Appendix S2: Decomposition of Simpson index

We use Pélissier and Couteron’s [1] organization of data to derive the decomposition. We define an indicator $1_{sk}$ such that it takes a value of 1 if individual $k$ belongs to species $s$, and of zero otherwise. $1_{sk}$ follows a Bernoulli distribution: either an individual belongs to species $s$ or not. $\sum_{k=1}^{n} 1_{sk}$ follows a binomial distribution of variance $np_s(1-p_s)$. A standard partitioning of variance can be applied. Within each community, variance is $n_i p_{si}(1-p_{si})$. The expectation of $\sum_{k=1}^{n} 1_{sk}$ is $np_s$ for pooled data, and that of within-community $\sum_{k=1}^{n_i} 1_{sk}$ is $n_i p_{si}$. We have then:

$$p_s(1-p_s) = \sum_i \frac{n_i}{n} [p_{si}(1-p_{si}) + (p_{si} - p_s)^2]$$

Summation of $p_s(1-p_s)$ over species provides Simpson’s entropy:

$$2H_\gamma = \sum_i \frac{n_i}{n} 2H_\alpha + \sum_i \frac{n_i}{n} \sum_s (p_{si} - p_s)^2$$

Assuming $2H_\gamma = 2H_\alpha + 2H_\beta$ and $\alpha$ entropy is the weighted sum of within-community $\alpha$ entropy values, the second term of (2) can be identified to $\beta$ entropy in the classical additive partitioning. $2H_\beta$ is the weighted sum of contributions of communities, denoted $2H_\beta = \sum_s (p_{si} - p_s)^2$. This additive partitioning of entropy is that of Nei [2] among others. As shown by Jost [3,4], $2H_\beta$ is constrained by $2H_\alpha$ since $\gamma$ entropy is limited to 1. Using $2H_\beta$ or $G_{ST} = 2H_\beta / 2H_\gamma$ as a measure of differentiation has been shown to be erroneous [5-7].

References: