Resolving the species pool dependence of beta-diversity using coverage-based rarefaction

Thore Engel (ORCID)[1,2], Shane A. Blowes (ORCID)[1,2], Daniel J. McGlinn (ORCID)[3], Felix May (ORCID)[4], Nicholas J. Gotelli (ORCID)[5], Brian J. McGill (ORCID)[6], Jonathan M. Chase (ORCID)[1,2]

Affiliations

1. German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, 04103, Leipzig, Germany
2. Institute of Computer Science, Martin Luther University Halle-Wittenberg, 06120, Halle (Saale), Germany
3. Department of Biology, College of Charleston, Charleston, SC
4. Leuphana University of Lüneburg, Lüneburg, Germany
5. Department of Biology, University of Vermont, Burlington, Vermont
6. School of Biology and Ecology, and Senator George J. Mitchell Center of Sustainability Solutions, University of Maine, Orono, Maine

Corresponding author: Thore Engel (thore.engel@posteo.de)

Abstract

1. Biodiversity is non-randomly distributed in space and understanding how spatial structure of species diversity responds to ecological, biogeographic and anthropogenic drivers is one of the major quests of modern ecology. However, metrics of community differentiation such as Whittaker's beta-diversity fail to unambiguously capture species turnover when they are compared across assemblages with changing species pool sizes, species abundance distributions and numbers of individuals. This is due to an effect of changing sample completeness.

2. Here, we propose a method that combines the benefits of individual-based and coverage based rarefaction to standardize beta-diversity estimates by sample completeness at the gamma-scale. We test the performance of our metric, β_C, using simulations of spatially explicit communities that change both species pool size and intraspecific spatial aggregation. Additionally, we re-analyse the "Gentry" plot dataset to investigate patterns of tree beta-diversity along a latitudinal gradient.

3. We find that our metric, β_C, captures changes in intraspecific aggregation independent of changes in the species pool size, while Whittaker's beta-diversity and standardized estimates based on the number of individuals respond to species pool size as well as spatial aggregation. Our coverage-based estimates of beta-diversity do not exhibit a latitudinal pattern in the empirical tree dataset.

4. Our simulations show that coverage-based standardization of beta-diversity offers a powerful tool for the unraveling of spatial and compositional biodiversity change. Potential applications of the approach range from a better description of basic biogeographic diversity patterns to the consolidation of local and regional diversity trends in the current biodiversity crisis.

Keywords: beta-diversity, coverage, measurements of biodiversity, null models, rarefaction, sample completeness, spatial aggregation

Introduction

Organisms are non-randomly distributed across the globe and understanding spatial patterns of species diversity and composition from ecological samples is one of the most central questions driving ecological research (Gaston, 2000; McGill, 2011a; Worm and Tittensor 2018). A fundamental feature of biodiversity that makes it different from other response variables in ecology is that it is scale-dependent (Rosenzweig, 1995; McGill 2011b; Chase and Knight 2013). That is, local biodiversity increases at a decelerating rate with increasing spatial scale (e.g., the species-area relationship), and therefore many patterns of biodiversity depend critically on scale (e.g., Rahbek 2005; Harrison et al. 2006; Chase et al., 2018). As a result of the growing awareness of the importance of spatial scale in biodiversity measurements and interpretations of possible mechanisms, there has been a steady increase in interest in measuring and comparing spatial beta-diversity (Anderson et al. 2011; De Cáceres et al. 2012, Sreekar et al., 2018; Xing & He 2019). Beta-diversity, as first defined by Whittaker (1960, 1972), offers a mathematical link between local (i.e. alpha) and regional (i.e. gamma) species diversity, and quantifies how species composition varies from site to site (i.e., species turnover) within a region. In addition to providing important insights regarding how biodiversity scales with area, spatial beta-diversity can, for example, shed light onto the metacommunity processes that shape biological assemblages (Chase & Myers, 2011), inform biodiversity conservation (Socolar, Gilroy, Kunin & Edwards, 2016), and help understand the provisioning of ecosystem functions and services (Mori, Isbell, & Seidl, 2018).

Although beta-diversity is conceptually quite important, in practice, its measurement has been highly ambiguous. First, there are multiple metrics and approaches for measuring beta-diversity that are often used interchangeably, despite their sensitivity to very different aspects of community differences (Legendre, Borcard & Peres-Neto, 2005; Baselga, 2010; Tuomisto 2010a, b; Anderson et al., 2011; Chao & Chiu, 2016). Second, meaningful inference from changes of beta-diversity is often complicated due to the entanglement of multiple components of species diversity that collectively influence sample-derived estimates of beta-diversity. Beta-

diversity is commonly thought to represent non-random spatial structure in species diversity (i.e. intraspecific spatial aggregation or species turnover in space). But it also is inherently a bridge between the local and regional scales, and as such it is unavoidably intertwined with the nature of biodiversity at these scales. For example, like virtually all diversity metrics, it responds to the size of the regional species pool, the shape of the regional species abundance distribution (SAD), and the number of individuals captured by the samples (McGill, 2011, Chase & Knight, 2013, Chase et al, 2018). Here, we argue that these non-spatial components of beta-diversity confound beta-diversity through their effects on sample completeness (Colwell & Coddington, 1994). For instance, regions with very large species pools typically exhibit high beta-diversity simply because local samples capture a small and incomplete portion of their diversity, which can lead to strong, but random, differentiation among samples (Kraft et al., 2011). Likewise, undersampling can inflate beta-diversity when there are many undetected rare species in an assemblage (e.g., when species evenness is very low) or when the sample size is very small (i.e., few individuals are captured by a sample) (Barwell, Isaac & Kunin 2015).

Although such sampling effects are ubiquitous in ecological studies (Colwell & Coddington, 1994; Gotelli & Colwell, 2001), they are not well addressed with respect to beta-diversity (Wolda 1981; Beck, Holloway, & Schwanghart, 2013, Chao & Chiu, 2016). There have been attempts to develop beta-diversity indices that account for unseen species (Chao, Chazdon, Colwell, & Shen, 2005). However, evidence for their robustness has been questioned using simulated and empirical data (Beck Holloway, & Schwanghart, 2013; Cardoso, Borges, & Veech, 2009). While single-scale diversity estimates can be standardized using coverage-based rarefaction which accounts for differences in species pool size (e.g., Chao and Jost 2012), this technique has not been applied to beta-diversity. Instead, empirical studies often adopt null model approaches to correct for species pool effects (and associated levels of undersampling), when evaluating beta-diversity patterns (Chase et al., 2011). The intent of such null models is to disentangle how much of the observed beta-diversity is due to non-random patterns of species distributions (i.e., patterns of intraspecific aggregation). While useful, for example, in disentangling the drivers of beta-diversity along large-scale gradients in gamma diversity (e.g., Kraft et al., 2011; Myers et al., 2013; Stegen et al. 2013, Xu, Chen, Liu & Ma, 2015), there is debate about how to best generate the null distributions for these models. For example, null distributions can be generated maintaining alpha diversity, gamma diversity, abundance structures and/or numbers of individuals (Mori, Fujii, Kitagawa & Koide, 2015; Qian, Wang & Zhang, 2012; Kraft et al., 2012; Brocklehurst, Day & Fröbisch, 2018). Although there is an emerging consensus that abundance-based null models are superior to incidence-based ones (Tucker et al., 2016), these approaches have been criticized because they ignore the completeness of the samples (Ulrich et al. 2017, Sreekar et al. 2018).

