

Beta Diversity

SESYNC
2 February 2017

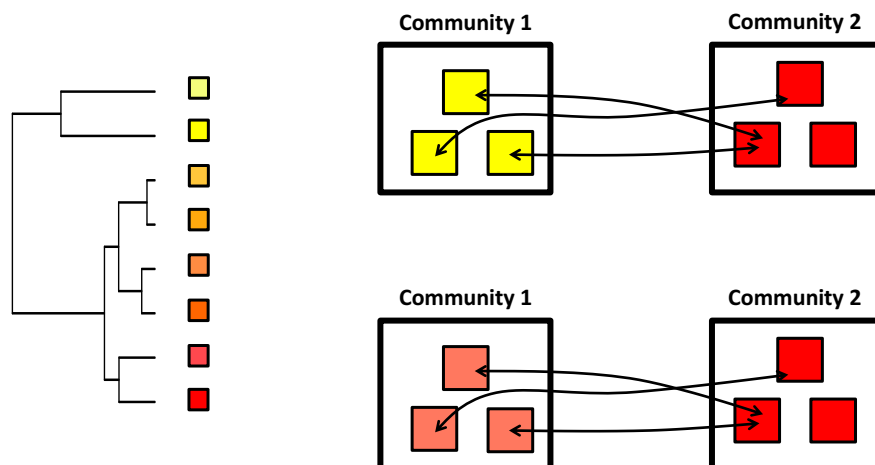
What is *Phylogenetic* beta diversity?

- The phylogenetic dissimilarity between two or more communities
- As with species beta diversity, I think of phylogenetic beta diversity as a general concept that can be measured in a couple of different ways

Why *Phylogenetic* beta diversity?

- Species names hold little information about the ecology of species.
- Phylogenies hold more information and may help provide stronger inferences regarding the role of random and non-random processes governing plant community structure

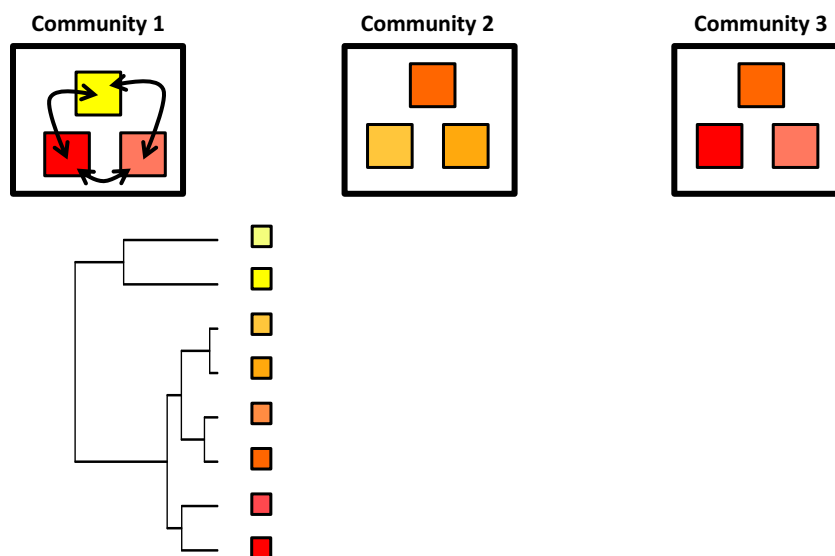
Why Phylogenetic Beta Diversity?



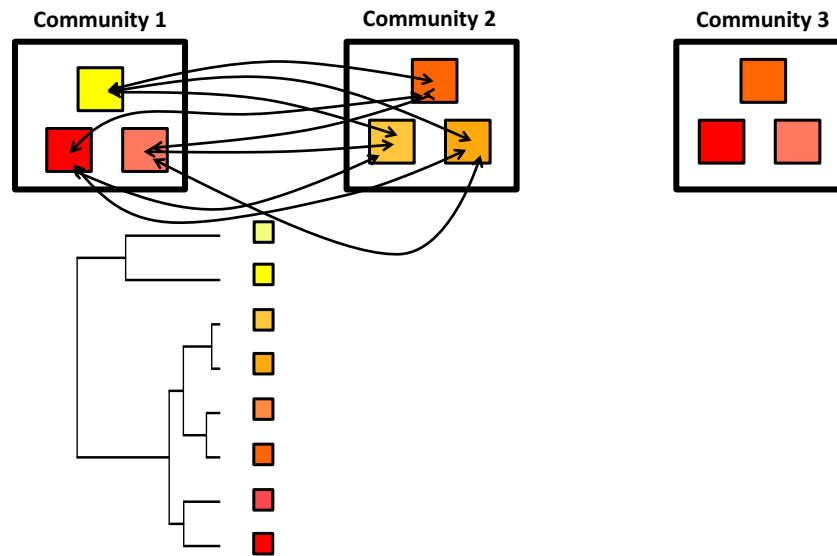
How do we calculate *Phylogenetic* beta diversity?

