ECOLOGICAL
SOUNDINGS

The tangled link between *β*- and *γ*-diversity: a Narcissus effect weakens statistical inferences in null model analyses of diversity patterns

Werner Ulrich[1]*, Andres Baselga[2], Buntarou Kusumoto[3], Takayuki Shiono[3], Hanna Tuomisto[4] and Yasuhiro Kubota[3]

[1]*Chair of Ecology and Biogeography, Nicolaus Copernicus University Toruń, Toruń, Poland,* [2]*Facultad de Biología, Universidad de Santiago de Compostela, Santiago de Compostela, Spain,* [3]*Faculty of Science, University of the Ryukyus, Ryukyus, Japan,* [4]*Department of Biology, University of Turku, Turku, Finland*

*Correspondence: Werner Ulrich, Chair of Ecology and Biogeography, Nicolaus Copernicus University, Lwowska 1, 87–100 Toruń, Poland.
E-mail: ulrichw@umk.pl

Abstract

Understanding the structure of and spatial variability in the species composition of ecological communities is at the heart of biogeography. In particular, there has been recent controversy about possible latitudinal trends in compositional heterogeneity across localities (β-diversity). A gradient in the size of the regional species pool alone can be expected to impose a parallel gradient on β-diversity, but whether β-diversity also varies independently of the size of the species pool remains unclear. A recently suggested methodological approach to correct latitudinal β-diversity gradients for the species pool effect is based on randomization null models that remove the effects of gradients in α- and γ-diversity on β-diversity. However, the randomization process imposes constraints on the variability of α-diversity, which in turn force γ- and β-diversity to become interdependent, such that any change in one is mirrored in the other. We argue that simple null model approaches are inadequate to discern whether correlations between α-, β- and γ-diversity reflect processes of ecological interest or merely differences in the size of the species pool among localities. We demonstrate that this kind of Narcissus effect may also apply to other metrics of spatial or phylogenetic species distribution. We highlight that Narcissus effects may lead to artificially high rejection rates for the focal pattern (Type II errors) and caution that these errors have not received sufficient attention in the ecological literature.

Keywords

Diversity, diversity partitioning, latitudinal gradient, null model, species pool, statistical inference.

Understanding which factors influence the assembly of ecological communities and their spatial variability has been at the centre of biogeographical research, especially since Whittaker (1965) introduced the concept of β-diversity and Diamond (1975) his assembly rules. A major question is to assess whether an observed distribution of species occurrences and absences among sites can be traced back to underlying ecological processes or whether it might stem from chance processes alone. Statistical standards have been proposed that compare the observed distribution with null models, accounting particularly for differences in species abundances and sample size (Gotelli & Ulrich, 2012; Bennett

& Gilbert, 2016). However, there is a long-standing discussion about how to construct these null models. The challenge is that the observed pattern of interest (e.g. species co-occurrences) should be randomized while other relevant aspects of the dataset (e.g. species richness and the total number of species occurrences) should be preserved and thereby serve as boundary conditions, i.e. define the scope of interpretation of the results. Several studies have assessed the probability that a pattern is identified where none exists (Type I errors) for important combinations of co-occurrence and β-diversity metrics and null models (reviewed in Ulrich & Gotelli, 2013). However, when null models are constrained, the

boundary conditions themselves might smuggle in unforeseen constraints on the null distribution of the pattern of interest, i.e. produce a Narcissus effect (Colwell & Winkler, 1984). This could increase the similarity between the observed and null-modelled species distributions and lead to artificially high rejection rates for the focal pattern (Type II errors). So far, Type II error probabilities have not received the necessary attention in the ecological literature. Here we address this problem and show that a Narcissus effect is inevitable in null model analyses of some kinds of metrics, such as those related to β-diversity.

Most animal and plant taxa show a gradient of increasing species diversity towards the lower latitudes. As a consequence, the species richness per sampling unit (α-diversity) in datasets collected at different latitudes generally increases with increasing total species richness observed when all sampling units in the dataset are pooled (γ-diversity). However, the latitudinal gradient in α-diversity is often less steep than the corresponding gradient in γ-diversity (Hillebrand, 2004; Brown, 2014) making β-diversity (= γ/α), which reflects the degree of compositional variability among the sampling units, increase towards the equator as well. There has been some controversy about the interpretation of latitudinal trends in β, because this might or might not be indicative of real changes in community assembly processes across latitudes (Kraft *et al.*, 2011; Qian & Xiao, 2012; Qian *et al.*, 2012, 2013; Tuomisto & Ruokolainen, 2012; Barton *et al.*, 2013; Brown, 2014; Fraser *et al.*, 2014; Xu *et al.*, 2015). Kraft *et al.* (2011) have argued that the observed latitudinal trend in β-diversity is merely a statistical artefact that arises from the definition of β as the ratio between γ and α, and that the trend disappears if the data are properly corrected for variation in γ. Others have questioned the adequacy of the correction methods that have been applied (Qian *et al.*, 2012, 2013; Tuomisto & Ruokolainen, 2012) and have even claimed that the correction removed a real trend, which remains detectable if an improved correction method is applied (Qian *et al.* 2013).

Raw scores of β-diversity metrics are constrained by the underlying regional distribution of species abundances that influence local colonization probabilities and therefore α-diversity (Kraft *et al.*, 2011). To control for this, and to remove the effect of γ-diversity on β-diversity, many authors (e.g. Qian *et al.*, 2013; Stegen *et al.*, 2013; Xu *et al.*, 2015) have used effect sizes derived from null models that resampled local communities from a hypothesized regional abundance distribution. However, null model comparisons might be less powerful for this purpose than has been assumed. We exemplify this using proportional species turnover $\beta_P = 1 - (\alpha/\gamma) = 1 - (1/\beta_M)$, where β_M is the classical multiplicative β-diversity (= γ/α). β_P has been a popular measure in this kind of study and quantifies the proportion of observed species that are not contained in a local community of average richness α (Tuomisto, 2010).

Null models based on a randomization of species occurrences in a species × sites matrix with γ species, n sites and k occurrences constrain either γ or k (or both) to remain consistent with observed patterns. As the quotient k/n is equivalent to the average number of species occurrences per site, it also equals α (Rodríguez & Arita, 2004). Substituting k/n for α in the equation for β_P shows that β_P is the complement of matrix fill f_{obs}, which in turn is the proportion of occupied cells in a matrix, k, divided by matrix size, γn (Routledge, 1977):

$$\beta_P = 1 - \frac{k}{\gamma n} = 1 - f_{obs}. \tag{1}$$

A null model might now fix the number of occurrences k and allow for variability in both n and γ. As variability in n implies either the use of subsamples or the creation of artificial sites of unknown composition, most null models have been based on variability of total species richness only. Using a variable γ_{null} can be justified by the fact that the observed γ is only one out of many possible samples of the regional species pool γ_{reg} in a focal set of sites. Thus, the simulated random variable γ_{null} should be constrained between α and γ_{reg}.

The statistical behaviour of constrained variables has been studied in detail in connection with theories of life-history invariants and metabolic theory (Cipriani & Collin, 2005; Nee *et al.*, 2005). That work has demonstrated that two random variables, one of which is nested within a second, inevitably become correlated. Modifying Nee *et al.* (2005), we can envision that γ_{null} (if nested within γ_{reg} and larger than α) is given by $\gamma_{null} = \alpha + p \times \gamma$ with $0 < p < (\gamma_{reg} - \alpha)/\gamma$. With increasing numbers of random matrices calculated for statistical inference, the random variable p will asymptotically reach its central tendency $E[p]$. With $\alpha = k/n$ and $f_{obs} = k/(n\gamma)$ the expected effect size $\Delta\beta_P = \beta_P(obs) - \beta_P(null)$ is estimated by

$$
\begin{aligned}
\Delta\beta_P &= 1 - f_{obs} - (1 - f_{null}) \\
&= 1 - f_{obs} - \frac{E[p]\gamma}{(k/n) + E[p]\gamma} \\
&= 1 - f_{obs} - \frac{E[p]}{f_{obs} + E[p]} \\
&= \frac{f_{obs}}{f_{obs} + E[p]} - f_{obs}.
\end{aligned}
\tag{2}
$$

The first line of equation (2) shows that any null model that fixes matrix fill does not work in combination with β_P as the resulting effect sizes are always zero. The last line of equation (2) shows that fixing k and allowing γ_{null} to vary does not help to counteract the effect of species pool size either, because the effect size $\Delta\beta_P$ is independent of species richness. It is instead fully determined by the fill of the observed matrix and the null model parameter $E[p]$ that specifies the method of resampling. When the same resampling method is used to compare matrices, $\Delta\beta_P$ therefore becomes a simple function of the observed fill and, hence, of the observed ratio α/γ (= 1 − β_P). Such a null model is not
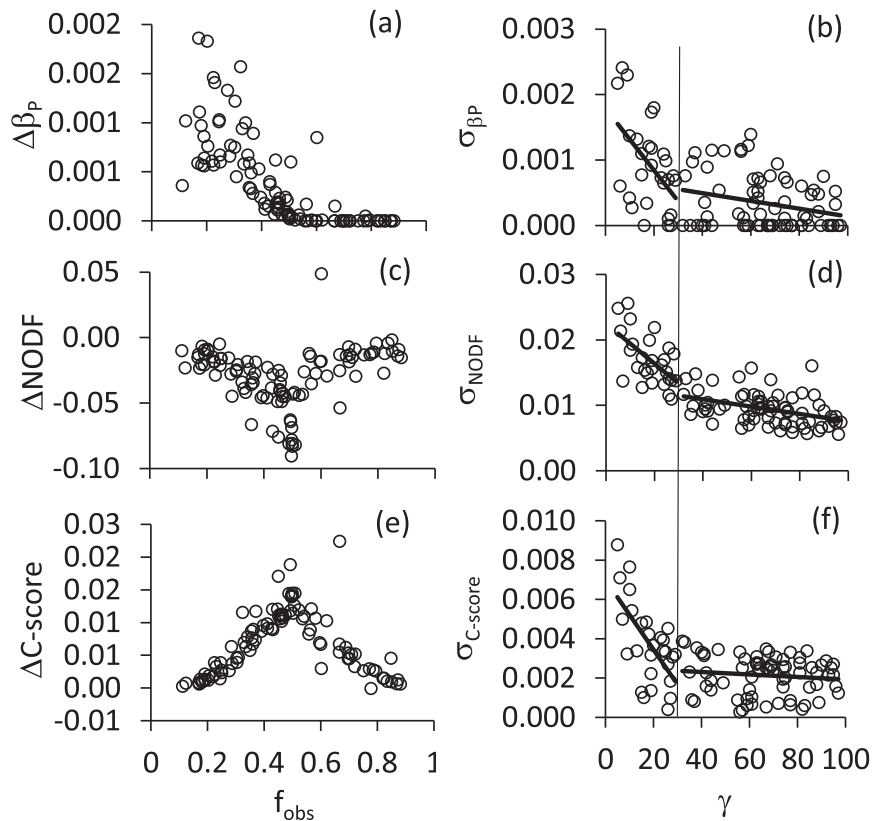
Figure 1 Simulation study based on 100 artificial matrices of very different structure (from highly nested to highly segregated) with expected γ-diversity (γ_{null}) = 5–100 species, n = 30 sites and observed fill (f_{obs}) = 0.1–0.9. Datasets were generated as in Ulrich & Gotelli (2013). Effect sizes (a, $\Delta\beta_P$; c, ΔNODF; e, ΔC-score) vary with matrix fill. Null model standard deviations σ (b, d, f) decline with species richness. The null model resampled species occurrences proportional to observed row and column marginal totals until the observed species richness was reached (the proportional null model algorithm of Ulrich & Gotelli, 2012). This algorithm allows matrix fill to vary. The vertical line in b, d, and f indicates that the number of species γ exceeds the number of sites n. Above this threshold σ becomes largely independent of γ_{null}. Linear ordinary least squares regressions for γ < 30 are (b) r^2 = 0.26, (d) r^2 = 0.34, (f) r^2 = 0.41, and for γ > 30 they are (b) r^2 = 0.07, (d) r^2 = 0.16, (f) r^2 = 0.01.

able to account successfully for variation in total species richness in the dataset. In particular, it is not able to counteract the problem that the degree of undersampling increases with species pool size, which was an important determinant of β_P in the study by Kraft *et al.* (2011) (Tuomisto & Ruokolainen, 2012).

The above results imply that a meaningful null model for testing patterns in β-diversity and related measures needs to allow for variability in matrix fill. In agreement with this, various authors have used null models that fix γ to obtain expected values of β_P while allowing occurrences (and hence matrix fill) to vary according to individual-based samples from either observed (Qian *et al.*, 2013) or assumed (Kraft *et al.*, 2011; Xu *et al.*, 2015) species abundance distributions.

However, even these null models suffer from internal constraints. Assume a non-degenerate matrix where each site is occupied by at least one species. Matrix fill is then constrained between $f_{min} = \max(1/n; 1/\gamma)$ and $f_{max} = 1$. Thus $f_{null} = f_{min} + (1 - f_{min})p$; with $0 < p < 1$. Using the same

argument as above (equation 2) the expected effect size $\Delta\beta_P$ equals

$$\begin{aligned}\Delta\beta_P &= 1 - f_{obs} - (1 - f_{null}) \\ &= 1 - f_{obs} + (E[p] - 1)(1 - f_{min}).\end{aligned} \tag{3}$$

For sufficiently large matrices f_{min} is small and equation (3) becomes

$$\Delta\beta_P \approx E[p] - f_{obs}. \tag{4}$$

Again, $\Delta\beta_P$ turns out to be a simple decreasing function of observed matrix fill independent of the null model, as shown in Fig. 1(a) for a null model that resampled the matrix proportional to observed totals of species and site occurrences (Ulrich & Gotelli, 2012).

The random variable p that defines matrix fill might be a function of species richness. This is the case in null models that resample the matrix according to observed incidence or

abundance distributions until the observed richness is reached, such as those used by Kraft *et al.* (2011). Consequently, a link between *p* and γ makes Δβ$_P$ dependent on γ irrespective of whether the null model invokes global (Kraft *et al.*, 2011) or local (Qian *et al.*, 2013) marginal totals or abundance distributions for randomization.

Statistical inference in null model analyses relies on the variance σ^2 of the null distribution using either confidence limits or standardized effect sizes (SES = Δβ$_P$/σ). The variance of the null model of equation (3) equals

$$\sigma^2 = \sigma^2[(p-1)(1-f_{min})] = (1-f_{min})\sigma^2(p) \qquad (5)$$

and is thus proportional to $1 - f_{min}$. Therefore, if *p* is independent of matrix properties the variance of the null distribution should be independent of matrix fill and of species richness if γ > *n* (marked by the vertical line in Fig. 1). Together, equations (4) imply that means and variances of the null distributions are constrained by observed matrix fill and size and are less variable than required for unbiased statistical testing.

On the basis of the above, we argue that previous tests for gradients in β-diversity using a null model approach to account for differences in total species richness did not effectively remove the relationships that differences in species pool size impose between α, β and γ. Qian *et al.* (2013) have already highlighted that a null model designed to test for variation in β-diversity should not smuggle in mechanisms that generate the pattern under study. Given that any of the variables α, β and γ can be calculated if the other two are known, any null model that constrains one of the three terms is prone to induce just this version of the Narcissus effect (Colwell & Winkler, 1984).

An additional concern is that such Narcissus effects are not limited to studies related to β-diversity, but may apply to any null model analysis of metrics that are constrained by matrix fill and size. These include metrics designed to quantify the spatial or temporal geometry of species distributions, such as the widely used NODF metric of the degree of nestedness (Almeida-Neto *et al.*, 2008), the *C*-score of pairwise species co-occurrences (Stone & Roberts, 1990; Ulrich & Gotelli, 2013), the Sørensen and Jaccard indices of compositional similarity (Arita, 2016) and measures of phylogenetic relatedness (Tucker *et al.*, 2016).

The Sørensen and Jaccard indices are monotonic transformations of β (e.g. Tuomisto, 2010) so they, too, are fully determined by matrix fill and size (Arita, 2016), and eqs. (2–5) and Fig. 1a, b apply to them as well. Both NODF ('nestedness measure based on overlap and decreasing fills') and the *C*-score are matrix size-normalized counts of pairwise patterns of co-occurrence (overlap and reciprocal exclusion, respectively). Any randomization that constrains these totals also constrains these metrics. As there is no simple analytical solution similar to equations 2 and 4, we used simulations to show that effect sizes and variances of the null distributions are still correlated with matrix fill and

richness, respectively (Fig. 1), which can bias statistical inference.

In particular, we notice that the standard deviation of the null distributions remains approximately constant and independent of matrix size for γ > *n* (Fig. 1b, d, f) as predicted by equation (5). Indeed, several studies have already reported that effect sizes of various co-occurrence metrics show a dependence on species richness (the temperature metric of nestedness, Ulrich & Gotelli, 2007a; metrics of species turnover, Ulrich & Gotelli, 2007b; metrics of phylogenetic evenness, Helmus *et al.*, 2007). Originally, these findings were interpreted in terms of metric and/or null model deficiency, but we argue that they are a logical consequence of how a metric is normalized. If this normalization invokes marginal totals, the Narcissus effect inevitably lurks around the corner.

In conclusion, in our opinion it is not possible to remove the effects of the regional species pool size on β and patterns of species co-occurrences by only focusing on matrix geometry in a single dataset. This is because geometry, γ and β are properties of the same dataset and therefore reflect the same external processes. The β-diversity quantifies the heterogeneity in the distributions of the observed γ species among the inventoried *n* local assemblages. To determine if this heterogeneity is merely noise caused by incomplete random sampling from the regional species pool, or reflects true biological differentiation of assemblages in response to a heterogeneous environment or other ecological factors, it is necessary to have external data on the sizes of the relevant species pools and on the variation in the relevant environmental or ecological factors among the sites. Null model analysis alone is not an adequate solution to this problem.

ACKNOWLEDGEMENTS

REFERENCES

Almeida-Neto, M., Guimarães, P., Guimarães, P.R., Jr, Loyola, R.D. & Ulrich, W. (2008) A consistent metric for nestedness analysis in ecological systems: reconciling concept and quantification. *Oikos*, **117**, 1227–1239.

Arita, H.T. (2016) Multisite and multispecies measures of overlap, co-occurrence, and co-diversity. *Ecography*, doi: 10.1111/ecog.01942.

Barton, P.S., Cunningham, S.A., Manning, A.D., Gibb, H., Linsenmayer, D.B. & Didham, R.K. (2013) The spatial scaling of beta diversity. *Global Ecology and Biogeography*, **22**, 639–647.

Bennett, J.R. & Gilbert, B. (2016) Contrasting beta diversity among regions: how do classical and multivariate approaches compare? *Global Ecology and Biogeography*, **25**, 368–377.

Brown, J.H. (2014) Why are there so many species in the tropics? *Journal of Biogeography*, **41**, 8–22.

Cipriani, R. & Collin, R. (2005) Life-history invariants with bounded variables cannot be distinguish from data generated by random processes using standard analyses. *Journal of Evolutionary Biology*, **18**, 1613–1618.

Colwell, R.K. & Winkler, D.W. (1984) A null model for null models in biogeography. *Ecological communities: conceptual issues and the evidence* (ed. by D.R. Strong Jr, D. Simberloff, L.G. Abele and A.B. Thistle), pp. 344–359. Princeton University Press, Princeton, NJ.

Diamond, J.M. (1975) Assembly of species communities. *Ecology and evolution of communities* (ed. by M.L. Cody and J.M. Diamond), pp. 342–444. Harvard University Press, Cambridge, MA, USA.

Fraser, D., Hassall, C., Gorelick, R. & Rybczynski, N. (2014) Mean annual precipitation explains spatiotemporal patterns of Cenozoic mammal beta diversity and latitudinal diversity gradients in North America. *PLoS One*, **9**, e106499.

Gotelli, N.J. & Ulrich, W. (2012) Statistical challenges in null model analysis. *Oikos*, **121**, 171–180.

Helmus, M.R., Bland, T.J., Williams, C.K. & Ives, A.R. (2007) Phylogenetic measures of biodiversity. *The American Naturalist*, **169**, E68–E83.

Hillebrand, H. (2004) On the generality of the latitudinal diversity gradient. *The American Naturalist*, **163**, 192–211.

Kraft, N.J.B., Comita, L.S., Chase, J.M., Sanders, N.J., Swenson, N.G., Crist, T.O., Stegen, J.C., Vellend, M., Boyle, B., Anderson, M.J., Cornell, H.V., Davies, K.F., Freestone, A.L., Inouye, B.D., Harrison, S.P. & Myers, J.A. (2011) Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science*, **333**, 1755–1758.

Nee, S., Colegrave, N., West, S.A. & Grafen, A. (2005) The illusion of invariant quantities in life history. *Science*, **309**, 1236–1239.

Qian, H. & Xiao, M. (2012) Global patterns of the beta diversity–energy relationship in terrestrial vertebrates. *Acta Oecologica*, **39**, 67–71.

Qian, H., Wang, X. & Zhang, Y. (2012) Comment on 'Disentangling the drivers of β diversity along latitudinal and elevational gradients'. *Science*, **30**, 1573.

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

Rodríguez, P. & Arita, H.T. (2004) Beta diversity and latitude in North American mammals: testing the hypothesis of covariation. *Ecography*, **27**, 547–556.

Routledge, R.D. (1977) On Whittaker's components of diversity. *Ecology*, **58**, 1120–1127.

Stegen, J.C., Freestone, A.L., Crist, T.O., Anderson, M.J., Chase, J.M., Comita, L.S., Cornell, H.V., Davies, K.F., Harrison, S.P., Hurlbert, A.H., Inouye, B.D., Kraft, N.J.B., Myers, J.A., Sanders, N.J., Swenson, N.G. & Vellend, M. (2013) Stochastic and deterministic drivers of spatial and temporal turnover in breeding bird communities. *Global Ecology and Biogeography*, **22**, 202–212.

Stone, L. & Roberts, A. (1990) The checkerboard score and species distributions. *Oecologia*, **85**, 74–79.

Tucker, C.M., Cadotte, M.W., Carvalho, S.B., Davies, T.J., Ferrier, S., Fritz, S.A., Grenyer, R., Helmus, M.R., Jin, L.S., Mooers, A.O., Pavoine, S., Purschke, O., Redding, D.W., Rosauer, D.F., Winter, M. & Mazel, F. (2016) A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews*, doi: 10.1111/brv.12252.

Tuomisto, H. (2010) 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**, 2–22.

Tuomisto, H. & Ruokolainen, K. (2012) Comment on 'Disentangling the drivers of β diversity along latitudinal and elevational gradients'. *Science*, **30**, 1573.

Ulrich, W. & Gotelli, N.J. (2007a) Null model analysis of species nestedness patterns. *Ecology*, **88**, 1824–1831.

Ulrich, W. & Gotelli, N.J. (2007b) Disentangling community patterns of nestedness and species co-occurrence. *Oikos*, **116**, 2053–2061.

Ulrich, W. & Gotelli, N.J. (2012) A null model algorithm for presence–absence matrices based on proportional resampling. *Ecological Modelling*, **244**, 20–27.

Ulrich, W. & Gotelli, N.J. (2013) Pattern detection in null model analysis. *Oikos*, **122**, 2–18.

Whittaker, R.H. (1965) Dominance and diversity in land plant communities. *Science*, **147**, 250–260.

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**, 1170–1180.

BIOSKETCH

The authors are particularly interested in the assembly of ecological communities at different spatial and temporal scales and in the variability of species richness at a global scale. They try to disentangle the interplay between patterns of species associations in plant and animal communities, environmental (and human) impact, and evolutionary constraints.

Editor: Pedro Peres-Neto