Here, we develop a rarefaction approach that is conceptually similar to an abundance-based null model in order to standardize beta-diversity estimates by the completeness of the gamma scale sample. To achieve this, we build on our previous development of beta-diversity within the Measurements of Biodiversity ("MoB") framework (Chase at el, 2018; McGlinn et al., 2019), and combine it with coverage-based rarefaction (Chao and Jost 2012). In the MoB framework, which itself is built upon previous work with individual-based rarefaction curves (e.g., Gotelli and

Colwell 2001), we compare diversity values from individual-based rarefaction curves taken from smaller (local) and larger (regional) scale samples to estimate the degree to which species are aggregated (clumped) in the landscape as a measure of beta-diversity (see also Olszewski 2004, Dauby and Hardy 2012). While the resulting metrics of beta-diversity are useful for comparing patterns of intraspecific aggregation within a given regional species pool (i.e., comparing observational or experimental results from a single region), we will show that they still retain an undersampling bias when used to compare across species pools of different sizes. As a solution, we show that a coverage-based beta-diversity metric enables comparisons of the magnitude of intraspecific aggregation notwithstanding changes in the species pool. We test our method on simulated spatial point patterns with different degrees of spatial structure (intraspecific spatial aggregation) and varying species pool sizes, and apply it to a case study of an empirical dataset of trees along a latitudinal gradient.

Calculating beta-diversity from rarefaction curves

Individual-based rarefaction (IBR) is a common tool to standardize species richness estimates from different sample sizes (Hurlbert, 1971; Gotelli & Colwell, 2001). Furthermore, IBR can also be used to examine beta-diversity by comparing curves estimated from smaller (alpha diversity) samples nested within larger (gamma diversity) samples (Olszewski 2004, Dauby and Hardy 2012, Chase et al. 2018, McGlinn et al. 2019). IBR curves estimate the non-linear scaling relationship between the number of individuals in a sample and species richness (i.e., rarefied richness). The shape of the curve is determined by the SAD (i.e., species pool size and evenness), and the slope at any point along the curve is related to the sample completeness corresponding to the sample size at that point. For a given sample, the IBR curve can be calculated using the following formula derived from the hypergeometric probability distribution (Hurlbert, 1971):

$$S_n = S_{obs} - \sum_{X_i \geq 1} \frac{\binom{N-X_i}{n}}{\binom{N}{n}} \qquad \text{(Eqn. 1)}$$

S_n is the rarefied richness, or the expected number of species found in a subsample of n individuals ($n<N$). S_{obs} is the observed total number of species, N is the total number of individuals in the sample and X_i is the number of individuals of the ith species.

By constructing IBR curves at two or more nested spatial scales, we can assess intraspecific spatial aggregation and beta-diversity. Like most classical approaches to beta-diversity, we define alpha as the mean diversity of the sample scale or the scale of a localized subset of samples, and gamma as the total diversity of multiple samples or local subsets of samples (ref). Accordingly, for a given assemblage with spatially replicated samples, the alpha scale IBR curve is made up of the mean S_n values across all samples (αS_n) and the gamma scale curve consists of the S_n values of the pooled samples (γS_n). While the alpha scale is affected by spatial turnover within the assemblage, the gamma scale breaks up any lower scale spatial structure by randomly accumulating individuals from all the samples. If species are distributed

4

randomly among samples (i.e., no aggregation), the alpha and gamma scale IBR curves sit on top of each other. Downward and upward deviations of the alpha scale curve, then, would be interpreted as intraspecific aggregation and overdispersion, respectively (Chase et al., 2018). The gamma scale IBR is conceptually very similar to abundance based null models (e.g. Kraft et al., 2011), but it uses an analytical formula for rarefaction rather than a shuffling algorithm.

Using IBR curves, Figure 1 illustrates how beta-diversity of a reference assemblage (Fig. 1A) responds to changes in the size of the species pool (Fig. 1B), the numbers of individuals (Fig. 1C) and intraspecific spatial aggregation (Fig. 1D). Traditional (i.e., Whittaker 1972) beta-diversity ($\beta = \frac{\gamma}{\alpha}$) is represented as the height ratio of the two curves at the respective right-hand end of the curves (dashed horizontal lines). In each of the four examples, there is beta-diversity, but for very different underlying reasons. Only the assemblage underlying Fig. 1D exhibits the sort of spatial turnover in species composition (due to aggregation) that we usually envision as an underlying mechanism for beta-diversity. In the other cases (A-C), beta-diversity only emerges due to sampling (i.e., a more-individuals effect between alpha and gamma scale). In order to evaluate the degree of intraspecific aggregation in a region, Chase et al. (2018) recommend that beta-diversity should be calculated from alpha and gamma scale richness values rarefied to a common number of individuals rather than from the observed species richness. The resulting ratio, termed β_{S_n}, quantifies the deviation of the alpha curve from the gamma curve at a common number of individuals (*n*) as:

$$\beta_{S_n} = \frac{\gamma_{S_n}}{\alpha_{S_n}} \qquad \text{(Eqn. 2)}$$

Where *n* corresponds to the number of individuals observed at the alpha scale. β_{S_n} is related to the metric developed by Olszewski 2004 (see also Dauby and Hardy 2012), who compared the differences of these curves at their base, rather than at *n* individuals as we have done here. β_{S_n} is also conceptually similar to abundance-based null models that tease apart the influence of aggregation from sampling effects in determining patterns of beta-diversity when comparing among sites within a regional species pool (e.g., Chase 2010, Myers et al., 2015). When assemblages have a random spatial structure, β_{S_n} is expected to equal 1 regardless of species pool and sample size (Fig 1 A-C). Conversely, β_{S_n} values larger than 1 reflect spatial aggregation or species turnover among sites in the region (Fig 1 D).
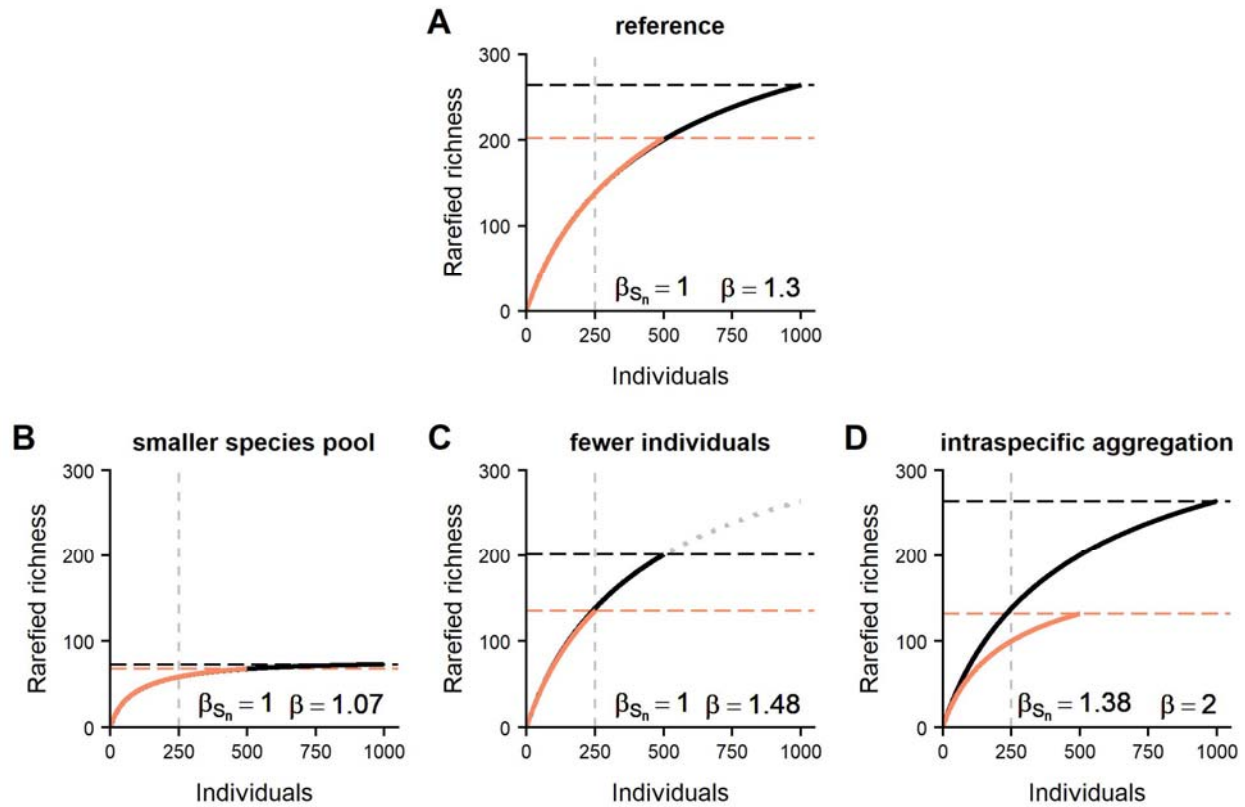
Figure 1: Examples of two-scale individual-based rarefaction (IBR) curves for (A) a hypothetical species pool size of 450, and how they respond to, (B) reduced species pool size / altered SAD (species pool size of 80), (C) reduced numbers of individuals and (D) and changes to patterns of within species aggregation. Orange curve: alpha scale, black curve: gamma scale. Dashed lines represent observed species richness at alpha and gamma scale and Whittaker's beta-diversity (β) can be illustrated as the height ratio of the two.   values are calculated for N = 250 individuals on all panels (dashed vertical lines). Dotted grey curve in panel C: reference IBR (from A) for easier comparison.

While   is a useful metric for estimating the degree of aggregation among treatments or observations within a given species pool, it suffers from the same bias associated with sampling that influences most measures of biodiversity. Thus, as we will illustrate below,   is less useful in comparing the degree of aggregation among regions where the size of the regional species pool changes (e.g., along biogeographical gradients).

The deviation between alpha and gamma scale IBR curves is due to spatial structure (i.e.   ≠ 1). However, the magnitude of this deviation is contingent on the size of the species pool. To visualize this interaction, consider two assemblages each composed of two patches, but which differ in the size of their regional species pool (500 vs 100 species Fig. 2A and 2B, respectively). Supposing that both assemblages have complete species turnover between their respective patches (i.e., species are maximally aggregated at the patch scale), Figure 2 shows the IBR

6

curves that we would expect to find if we sampled 500 individuals from each patch in the large (Fig. 2A) and small (Fig. 2B) species pools. While both assemblages are maximally and equally structured at the patch scale, is substantially higher in the community with the small species pool than in the community with the large species pool. This is because for equal sample sizes (i.e., equal numbers of individuals), samples from a large species pool are less complete than samples from a small species pool (i.e., further from a hypothetical asymptote; Chao and Jost, 2012). Thus, there is a downward bias in that increases with decreasing completeness of the gamma-scale estimate (Fig. 2C).
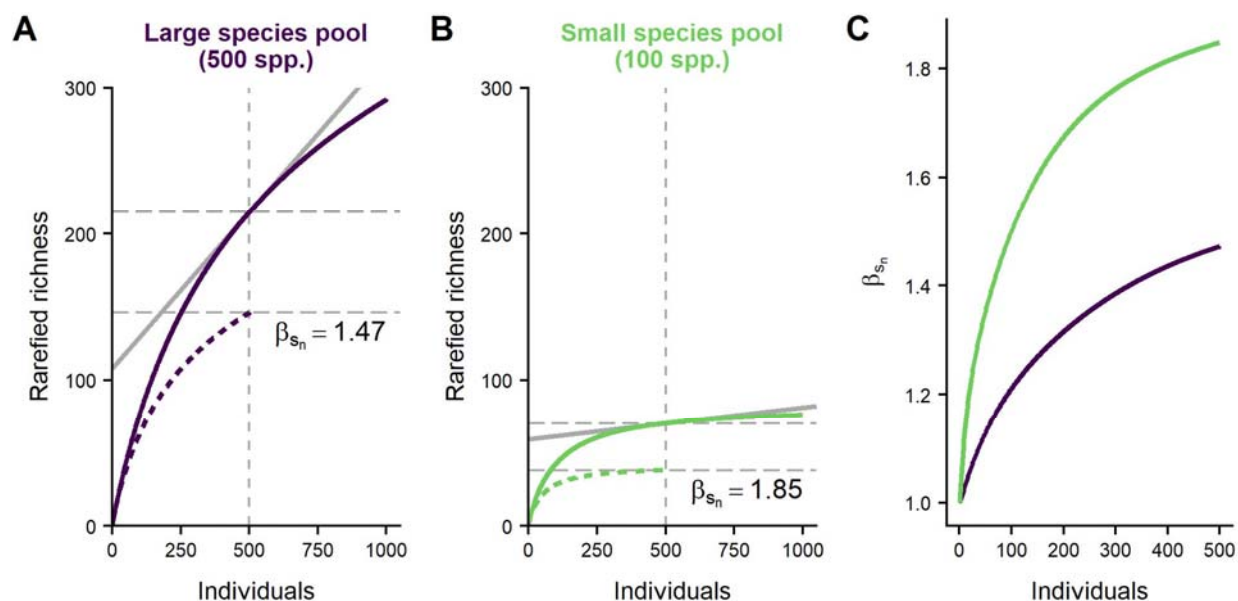

Figure 2: is affected by species pool size in aggregated communities. Two-scale IBR curves for (A) a large species pool and (B) a small species pool. Solid curve: gamma scale; dashed curve: alpha scale; dashed vertical grey line: number of individuals (n=500) used for the calculation of (i.e,. the ratio of the horizontal dashed lines) . Grey solid line shows the slope of the gamma scale IBR that relates to sample coverage (the steeper the slope of this line, the less complete the sample). (C) plotted as a function of the corresponding sample size (n).

Method

We propose an approach that accounts for the sample completeness of an assemblage, in order to evaluate the degree of aggregation leading to beta-diversity when species pool size differs. Following Chase et al. (2018), we use IBR within assemblages to standardize alpha and gamma diversity estimates to a common number of individuals *n*. This removes the influence of the more-individuals effect influencing beta-diversity that comes about due to the difference in sample size between the alpha and gamma scales. However, because we are interested in cases where the species pool size varies, rather than keeping sample-size constant across assemblages (via *n*), we instead maintain a consistent sample coverage at the gamma scale. Sample coverage is a measure of sample completeness that ranges from 0 to 1 and refers to the "proportion of the total number of individuals in a community that belong to the species

7

represented in the sample" (Chao & Jost, 2012). Coverage can be estimated from the number of rare species in a sample (Good, 1953; Chao & Shen, 2010). As it is directly related to the slope of the IBR curve, it can also be interpolated to any smaller than observed sample size, n, using the following equation (Chao & Jost, 2012) :

$$C_n = 1 - \sum_{X_i \geq 1} \frac{X_i}{N} \frac{\binom{N-X_i}{n}}{\binom{N}{n}} \qquad \text{(Eqn 3)}$$

C_n is the expected coverage for a subsample of sample size n. N is the total number of individuals in the sample and X_i is the number of individuals of the ith species. Eqn 3 can be used to assign coverage estimates to any sample size used in IBR, which is the basis for coverage-based rarefaction. For coverage-based standardization, one has to rearrange Eqn 3 to determine how many individuals (n) are necessary to obtain a given target coverage C_{target} (usually the smallest coverage of all communities). This is computed numerically by calculating C_n for every possible n between 1 and N, and choosing the n that minimises the difference between C_n and C_{target}. Subsequently, IBR is applied to standardize the diversity estimate to a sample size of n and thus the desired coverage level C_{target} (e.g. Hsieh, Ma, & Chao, 2016). Here, we adopt the same procedure to standardize beta-diversity.

Figure 3 illustrates the two-scale IBR curves corresponding to hypothetical large (Fig. 3A) and small (Fig. 3B) species pools. Again we quantify the deviation between gamma and alpha scale as β_{S_n} at a common n. However, we allow n to vary between the two scenarios to maintain a constant gamma-scale coverage (indicated by the slope of the tangential lines). The resulting pair of β_{S_n} values are very similar (compare with Fig. 2), which represents the fact that both scenarios are equally aggregated at the patch scale. The advantage of matching β_{S_n} values by gamma scale coverage becomes particularly clear when we consider the entire scaling relationship of β_{S_n} (Fig. 3 C). If we quantify β_{S_n} for every possible value of n and plot them against gamma-scale expected coverage, the values from large and small species pool fall on approximately the same line and the species pool dependence vanishes.

We refer to β_{S_n} controlled by sample coverage as β_C. To compare β_C across multiple assemblages (e.g. with different species pools), we suggest the following protocol:

1. Determine the appropriate target coverage value C_{target} for the standardization:
    1.1. For each assemblage j, determine the smallest number of individuals observed at the alpha scale and call it N_{min_j}
    1.2. Using Eqn 3, estimate the expected gamma-scale coverage C_n corresponding to N_{min_j} individuals.
    1.3. Let C_{target} be the smallest of the C_n values across all assemblages.

2. Calculate β_C:

2.1.   For each assemblage use the inverse of Eqn 3 to estimate the sample size n_j corresponding to a gamma-scale coverage of C_{target}. Exclude all assemblages for which $n_j \leq 1$.

2.2.   Rarefy gamma and alpha scale species richness to n_j individuals (Eqn 1) to get $^{\gamma}S_{n,j}$ and $^{\alpha}S_{n,j}$.

2.3.   Calculate β_C as:

$$\beta_{C_J} = \frac{^{\gamma}S_{n,j}}{^{\alpha}S_{n,j}} \quad \text{(Eqn 4)}$$

We implemented this procedure in an R package available at https://github.com/T-Engel/betaC. It provides the function "C_target" that can be used for steps 1.1 and 1.2. Furthermore, it contains the function "beta_C" that carries out step 2. These functions adopted code that was originally written by Hsieh et al. (2016) for the R package iNEXT.

Our approach can be applied to a wide range of ecological datasets where changes in spatial structure of species diversity and species pools are entangled along a common gradient. Like all IBR methods, our approach requires species abundance data (e.g., site by species matrices with abundance data). Every assemblage should have spatially replicated samples, so that one can define at least two nested sampling scales (alpha and gamma). We assume that the sampling design is standardized across all assemblages. This means there should be a consistent number of samples per assemblage and every sample should have the same effort (e.g., plot size, trap nights, etc.). Furthermore, we assume a consistent spatial extent. If the number of samples changes from one assemblage to the next, we suggest that users take a spatially constrained subset of the samples to keep extents as consistent as possible. Likewise, other deviations from the 'ideal' scenario, such as correcting for unequal sampling effort, can be accomplished prior to deriving the values described below. Similar to Whittaker's (1962) multiplicative beta, β_C is calculated as a ratio of gamma over alpha. This makes β_C more intuitive than metrics of beta-deviation that are calculated in null model approaches (e.g., Kraft et al. 2011) because it can be interpreted as an effective number of distinct communities (sensu, Jost, 2007).
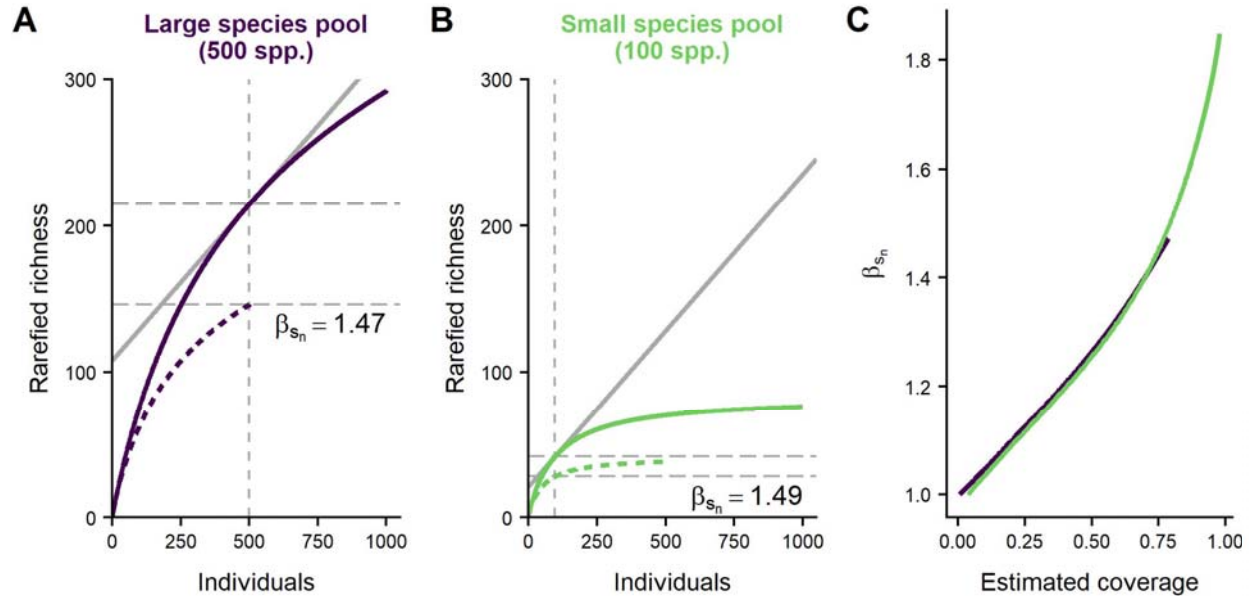
Figure 3: is unaffected by species pool size in aggregated communities when it is matched according to sample coverage (compare with figure 2). Two-scale IBR curves for (A) a large species pool and (B) a small species pool. Solid curve: gamma scale; dashed curve: alpha scale; dashed vertical grey line: number of individuals used for the calculation of (i.e. the ratio of the horizontal dashed lines). Grey solid line shows the slope of the gamma scale IBR that relates to sample coverage. C) plotted as a function of coverage calculated at the gamma scale.