- Many methods (>10?) exist currently and some are very similar or nearly identical
 - I consider there to be two general classes of phylogenetic beta diversity metrics (Swenson 2011 PLoS One)
 - Pairwise
 - Nearest Neighbor

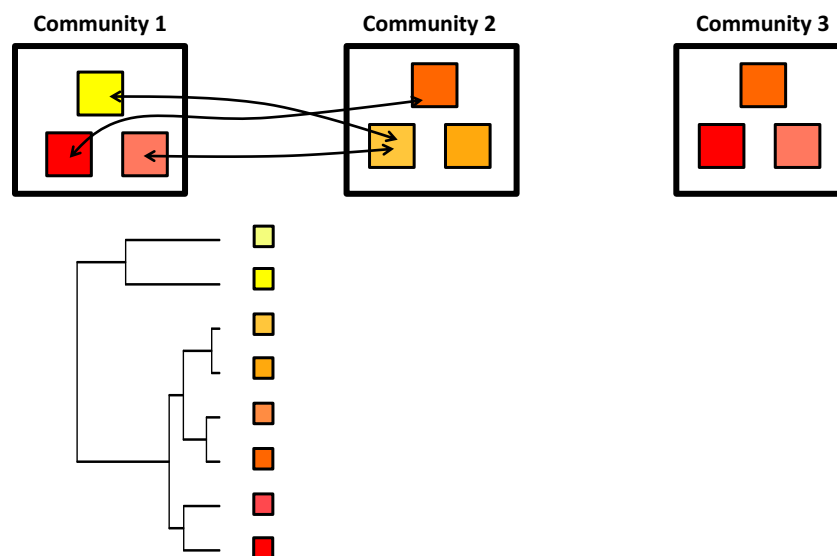
Phylogenetic Alpha Diversity Along Environmental Gradient



Phylogenetic Beta Diversity Along Environmental Gradient (Pairwise)



Phylogenetic Beta Diversity Along Environmental Gradient (Nearest Neighbor)



The History of Phylogentic Beta Diversity

- Phylogenetic Beta Diversity papers have been published as far back as 2002 --- 1982 if we consider Rao.
- This work is generally not acknowledged or known because it primarily was done by microbial ecologists
- Microbial ecologists need to build phylogenetic trees for their work and therefore all of their diversity metrics are phylogenetic.

The History of Phylogentic Beta Diversity

- Phylogenetic Beta Diversity was 'discovered' by plant ecologists around 2008
 - Bryant et al. 2008 PNAS
 - Graham and Fine 2008 Ecology Letters

A Comparison of Metrics

Phylogenetic Beta Diversity Metrics, Trait Evolution and Inferring the Functional Beta Diversity of Communities

Nathan G. Swenson*

Department of Plant Biology, Michigan State University, East Lansing, Michigan, United States of America

Abstract

The beta diversity of communities along gradients has fascinated ecologists for decades. Traditionally focused on the species composition of communities, but researchers are becoming increasingly interested in phylogenetic composition in the hope of achieving mechanistic insights into community structure. Phylogenetic beta diversity has been published, but few empirical studies have been published, from such phylogenetic studies critically rely on the pattern of trait evolution. The present work phylogenetic dissimilarity of 96 tree communities in India. The work compares and contrasts eight dissimilarity, considers the role of phylogenetic signal in trait data and shows that environmental spatial distance is the best correlate of phylogenetic dissimilarity in the study system.

Citation: Swenson NG (2011) Phylogenetic Beta Diversity Metrics, Trait Evolution and Inferring the Functional Beta Diversity. *PLoS ONE* 6(2): e21264. doi:10.1371/journal.pone.0021264

Editor: Andrew Hector, University of Zurich, Switzerland

Received: January 18, 2011; **Accepted:** May 27, 2011; **Published:** June 24, 2011

Copyrights: © 2011 Nathan G. Swenson. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by Michigan State University. The funders had no role in the study design, data collection and preparation of the manuscript.

Competing Interests: The author has declared that no competing interests exist.

* E-mail: swenson@msu.edu

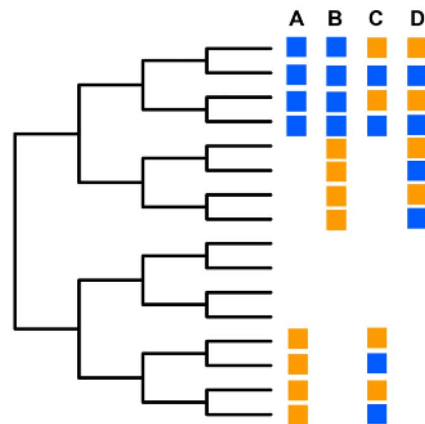


Figure 1. An example of four pairs of hypothetical communities and types of phylogenetic beta diversity. The species in a

A Comparison of Metrics

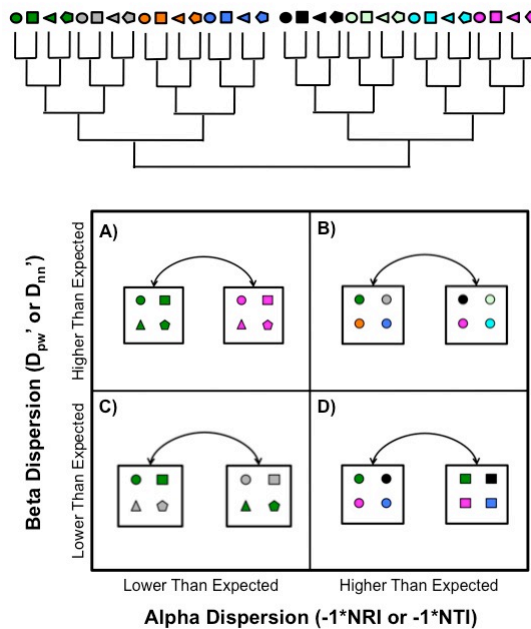
Table 3. A correlation analysis of different metrics of species and phylogenetic community dissimilarity.

	Jaccard	Bray-Curtis	PhyloSor	UniFrac	Dnn	Dnn'	Dpw	Dpw'	Rao's D	Rao's H
Jaccard	-	-	-0.59--0.55	0.592--0.598	0.429--0.434	0.592--0.597	0.170--0.178	0.343--0.348	0.344--0.349	0.280--0.284
Bray-Curtis	0.988	-	-0.626--0.622	0.642--0.646	0.478--0.482	0.650--0.656	0.183--0.188	0.342--0.347	0.343--0.349	0.108--0.113
PhyloSor	-0.57	-0.624	-	-0.993--0.989	-0.905--0.902	-0.805--0.801	-0.116--0.112	-0.155--0.147	-0.155--0.150	-0.366--0.362
UniFrac	0.596	0.644	-0.991	-	0.869--0.874	0.775--0.779	0.106--0.109	0.149--0.153	0.149--0.153	0.342--0.347
Dnn	0.431	0.48	-0.904	0.872	-	0.765--0.770	0.063--0.061	0.074--0.078	0.073--0.077	0.465--0.472
Dnn'	0.594	0.653	-0.803	0.778	0.769	-	0.158--0.163	0.239--0.244	0.239--0.244	0.408--0.414
Dpw	0.174	0.186	-0.114	0.108	0.065	0.161	-	0.564--0.569	0.563--0.569	0.049--0.053
Dpw'	0.346	0.346	-0.152	0.151	0.076	0.242	0.566	-	1--1	0.099--0.112
Rao's D	0.346	0.346	-0.152	0.151	0.076	0.242	0.566	1	-	0.098--0.112
Rao's H	0.283	0.11	-0.364	0.344	0.47	0.411	0.051	0.11	0.11	-

The lower triangle cell values are Pearson's r values being calculated from the Phylomatic phylogeny. Values in the upper triangle are 95% confidence intervals of r values calculated from the 100 randomly resolved phylogenies.
doi:10.1371/journal.pone.0021264.t003

Scale, Alpha and Beta

- Swenson et al. 2012b Ecology
- BCI, Luquillo, Dinghushan, Gutianshan, Wisconsin, Maryland
- How are phylogenetic and functional alpha and beta related and how does that change with latitude?



Functional Beta Diversity

- There are few functional beta diversity papers published.
- The majority of these papers are on metric development and less focus has been placed on biological processes
- The metrics that have been developed are generally identical to those used in phylogenetic beta diversity studies.

