITED

Chao, A., and L. Jost. 2012. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. Ecology, 93:2533-2547.

Colwell, R. K., A. Chao, N. J. Gotelli, S. Y. Lin, C. X. Mao, R. L. Chazdon, and J. T. Longino. 2012. Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. Journal of Plant Ecology 5:3-21.