Proof of concept using simulations

To test the performance of , we simulated spatially explicit assemblages that varied in the size of the species pool and the degree of intraspecific aggregation using multivariate spatial point patterns. We used the R package mobsim (May et al., 2018) to carry out the simulations. For each simulated assemblage, we first drew 4000 individuals from a lognormal SAD that was parameterized with a given species pool size and a coefficient of variation of 1. Then, we used the Thomas cluster process to distribute individuals in space, varying the degree of intraspecific spatial aggregation through the parameter for the number of conspecific clusters. The simulation was parameterized in a full-factorial design, where the species pool size encompassed every integer between 10 and 500, and the number of clusters was set to 1 (i.e., extreme intraspecific aggregation), 4, 10 and 20. We also implemented a random Poisson process to simulate completely random spatial distributions for species as an additional level of within species aggregation. Each combination of species pool and aggregation was replicated 3 times yielding a total of 7365 simulated communities. To sample from the regional communities, we placed 4 sample quadrats into each simulated community (Fig. 4 A, B). We calculated Whittaker's (−), and among the 4 samples and examined their response to changes in species pool size and aggregation. Following the protocol above, was set to 0.55 and the sample

10

size for β_{S_n} was 50. The R code for the simulation is available on github (https://github.com/T-Engel/betaC) and will be archived upon acceptance of this manuscript.

All three beta-diversity indices (Whittaker's β, β_{S_n}, and β_C) were influenced by intraspecific community aggregation. Additionally, Whittaker's β and β_{S_n} were affected by the changes in species pool size, whereas β_C was insensitive to this factor. While both Whittaker's β and β_{S_n} responded to the degree of spatial aggregation and species pool, they did so in contrasting ways. For Whittaker's β, sensitivity to the size of the species pool diminished with increasing aggregation; for β_{S_n}, the effect of the species pool increased with increasing aggregation (Fig. 4 C). Only β_C measures the spatial structure of the simulated communities independently of the species pool size, and is therefore recommendable for quantifying beta-diversity across regional species pools of varying sizes.
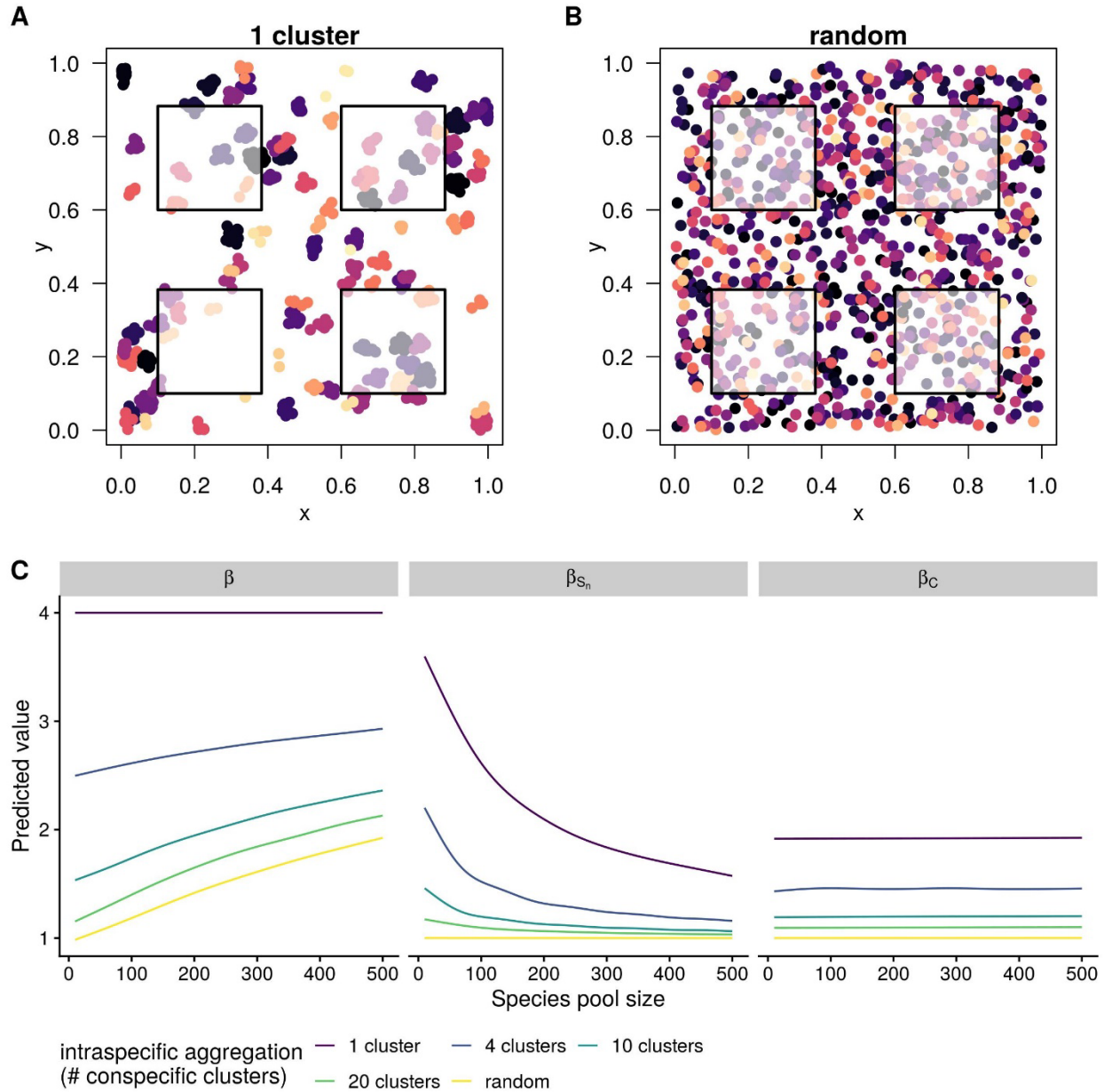
Figure 4: Simulated assemblages and the response of Whittaker's β, β_{S_n} and β_C to changes in aggregation and species pool size. (A) Assemblage with extreme within species aggregation (1 cluster per species) and (B) assemblage where species have random spatial distribution. Species pool, SAD and numbers of individuals are constant between (A) and (B). Squares represent sample quadrats. (C) Beta-diversity metrics calculated from the samples as a function of species pool size. Each line shows the predicted values from a GAM fit to each beta-diversity metric with species pool size as the predictor; the GAM estimated separate smoothers for each level of intraspecific aggregation.

Empirical case study

We applied our approach to a dataset that has been used frequently to examine patterns of beta-diversity across biogeographic (specifically latitudinal) gradients—the Gentry Forest plot dataset (Gentry, 1988; Phillips & Miller, 2002). This dataset has been used to disentangle changes of beta-diversity and species pool using different null model approaches (Kraft et al, 2011; Qian & Song, 2012; Qian et al., 2013; Xu et al., 2015). We obtained the dataset through the "BIEN" library in R (Maitner et al., 2018). We restricted our analysis to the plots located in the Americas in order to focus explicitly on the influence of latitude rather than other biogeographic effects, and only included plots where data from 10 subplots were available and where each subplot contained at least 5 stems. This resulted in 140 plots with a latitudinal range from -40.717°S to 45.553°N; total species richness was strongly negatively correlated with latitude ranging from 12 to 257 (Fig. 5A).

We defined the plot scale (i.e., all 10 subplots) as the gamma scale, and two neighboring subplots as the alpha scale. Previous analyses have used one subplot as the alpha scale, but we used two to achieve a better sample completeness (nevertheless, our results were qualitatively comparable regardless of the alpha-scale we used; not shown). We computed Whittaker's β, β_{S_n}(n = 12) and $\beta_C (C_{target}= 0.14)$. First, as previously shown (e.g., Kraft et al. 2011), we found that Whittaker's beta-diversity declined with latitude (Fig. 5B). In addition to a potential signal of spatial structure, this effect includes the differences in sample size between alpha and gamma scale, as well as changes in species pool size and associated changes in sample completeness. When we control for number of individuals using β_{S_n}, the trend is reversed and there is an increase in beta-diversity with increasing latitude (Fig. 5C). However, as we showed above, these results are still affected by the effect of the species pool and undersampling because sampling coverage among these plots differs. Fig. 5D shows the latitudinal gradient in coverage C_n corresponding to the sample size of 12 that was used for the calculation of β_{S_n}. In some of the temperate plots, 12 stems can correspond to 75% and more of the species pool estimated from the sample, whereas in many of the species-rich tropical plots this sample size only captures as little as 5% of the species (Fig. 5D). β_C, which controls for this difference in sample completeness, showed no change along the latitudinal gradient (Fig. 5E). We conclude that at the given scale there is no change of spatial aggregation along this gradient. But this does not mean that the spatial structure is random. As indicated by $\beta_C > 1$, we find a signal for spatial aggregation across the gradient even at this small scale and sample completeness.

The target coverage of 14 % is very low for this dataset, which supports the argument that these plots are most likely too small (especially in the tropics) to appropriately compare beta-diversity trends (Tuomisto & Ruokolainen, 2012). Nevertheless, they still provide an important test case for the analyses we advocate here, particularly given the previous analyses of these data and the discussion regarding appropriate null models that ensued (Kraft et al, 2011; Qian & Song, 2012; Tuomisto & Ruokolainen, 2012; Kraft et al., 2012; Qian et al., 2013; Xu et al., 2015). Although we come to qualitatively similar results as the abundance-based null models of Kraft et al. (2011) and Xu et al. (2015), our method has the advantage of explicitly incorporating an

estimate of sample completeness. This means that users are confronted with the quality of their data and have to interpret the results in the light of any limitations.
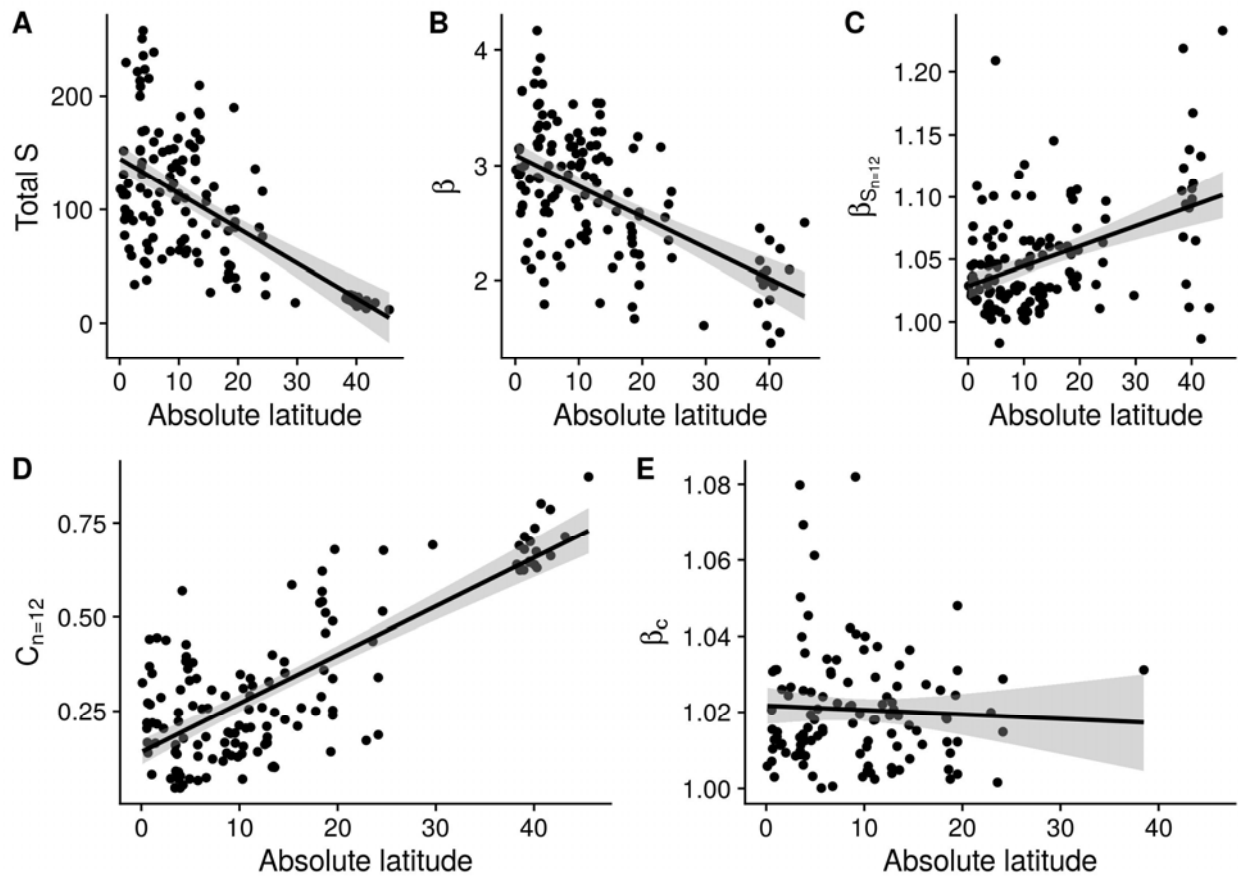

Figure 5: Entangled changes of species pool and spatial aggregation along the latitudinal gradient of Gentry forest plots. (A) Total species richness (i.e. gamma), (B) Whittaker's multiplicative , (C) (n = 12), D) the estimated coverage (i.e., sample completeness) corresponding to the sample size of n = 12, and E) (C = 0.14) as a function of absolute latitude. Solid lines show simple linear regressions. All except panel E were significant (p < 0.05).

Discussion and conclusions

Adjusting sample size from one location to the other to achieve similar sample completeness is not a new idea for the estimation of beta-diversity, but there has been a lack of statistical methods to achieve this (Colwell and Coddington, 1994). Building on previous work using rarefaction and coverage-based tools (e.g., Colwell and Coddington, 1994; Chao and Jost, 2012; Chao et al., 2014; Chase et al., 2018; McGlinn et al., 2019), we provide an approach that standardizes beta-diversity by sample coverage to quantify the degree of spatial structuring (aggregation). Our simulations of spatially explicit assemblages show that the resulting metric, is influenced by changes in spatial structure, but remains unaffected by changes in species pool size. This allows for comparisons of spatial structure along large biogeographic gradients

with changing species pools. In our empirical case study, we show that spatial structure in tree communities did not change along a latitudinal gradient of forest plots.

Although we standardize by sample coverage, it is nonetheless important to ensure that the samples have somewhat meaningful and comparable completeness. As our empirical case study shows, this can be challenging when the differences in species pool size are very large and sample sizes (e.g., plot size) are relatively small. Within our method, the target coverage value used for the standardization is set by the site with the lowest gamma-scale completeness, which for the Gentry Forest Plot data presented above (Gentry 1988, Phillips & Miller, 2002) was 0.14, which is quite low. For some of the species poor temperate plots, this meant that sample size had to be scaled down to as little as one or two individuals (i.e., n_j) to achieve the same value of coverage. From a theoretical point of view, the n_j value should be 2 or higher because only then can there be a deviation between alpha and gamma scale IBR curve; rarefying down to one individual always yields 1 species. As a result, many of the low diversity plots had to be excluded, as β_C for this very low coverage was not defined. In a separate analysis, we increased the target coverage value to 50 %, which meant that many of the species rich plots had to be excluded because their sample coverage was not high enough (see Supporting Information S1). Having controlled for variation in sample completeness, we find no systematic change of spatial structure along the latitudinal gradient. Nonetheless, we emphasize that our empirical case study merely serves as a proof of concept, and we share the concern that due to the small plot size, these data have limited power to reveal any changes in assembly processes along the latitudinal gradient (Tuomisto & Ruokolainen, 2012).

β_C quantifies intraspecific spatial aggregation whilst controlling for the fraction of the estimated species pool in the samples (i.e., coverage). Others have suggested that beta-diversity should be calculated from only a subset of the most dominant species in a sample, but this approach still responds to changes in species pool size and the SAD (Roden et al., 2018). In contrast, our coverage-based method makes use of the SAD of the samples (i.e., the shape of the IBR curve) to make inferences for a constant fraction of the species pool. Ulrich et al. (2017) argue that null models cannot disentangle species pool and aggregation unless they incorporate external data on the sizes of the relevant species pools. However, we argue that the sample itself can provide an estimate of its completeness (Good, 1953, Chao and Jost 2012). Nevertheless, coverage-based estimates, implicitly assume that there is an asymptote in species diversity (Chao and Jost, 2012). While this assumption is mathematically convenient and the asymptote can be approximated for a given sample, it cannot strictly be true (i.e., we will always find more species with more samples until the entire global pool is sampled; Williamson et al,. 2001). The other important assumption is that the gamma scale is a somewhat random subset of the species pool, which is probably not the case when there is spatial structure at larger scales and when the number of samples is relatively small. Nevertheless, as we presented it here, β_C does not require estimation of the asymptote itself, but merely uses interpolation to develop a useful approximation of sample completeness. Our approach could be extended to encompass extrapolation, but this should be done with caution (Chao et al., 2014).

In conclusion, we have developed an approach to address species pool dependence and associated sampling effects in beta-diversity. Unlike other null model approaches, we explicitly treat species pool dependence as a problem of undersampling, and address it by standardizing beta-diversity by sample coverage. As a result, this approach allows deeper insights into the scale-dependent and multivariate nature of biodiversity patterns (Chase et al., 2018). Applications could, for example, shed light onto the assembly processes that govern biological (meta-)communities along biogeographic gradients or contribute to a better understanding of the scale-dependent biodiversity trends observed during the current biodiversity crisis.

Acknowledgements

Author's contributions

TE conceived the idea. All authors helped to develop and refine the approach. TE, SB and JMC designed the simulations and the empirical case study. TE wrote the R code and the first draft of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data availability

R code and simulated data are available at https://github.com/T-Engel/betaC and will be archived upon acceptance of this manuscript.

References

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., ... & Harrison, S. P. (2011). Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology letters*, *14*(1), 19-28.

Barwell, L. J., Isaac, N. J., & Kunin, W. E. (2015). Measuring β-diversity with species abundance data. *Journal of Animal Ecology*, *84*(4), 1112-1122.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global ecology and biogeography*, *19*(1), 134-143.

Beck, J., Holloway, J. D., & Schwanghart, W. (2013). Undersampling and the measurement of beta diversity. *Methods in Ecology and Evolution*, *4*(4), 370-382.

16

Brocklehurst, N., Day, M. O., & Fröbisch, J. (2018). Accounting for differences in species frequency distributions when calculating beta diversity in the fossil record. *Methods in Ecology and Evolution*, *9*(6), 1409-1420.

Cardoso, P., Borges, P. A. V., & Veech, J. A. (2009). Testing the performance of beta diversity measures based on incidence data: The robustness to undersampling. *Diversity and Distributions*, *15*(6), 1081–1090.

Chao, A., & Chiu, C. H. (2016). Bridging the variance and diversity decomposition approaches to beta diversity via similarity and differentiation measures. *Methods in Ecology and Evolution*, *7*(8), 919-928.

Chao, A., & Jost, L. (2012). Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, *93*(12), 2533-2547.

Chao, A., & Shen, T. J. (2010). User's guide for program SPADE (Species prediction and diversity estimation). *Taiwan: National Tsing Hua University*.

Chao, A., Chazdon, R. L., Colwell, R. K., & Shen, T. J. (2005). A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology letters*, *8*(2), 148-159.

Chao, A., Chiu, C. H., & Hsieh, T. C. (2012). Proposing a resolution to debates on diversity partitioning. *Ecology*, *93*(9), 2037-2051.

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014). Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, *84*(1), 45–67.

Chase, J. M., & Knight, T. M. (2013). Scale-dependent effect sizes of ecological drivers on biodiversity: why standardised sampling is not enough. *Ecology letters*, *16*, 17-26.

Chase, J. M., & Myers, J. A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical transactions of the Royal Society B: Biological sciences*, *366*(1576), 2351-2363.

Chase, J. M., Kraft, N. J., Smith, K. G., Vellend, M., & Inouye, B. D. (2011). Using null models to disentangle variation in community dissimilarity from variation in α-diversity. *Ecosphere*, *2*(2), 1-11.

Chase, J. M., McGill, B. J., McGlinn, D. J., May, F., Blowes, S. A., Xiao, X., ... & Gotelli, N. J. (2018). Embracing scale-dependence to achieve a deeper understanding of biodiversity and its change across communities. *Ecology letters*, *21*(11), 1737-1751.

Colwell, R. K., & Coddington, J. A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *345*(1311), 101-118.

De Cáceres, M., Legendre, P., Valencia, R., Cao, M., Chang, L. W., Chuyong, G., ... & Kenfack, D. (2012). The variation of tree beta diversity across a global network of forest plots. *Global Ecology and Biogeography*, *21*(12), 1191-1202.

Ellison, A. M. (2010). Partitioning diversity 1. *Ecology*, *91*(7), 1962-1963.

Gaston, K. J. (2000). Global patterns in biodiversity. *Nature*, *405*(6783), 220.

Gentry, A. H. (1988). Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Annals of the Missouri botanical garden*, 1-34.

Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, *40*(3-4), 237-26

Gotelli, N. J., & Colwell, R. K. (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology letters*, *4*(4), 379-391.

Harrison, S., Davies, K. F., Safford, H. D., & Viers, J. H. (2006). Beta diversity and the scale-dependence of the productivity-diversity relationship: a test in the Californian serpentine flora. *Journal of Ecology*, *94*(1), 110-117.

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods in Ecology and Evolution*, *7*(12), 1451-1456.

Hurlbert, S. H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, *52*(4), 577-586.

Jost, L. (2007). Partitioning diversity into independent alpha and beta components. Ecology, 88(10), 2427–2439.

Kraft, N. J. B., Sanders, N. J., Stegen, J. C., Anderson, M. J., Crist, T. O., Cornell, H. V, … Swenson, N. G. (2012). Response to Comments on "Disentangling the Drivers of Diversity Along Latitudinal and Elevational Gradients." Science, 335(6076), 1573–1573.

Kraft, N. J., Comita, L. S., Chase, J. M., Sanders, N. J., Swenson, N. G., Crist, T. O., ... & Cornell, H. V. (2011). Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science, 333*(6050), 1755-1758.

Legendre, P., Borcard, D., & Peres-Neto, P. R. (2005). Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological monographs*, *75*(4), 435-450.

Maitner, B. S., Boyle, B., Casler, N., Condit, R., Donoghue, J., Durán, S. M., ... & McGill, B. (2018). The bien r package: A tool to access the Botanical Information and Ecology Network (BIEN) database. Methods in Ecology and Evolution, 9(2), 373-379.

May, F., Gerstner, K., McGlinn, D. J., Xiao, X., & Chase, J. M. (2018). mobsim: An r package for the simulation and measurement of biodiversity across spatial scales. *Methods in Ecology and Evolution*, *9*(6), 1401-1408.

McGill, B. J. (2010 b). Matters of scale. *Science*, *328*(5978), 575-576.

McGlinn, D. J., Xiao, X., May, F., Gotelli, N. J., Engel, T., Blowes, S. A., ... & McGill, B. J. (2019). Measurement of Biodiversity (MoB): A method to separate the scale-dependent effects of species abundance distribution, density, and aggregation on diversity change. *Methods in Ecology and Evolution*, *10*(2), 258-269.

Mori, A. S., Fujii, S., Kitagawa, R., & Koide, D. (2015). Null model approaches to evaluating the relative role of different assembly processes in shaping ecological communities. *Oecologia*, *178*(1), 261-273.

Mori, A. S., Isbell, F., & Seidl, R. (2018). β-diversity, community assembly, and ecosystem functioning. *Trends in ecology & evolution*, *33*(7), 549-564.

Myers, J. A., Chase, J. M., Jiménez, I., Jørgensen, P. M., Araujo-Murakami, A., Paniagua-Zambrana, N., & Seidel, R. (2013). Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. *Ecology letters*, *16*(2), 151-157.

Phillips O, Miller JS (2002) Global Patterns of Plant Diversity: Alwyn H. Gentry's Forest Transect Data Set. St. Louis, MO: Missouri Botanical Garden Press

Qian, H., & Song, J. S. (2012). Latitudinal gradients of associations between beta and gamma diversity of trees in forest communities in the New World. *Journal of Plant Ecology*, *6*(1), 12-18.

Qian, H., Chen, S., Mao, L., & Ouyang, Z. (2013). Drivers of β-diversity along latitudinal gradients revisited. Global Ecology and Biogeography, 22(6), 659–670.

Qian, H., Wang, X., & Zhang, Y. (2012). Comment on "Disentangling the Drivers of Diversity Along Latitudinal and Elevational Gradients." Science, 335(6076), 1573–1573.

Roden, V. J., Kocsis, Á. T., Zuschin, M., & Kiessling, W. (2018). Reliable estimates of beta diversity with incomplete sampling. *Ecology*, *99*(5), 1051–1062.

Rosenzweig, M. L. (1995). *Species diversity in space and time*. Cambridge University Press.

Socolar, J. B., Gilroy, J. J., Kunin, W. E., & Edwards, D. P. (2016). How should beta-diversity inform biodiversity conservation?. *Trends in ecology & evolution*, *31*(1), 67-80.

Sreekar, R., Katabuchi, M., Nakamura, A., Corlett, R. T., Slik, J. F., Fletcher, C., ... & Sun, I. F. (2018). Spatial scale changes the relationship between beta diversity, species richness and latitude. *Royal Society open science*, *5*(9), 181168.

Stegen, J. C., Freestone, A. L., Crist, T. O., Anderson, M. J., Chase, J. M., Comita, L. S., ... & Inouye, B. D. (2013). Stochastic and deterministic drivers of spatial and temporal turnover in breeding bird communities. *Global Ecology and Biogeography*, *22*(2), 202-212.

Tucker, C. M., Shoemaker, L. G., Davies, K. F., Nemergut, D. R., & Melbourne, B. A. (2016). Differentiating between niche and neutral assembly in metacommunities using null models of β-diversity. Oikos, 125(6), 778-789.

Tuomisto, H. (2010 a). A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, *33*(1), 2-22.

Tuomisto, H. (2010 b). A diversity of beta diversities: straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography*, *33*(1), 23-45.

Tuomisto, H., & Ruokolainen, K. (2012). Comment on "Disentangling the Drivers of Diversity Along Latitudinal and Elevational Gradients." Science, 335(6076), 1573–1573.

Whittaker, R. H. (1960). Vegetation of the Siskiyou mountains, Oregon and California. *Ecological monographs*, *30*(3), 279-338.

Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon*, *21*(2-3), 213-251.

Williamson, M., Gaston, K.J. and Lonsdale, W.M., 2001. The species–area relationship does not have an asymptote!. *Journal of Biogeography*, *28*(7), pp.827-830.

Wolda, H. (1981). Similarity indices, sample size and diversity. *Oecologia*, *50*(3), 296-302.

Worm, B. and Tittensor, D.P., 2018. A Theory of Global Biodiversity (MPB-60) (Vol. 79). Princeton University Press.

Xing, D., & He, F. (2019). Environmental filtering explains a U-shape latitudinal pattern in regional β-deviation for eastern North American trees. *Ecology letters*, *22*(2), 284-291.

Xu, W., Chen, G., Liu, C., & Ma, K. (2015). Latitudinal differences in species abundance distributions, rather than spatial aggregation, explain beta-diversity along latitudinal gradients. *Global Ecology and Biogeography*, *24*(10), 1170-1180.