Functional Beta Diversity – Main Metrics

- All can be weighted by abundance
- Pairwise metrics
 - Dpw, Rao's Quadratic Entropy
- Nearest neighbor metrics
 - Dnn

Functional Beta Diversity in Space – An Example

PROCEEDINGS
OF
THE ROYAL
SOCIETY **B**

Proc. R. Soc. B (2011) **278**, 877–884
doi:10.1098/rspb.2010.1369
Published online 22 September 2010

Deterministic tropical tree community turnover: evidence from patterns of functional beta diversity along an elevational gradient

Nathan G. Swenson^{1,*}, Pedro Anglada-Cordero² and John A. Barone³

¹*Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA*

²*Institute for Ecosystem Studies, University of Puerto Rico, Rio Piedras, Puerto Rico 00781*

³*Department of Biology, Columbus State University, Columbus, GA 31907, USA*

Functional Beta Diversity in Space – An Example

- Several leaf functional traits and wood density
- Measured at each elevation for each species
 - ‘population means’ not ‘species means’
- 15 plots from 300–1000m asl (0.1ha)

PROCEEDINGS
OF
THE ROYAL
SOCIETY **B**

Proc. R. Soc. B (2011) **278**, 877–884
doi:10.1098/rspb.2010.1369
Published online 22 September 2010

Deterministic tropical tree community turnover: evidence from patterns of functional beta diversity along an elevational gradient

Nathan G. Swenson^{1,*}, Pedro Anglada-Cordero² and John A. Barone³

¹*Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA*

²*Institute for Ecosystem Studies, University of Puerto Rico, Rio Piedras, Puerto Rico 00781*

³*Department of Biology, Columbus State University, Columbus, GA 31907, USA*

Functional Beta Diversity in Space – An Example



Functional Beta Diversity in Space – An Example

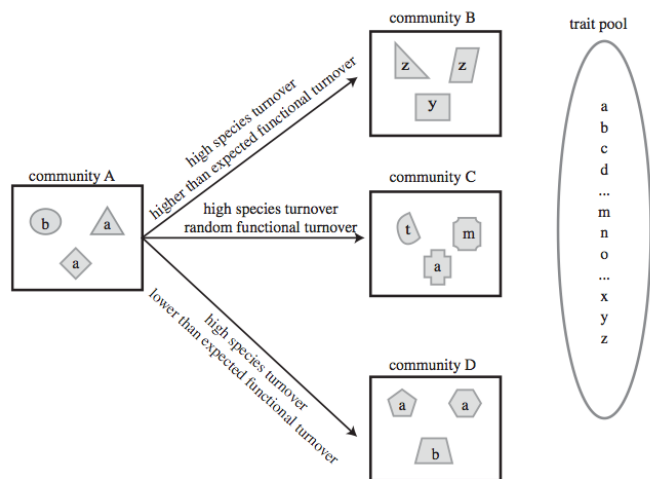


Figure 1. A cartoon depicting the species and functional turnover between a set of four hypothetical communities. The shapes

Functional Beta Diversity in Space – An Example

- 3 Measures of functional beta diversity
 - Functional Sorensen's
 - Nearest Neighbor (Dnn)
 - Pairwise (Dpw)
- Null model & SES

PROCEEDINGS
OF
THE ROYAL
SOCIETY

Deterministic tropical tree community turnover: evidence from patterns of functional beta diversity along an elevational gradient

Nathan G. Swenson^{1,*}, Pedro Anglada-Cordero² and John A. Barone³

¹Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA

²Instituto for Ecosystem Studies, University of Puerto Rico, Rio Piedras, Puerto Rico 00781

³Department of Biology, Columbus State University, Columbus, GA 31907, USA

Proc. R. Soc. B (2011) 278, 877–884
doi:10.1098/rspb.2010.1369
Published online 22 September 2010

$$F_{\text{SOR}} = \frac{BL_{k_1 k_2}}{(BL_{k_1} + BL_{k_2})1/2},$$

$$D_{\text{nn}} = \frac{\sum_{i=1}^{n_{k_1}} f_i \min \delta_{ik_2} + \sum_{j=1}^{n_{k_2}} f_j \min \delta_{jk_1}}{2},$$

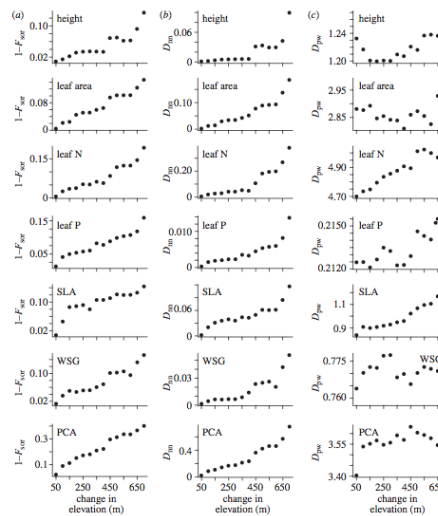
for species $i \neq$ species j

$$D_{\text{pw}} = \frac{\sum_{i=1}^{n_{k_1}} f_i \overline{\delta_{ik_2}} + \sum_{j=1}^{n_{k_2}} f_j \overline{\delta_{jk_1}}}{2}, \quad \text{for species } i \neq \text{species } j$$

$$\text{SES} = \frac{X_{\text{obs}} - \overline{X_{\text{null}}}}{\text{s.d.}(X_{\text{null}})},$$

Functional Beta Diversity in Space – An Example

- Functional distance decay
- Lowest elevation compared to all other elevations



Functional Beta Diversity in Space – An Example

- Higher than expected functional turnover with elevation.
- This can largely be explained by intra-specific trait variation at different elevations

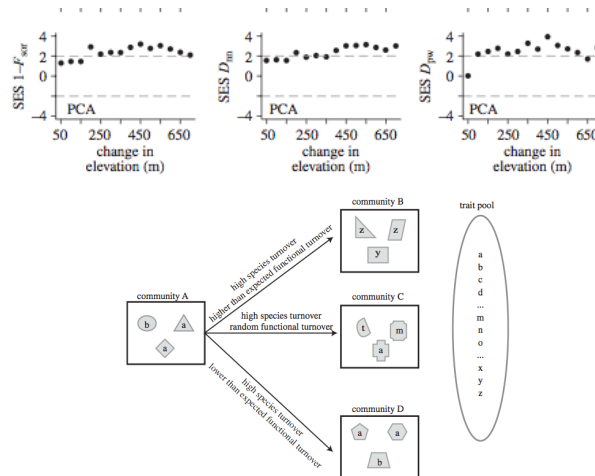


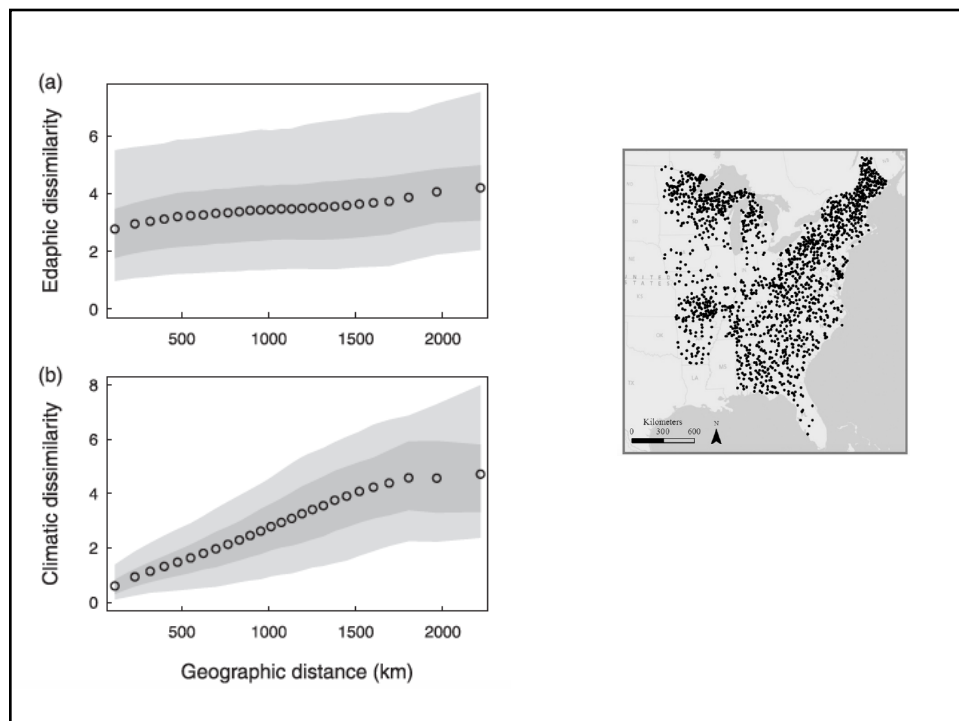
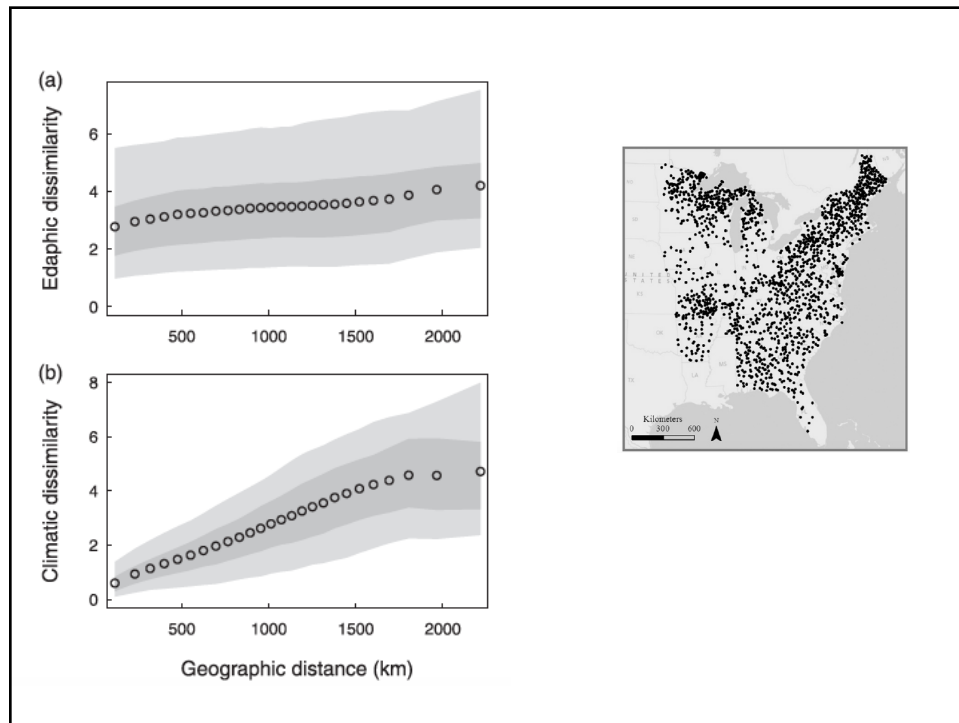
Figure 1. A cartoon depicting the species and functional turnover between a set of four hypothetical communities. The shapes

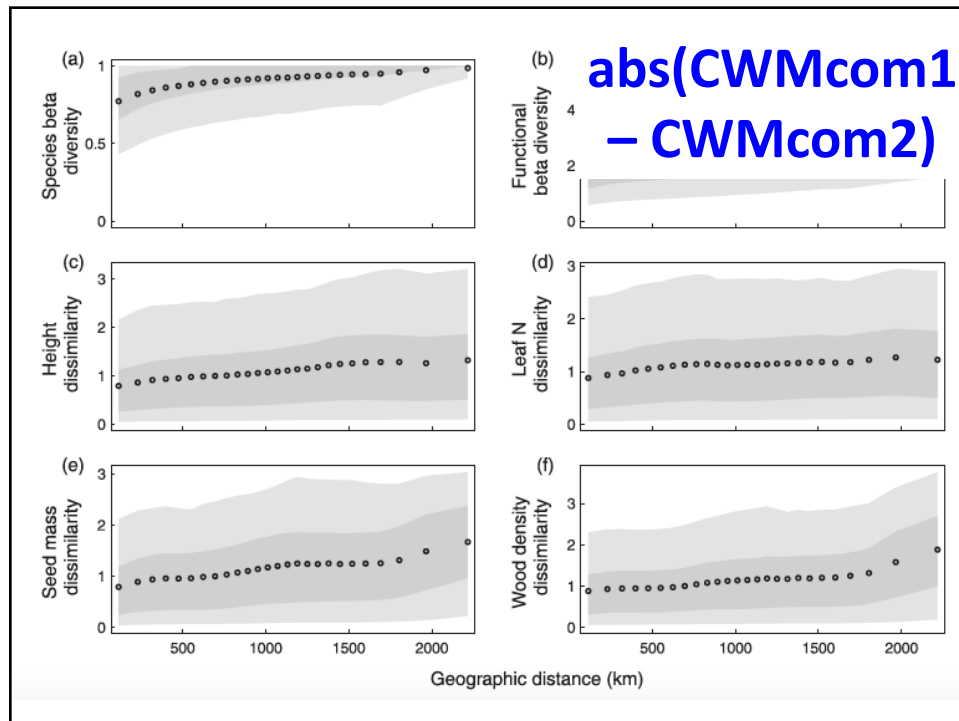
Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2013) **22**, 682–691



Functional beta-diversity patterns reveal deterministic community assembly processes in eastern North American trees

Andrew Siefert^{1*}, Catherine Ravenscroft¹, Michael D. Weiser² and Nathan G. Swenson³





“Major” Beta Diversity Measures

Their metric seeks to quantify the unique fraction of the phylogeny contained in each of the two communities or assemblages being compared, hence the name UniFrac. The formal equation for calculating UniFrac is:

$$\text{UniFrac}_{A,B} = \frac{PD_{A \cup B} - PD_{A \cap B}}{PD_{A \cup B}}$$

Where $PD_{A \cup B}$ is the phylogenetic diversity of the species in communities A and B combined and $PD_{A \cap B}$ is the phylogenetic diversity of the species shared between communities A and B.

Put in plain English, in order to calculate the UniFrac for two communities you must subtract the shared branch lengths between two communities from the sum of the total branch lengths for all species in both communities (i.e. the Faith's index for a 'community' containing all species found in your two communities). This difference is then divided by the sum of the total branch lengths in both communities. Thus, it is fairly easy to calculate in R. We start by first loading our example community data matrix for this chapter in to R.

It is often useful when calculating beta diversity to consider the relative abundances of the species in the communities being compared and to perhaps compare and contrast such measures to those that weight all species equally. The UniFrac metric is the only tree-based metric of which I am aware that has been extended to include the relative abundances of species into the calculation. Specifically, Luzopone et al. [116] extended their original unweighted UniFrac calculation to consider how many individuals from each community were subtended by the branches unique to each community. The calculation of the weighted UniFrac metric has two steps. The first step is to calculate a raw weighted UniFrac value (u) as follows:

$$u = \sum_i^n b_i \times \left| \frac{A_i}{A_T} - \frac{B_i}{B_T} \right|$$

Where b_i is the length of individual branch i of the n total branches in your phylogeny containing all species in your community data matrix or meta-community. Each individual branch (b_i) is multiplied by the absolute value of the difference between the abundance of the species subtended by branch i in community A (A_i) divided by the total abundance in community A (A_T) and the abundance of the species subtended by branch i in community B (B_i) divided by the total abundance in community B (B_T). Simplified, the length of each branch is multiplied by the absolute value of the differences in the relative abundance of the species in community A and B that are subtended by that branch summed over all branches. Given the calculation utilizes the relative abundances of species in each community we first transform our community data matrix to transform raw counts of species to relative abundances.

“Major” Beta Diversity Measures

The result is the raw weighted UniFrac value (u) for the first two communities in our community data matrix. This value is typically “normalized” by a scaling factor called D . In words, the scaling factor sums the relative abundance for each species in the two communities and multiplies this value by the branch length from the root of the phylogeny to the branch tip for that species and sums this calculation over all species.

$$D = \sum_j^n d_j \times \left(\frac{A_j}{A_T} + \frac{B_j}{B_T} \right)$$

Where d_j is the distance from the root to the branch tip for species j of n total species. The value for d_j will be constant for all species when using an ultrametric phylogeny. The abundances of species j in community A (A_j) and community B (B_j) are divided by the total abundance of all species in community A (A_T) and B (B_T).

We, therefore, now have the ability to calculate the ability to calculate a weighted and scaled version of the UniFrac metric. The weighting of UniFrac by abundance is a welcome improvement over the original UniFrac metric, but some conceptual and mathematical problems with this weighted metric have been pointed out [117]. One particularly troubling problem is that the weighted UniFrac does not equal the value from the original UniFrac calculation if abundances become binary. That is, if species are treated as either present or absent in the weighted equation the result would not equal that of the original equation. In other words, the weighted metric is not a natural extension of the original and the two are not coherent. An additional issue with the weighted UniFrac calculation above is that it may overly weight branches with large abundance proportions [117]. A solution to these two issues has been derived by Chen et al. [117] by altering the equation for weighted UniFrac to include a normalizing factor to the equation.

$$d^{(\alpha)} = \frac{\sum_i b_i (p_i^A + p_i^B)^\alpha \left| \frac{p_i^A - p_i^B}{p_i^A + p_i^B} \right|}{\sum_i b_i (p_i^A + p_i^B)^\alpha}$$

Where α is the normalizing factor that can range from zero to one, b_i is the length of branch i , and is the proportion of the abundances subtended by branch i for community A . This modified, and arguably much improved, version of the abundance-weighted UniFrac is available in the GUniFrac package.

“Major” Beta Diversity Measures

The next metric we will discuss is very similar to the unweighted UniFrac method. It was proposed years later by Bryant et al. [119] and is calculated as:

$$PhyloSor = 2 \times \frac{BL_{k_1 k_2}}{(BL_{k_1} + BL_{k_2})}$$

Where $BL_{k_1 k_2}$ is the Faith's Index of the species shared between two communities and BL_{k_1} and BL_{k_2} are the Faith's Indices for the two communities. The PhyloSor metric is a metric of similarity such that as the phylogenetic similarity of two communities increases the PhyloSor value increases. This is the opposite of the un-weighted UniFrac metric which is a dissimilarity metric. Nonetheless, as we will see below, the two metrics are highly correlated.

We can calculate the PhyloSor for two communities using all the calculations we did in the previous subsection for the un-weighted UniFrac. Specifically, all we need to

“Major” Beta Diversity Measures

The calculation of a distance-based metric to quantify the overall dissimilarity between two samples has its roots in Rao [68], which we will discuss above. The metrics utilized today are generally natural and minor extensions to the original Rao metric. For example, the pairwise phylogenetic or trait distance (D_{pw}) between two communities can be calculate as:

$$D_{pw} = \frac{\sum_i^{nk_1} \sum_j^{nk_2} \delta_{ij}}{nk_1 \times nk_2}$$

Where there are nk_1 species in community k_1 , nk_2 species in community k_2 , δ_{ij} is the phylogenetic or functional distance between species i in community k_1 to all species in community k_2 . In this instance species i or species j can occur in both communities and conspecific distances (i.e. zeros) are counted in the mean pairwise distances between communities.

“Major” Beta Diversity Measures

Next we would like to extend the D_{pw} metric to include the abundances of species. It is critical that the abundance-weighted metric is a natural mathematical extension of the original un-weighted metric and, as we have seen above, this is not always accomplished. With this in mind, we can define an abundance-weighted version of the pairwise dissimilarity (D_{pw}').

$$D_{pw}' = \sum_i^{nk_1} \sum_j^{nk_2} \delta_{ij} f_i f_j$$

“Major” Beta Diversity Measures

$$D_{nn} = \frac{\sum_i^{nk_1} \min \delta_{ik_2} + \sum_j^{nk_2} \min \delta_{jk_1}}{nk_1 + nk_2}$$

Where $\min \delta_{ik_2}$ is the minimum phylogenetic distance between species i in community k_1 and all species in community k_2 , $\min \delta_{jk_1}$ is the minimum phylogenetic distance between species j in community k_2 and all species in community k_1 and n is the number of species in the respective communities. This metric has been utilized in phylogenetic and functional beta diversity analyses and provides a refined or ‘terminal’ measure of dissimilarity [64, 80, 120-121]. As with the D_{nn} calculations above, conspecifics are included in this D_{nn} equation such that if two communities have the identical species composition the D_{nn} value will be zero.

“Major” Beta Diversity Measures

Like all of the other phylogenetic and functional diversity metrics presented in this book it is often very useful to weight the metrics by abundance. We will call the abundance-weighted version of the nearest neighbor metric D_{nn}' . The D_{nn}' equation for simply weights all of the nearest neighbor distances by the abundance of the focal species and averages across all individuals.

$$D_{nn}' = \frac{\sum_i^{nk_1} \min \delta_{ik_2} \times f_i + \sum_j^{nk_2} \min \delta_{jk_1} \times f_j}{\sum_i^{nk_1} f_i + \sum_j^{nk_2} f_j}$$

Where f_i is the abundance of species i in community k_1 and f_j is the abundance of species j in community k_2 . The original D_{nn} calculation and the present calculation include conspecific species, but this can be easily modified to remove conspecifics. The calculation of this metric in R first requires that we convert our community data matrix to relative abundances by dividing values in the rows by the sum of the values in that row.

Next we would like to extend the D_{pw} metric to include the abundances of species. It is critical that the abundance-weighted metric is a natural mathematical extension of the original un-weighted metric and, as we have seen above, this is not always accomplished. With this in mind, we can define an abundance-weighted version of the pairwise dissimilarity (D_{pw}').

$$D_{pw}' = \sum_i^{nk_1} \sum_j^{nk_2} \delta_{ij} f_i f_j$$

Where f_i is the abundance of species i in community one and f_j is the abundance of species j in community two. We could alternatively write the equation with the summations in the numerator going across all individuals instead of species and the with no abundance parameters. Now that we have defined an abundance-weighted pairwise dissimilarity metric that is a natural extension of the un-weighted metric, we can proceed to calculate it in R using an example quantifying the phylogenetic dissimilarity between two communities. We start by transforming our community data matrix that contains the raw count of individuals into relative abundances by dividing the values in each row by the sum of the values in that row.