For comparing phylogenetic diversity among communities, go ahead and use synthesis phylogenies

Daijiang Li[1], Lauren Trotta[1], Hannah E. Marx[2], Julie M. Allen[3], Miao Sun[3], Douglas E. Soltis[3], Pamela S. Soltis[3], Robert P. Guralnick[3], and Benjamin Baiser[1]

27 September, 2018

[1]Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32611, USA

[2]Department of Ecology and Evolutionary Biology, University of Arizona, PO Box 210088, Tucson, AZ 85721, USA

[3]Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA

Email: daijianglee@gmail.com (Corresponding author, ORCID: http://orcid.org/0000-0002-0925-3421)

Running headline: Phylogenetic diversity based on different trees

Abstract:

Should we build our own phylogenetic trees based on gene sequence data, or can we simply use available synthesis phylogenies? This is a fundamental question that any study involving a phylogenetic framework must face at the beginning of the project. Building a phylogeny from gene sequence data (purpose-built phylogeny) requires more effort and expertise than subsetting

an already available phylogeny (synthesis-based phylogeny). If phylogenetic diversity estimates based on these two types of phylogenies are highly correlated, using readily available synthesis-based phylogenies is justified for comparing phylogenetic diversity among communities. However, a comparison of how these two approaches to building phylogenetic trees influence the calculation of phylogenetic diversity has not been explicitly tested. We generated three purpose-built phylogenies and their corresponding synthesis-based trees (two from Phylomatic and one from the Open Tree of Life). We then used a simulation approach to generate 1000 communities with a fixed number of species per site and compared the effects of different trees on estimates of phylogenetic alpha and beta diversity using Spearman's rank-based correlation and linear mixed models. Synthesis-based phylogenies generally over-estimated phylogenetic diversity when compared to purpose-built ones. However, their resulting measures of phylogenetic diversity were highly correlated (Spearman's $r > 0.8$ in most cases). Mean pairwise distance (both alpha and beta) is the index that is most robust to the differences in tree construction that we tested. Measures of phylogenetic diversity based on the Open Tree of Life showed the highest correlation with measures based on the purpose-built phylogenies. For comparing phylogenetic diversity among communities, our results justify taking advantage of recently developed and continuously improving synthesis trees such as the Open Tree of Life.

Key words: alpha diversity, beta diversity, community phylogenetic structure, open tree of life, phylogenetic diversity, purpose-built phylogeny, synthesis tree.

Introduction

Phylogenies describe the evolutionary history of species and provide important tools to study ecological and evolutionary questions (Baum and Smith 2012). Recently, phylogenies have been used to better understand patterns of community assembly. The phylogenetic structure of ecological communities can lend insight into the processes by which local communities assemble

from regional species pools (Webb et al. 2002). For example, if closely related species are more likely to co-occur in the same habitats, we might suspect that these species share traits that allow them to have a positive growth rate under the environmental conditions in these habitats. To test whether closely related species are more or less likely to co-occur, one common approach is to calculate the phylogenetic diversity of communities and then compare the observed phylogenetic diversity with those expected by chance through different null models. There is a growing body of literature using this community phylogenetic approach, documenting the phylogenetic structure of ecological communities across taxa and scales (Webb et al. 2002, Cavender-Bares et al. 2006, Helmus et al. 2007, Vamosi et al. 2009, Cardillo 2011, Smith et al. 2014, Li et al. 2017, Marx et al. 2017).

As an important facet of biodiversity, phylogenetic diversity (Faith 1992) also plays a crucial role in conservation biology by complementing more traditional taxonomic measures of biodiversity (e.g., species richness). For example, two communities can have the same number of species but differ drastically in their phylogenetic diversity depending on relatedness of the constituent species. The community with higher phylogenetic diversity, representing taxa more distantly related to each other, is expected to be more stable and productive given its greater evolutionary potential to adapt to changing environmental conditions (Forest et al. 2007, Maherali and Klironomos 2007, Lavergne et al. 2010). Therefore, all else being equal, a community with higher phylogenetic diversity should have higher conservation priority.

The information gained from phylogenetic diversity analyses are only as good as the species composition data and the phylogenies from which they are generated. In this manuscript, we explore how tree generation affects these phylogenetic diversity metrics. Generally, ecologists and evolutionary biologists use two common approaches to build phylogenies for community phylogenetic analyses. The first approach is for a researcher to generate his/her own phylogenies for a set of target species based on gene sequence data. We refer to such phylogenies as purpose-built phylogenies. The second approach is to derive phylogenies based on available

synthesis trees, such as the Open Tree of Life[1], or classifications, such as the Angiosperm Phylogeny Group (APG IV et al. 2016), by pruning or sampling, respectively, from the resource so that the phylogeny contains only the target species. We refer to such phylogenies as synthesis-based phylogenies. To a certain extent, one can argue that a synthesis tree could be a purpose-built tree for a larger set of species, but the sources for deriving the synthesis-based trees vary in scope, methodology, assumptions, and content (see Materials and Methods for further description of source trees for synthesis-based phylogenies). From a researcher perspective, a purpose-built phylogeny is a major undertaking but offers potential to utilize taxonomic and phylogenetic expertise often needed in order to successfully construct trees. Synthesis trees, as compilations of peer-reviewed phylogenetic hypotheses, offer an immediately available, but typically less customizable output to researchers. We thus use these two terms (purpose-built and synthesis-based) to categorize the underlying methods and researcher cost-benefits to obtain phylogenies.

Generating a purpose-built tree requires more effort and expertise than subsetting a well-developed phylogeny or sampling from a classification. Generally, purpose-built trees are constructed by using newly generated sequence data and then combining those data with data already available on GenBank; although in many cases the researcher may simply use what is in GenBank. The first step requires gathering tissue for taxa of interest either from field or museum collections, extracting DNA from these tissue samples, and then identifying, amplifying, and sequencing appropriate loci. The gene regions selected are typically based on the taxa of interest and discipline-accepted standards. Resulting sequences are aligned in programs like MUSCLE (Edgar 2004). Sequences are also commonly sourced entirely or as an addition to sequence data already in databases like GenBank with the help of computational pipelines such as PHLAWD (Smith et al. 2009). Appropriate models of evolution for phylogenetic estimation are determined using programs like PartitionFinder (Lanfear et al. 2012) such that each gene region in a set of concatenated sequences can be treated separately. The most appropriate models of nucleotide

[1] https://tree.opentreeoflife.org/opentree

evolution are used to estimate phylogenies in Maximum Likelihood (ML) and/or Bayesian Inference (BI) frameworks in programs like RAxML (Stamatakis 2014), MrBayes (Ronquist and Huelsenbeck 2003), and BEAST (Drummond and Rambaut 2007). Depending on the desired application, it may be necessary to impose topological constraints to ease phylogenetic inference or fossil constraints to scale branch lengths to time. Statistics for clade support are calculated using bootstrap or jack-knifing techniques in an ML framework, and posterior probabilities in BI. Despite the fact that multiple software programs are available to help automate these processes (e.g., phyloGenerator (Pearse and Purvis 2013), SUPERSMART (Antonelli et al. 2017)), many decisions at different steps must be made based on expert knowledge (e.g., Which genes to select? How to select models? Which software program to use? How to estimate divergence time?).

Because of the effort, expertise, and cost required to generate purpose-built phylogenies, many community phylogenetic studies use a second approach: deriving phylogenies from available synthesis trees. Over the past few decades, tremendous advances in computational tools and increasingly available genetic sequence data have led to vastly improved phylogenies for plants (Zanne et al. 2014), birds (Jetz et al. 2012), fishes (Rabosky et al. 2013), and mammals (Bininda-Emonds et al. 2007, Fritz et al. 2009). Such advances in phylogenetics have facilitated the synthesis of all available information to make a comprehensive tree of life on Earth (Hinchliff et al. 2015). With these available synthesis trees and software programs such as Phylomatic (Webb and Donoghue 2005), ecologists can derive phylogenies for the species or communities they are interested in with less effort and limited cost. When different studies use the same synthesis tree to derive their phylogenies, their phylogenetic diversity results are comparable. Importantly, this may not be the case if they use purpose-built phylogenies. In addition, these approaches may avoid some issues when generating phylogenies from sequence data such as taxon sampling effects (Park et al. 2018). However, the tractability of phylogenies based on synthesis trees often comes with the cost of decreased resolution (e.g., increase in polytomies) of the resulting phylogenies compared with purpose-built ones; such trees also have taxonomic gaps, which are often filled using existing classifications to become comprehensive.

Previous studies have demonstrated that most phylogenetic diversity metrics are robust to terminal polytomies (Swenson 2009, Patrick and Stevens 2014, Boyle and Adamowicz 2015). These studies, however, used simulated phylogenies or compared different posterior purpose-built phylogenies. Therefore, they provided little practical advice about selecting between purpose-built and synthesis-based phylogenies for ecological studies. In this study, we compared phylogenetic diversity metrics calculated from purpose-built phylogenies and corresponding phylogenies derived from three commonly used sources. It is important to note that we do not treat the purpose-built phylogenies as a gold standard and we recognize that sampling bias of both taxa and genes, combined with variation introduced through the tree-building process (e.g., tree reconstruction methods, assessment of support, etc.), can compromise the accuracy of purpose-built phylogenies. However, these issues – and others – apply also to the source trees used for synthesis-based phylogenies, although perhaps at different scales. Our aim here is to quantify the influence of the two tree construction techniques on measures of phylogenetic diversity that are commonly employed in the rapidly growing field of community phylogenetics.

Materials and Methods

Purpose-built phylogenies

We collected three "purpose-built" phylogenies from published and unpublished sources. The first purpose-built phylogeny is for 540 plant taxa in the globally critically imperiled pine rockland ecosystem in South Florida, USA (Trotta et al. 2018). The second phylogeny consists of 1,064 alpine plant taxa in France (Marx et al. 2017). The third purpose-built phylogeny has 1,548 plant species with distributions in Florida, USA (Allen et al. in review). All three phylogenies were estimated from sequence data and were time-calibrated (i.e., chronograms). When using time-calibrated phylogenies, phylogenetic diversity measures the amount of evolution in time-units, and this is the measure we focus on here. For details about phylogeny building processes, see the Appendix.

Commonly available phylogenies

For each of the three purpose-built phylogenies, we generated four phylogenies based on different sources with which to compare phylogenetic alpha and beta diversity. The first two were generated using Phylomatic v4.2 (Webb and Donoghue 2005) using two different backbone trees: `R20120829` (APG III 2009) and `zanne2014` (Zanne et al. 2014). We call the first phylogeny `tree_apg` and the second one `tree_zanne`. The phylogeny `tree_zanne` has branch lengths because the backbone tree `zanne2014` was inferred from seven gene regions for >32k plant species and was time-calibrated using 'congruification' (Eastman et al. 2013). In contrast, the phylogeny `tree_apg` has no branch lengths and is based, not on the result of a phylogenetic analysis *per se*, but on a series of phylogenetic analyses as summarized by the Angiosperm Phylogeny Group III (2009). APG classification now updated as APG IV (2016), but Phylomatic uses APG III. To add branch lengths, we used the `bladj` algorithm in Phylocom (Webb et al. 2008) to convert the tree to a chronogram using a set of the minimum node ages given by Wikström et al. (2001).

The third phylogeny was derived from the Open Tree of Life (Hinchliff et al. 2015), a recent comprehensive phylogeny for ~ 2.3 million named species of life, including all eukaryotes, Archaea, and Bacteria. This phylogeny, which we call `tree_otl`, is a supertree constructed from available source trees, with missing species added based on taxonomy; this resulting tree therefore contains many polytomies and also did not include branch lengths. To calculate branch lengths, we first identified descendants for each of the internal nodes in `tree_otl` and then searched for their divergence time in the TimeTree of Life database (Kumar et al. 2017). The TimeTree database was compiled based on 3,163 studies and 97,085 species (as of October 10, 2017). For a pair of species included in this database, we extracted their average divergence time from all previous studies. Using the divergence date of internal nodes from the TimeTree database, we then determined branch lengths of `tree_otl` using Phylocom (Webb et al. 2008) and its `bladj` function. Recently, an updated phylogeny with branch lengths for seed plants based on the Open Tree of Life was published (Smith and Brown 2018); however, we did not use this seed plant phylogeny as a source

because it contains only seed plants, and our purpose-built phylogenies also contain other clades of vascular plants.

The fourth phylogeny was a random coalescent phylogeny generated using the `rcoal` function from the R package `ape` (Paradis et al. 2004). The random tree was then scaled to have a root age that was the average root age of `tree_apg`, `tree_zanne`, and `tree_otl`. Results based on the random phylogeny should not correlate with those based on other phylogenies.

Not every species from the purpose-built phylogenies was found in all of the synthesis phylogenies. For the pine rockland phylogeny, 514 out of 540 species (95.2%) were found in all phylogenies. For the alpine plant phylogeny, 994 out of 1064 species (93.4%) were found in all phylogenies. For the Florida flora phylogeny, 1472 out of 1548 species (95.1%) were found in all phylogenies. Therefore, we pruned the purpose-built phylogenies to have the same species as their corresponding synthesis tree. In practice, one could insert species that were missing from the derived phylogeny as polytomies in the same genus, so that all species could be included in the analysis.

Generation of community assemblages

For each purpose-built phylogeny, we simulated 1000 presence/absence site-by-species matrices. Each matrix has 30 sites, with species within each site randomly selected from the phylogeny tips representing the species pool. We fixed species richness of each site to be 50 to remove any effects of species richness on the phylogenetic diversity measures. Without setting all sites to have the same number of species, results based on different phylogenies will correlate with each other. For example, it is likely that results from `tree_random` will be highly correlated with results from other phylogenies (Appendix Fig. A1). This is because most phylogenetic diversity metrics correlate with species richness, which, in turn, will lead to correlations among them and confound the comparisons of effects of phylogeny *per se* on the measurement of phylogenetic diversity. Removing the constraint of using the same species richness does not affect our results and

conclusions (Appendix Fig. A1, A2). In our current setting, the maximum total number of species across 30 sites is 30 × 50 = 1500, which is similar to the number of tips in the largest purpose-built phylogeny in our study. We selected species from the species pool randomly because previous studies demonstrated that different approaches to species selection give similar results (Swenson 2009).

Phylogenetic diversity measurements

For each site-by-species matrix, we calculated alpha and beta phylogenetic diversity for each of the phylogenies using indices that are commonly used in community phylogenetic studies. For phylogenetic alpha diversity, we used Faith's PD (PD), mean pairwise distance (MPD), and mean pairwise distance between the closest relatives (MNTD). PD calculates the sum of the branch lengths of all species present in an assemblage (Faith 1992). We did *not* include the root of the phylogeny when calculating PD. MPD calculates the average pairwise distance between all species, and MNTD calculates the average pairwise distance between the closest relatives in an assemblage (Webb et al. 2002). We selected these three metrics for phylogenetic alpha diversity among the myriad of metrics available because they are most commonly used and represent different but complementary information about phylogenetic structure of communities (Miller et al. 2017, Tucker et al. 2017).

For phylogenetic beta diversity, we applied UniFrac (Unif), inter-assemblage MPD (MPD_beta), inter-assemblage MNTD (MNTD_beta), and phylogenetic community dissimilarity (PCD) to all possible unique combinations of assemblage pairs. Unif is derived from the Jaccard dissimilarity index and calculates the total branch length unique to each assemblage relative to the total branch length of all species in a pair of assemblages (Lozupone and Knight 2005). Therefore, it measures the fraction of evolutionary history unique to each assemblage. MPD_beta and MNTD_beta were derived from MPD and MNTD, respectively, but instead of comparing species within the same assemblage, they compare species from two different assemblages (Webb et al. 2008). PCD

measures pairwise phylogenetic dissimilarity between assemblages by asking how much of the variance of values of a hypothetical trait among species in one assemblage can be predicted by the values of species from another. PCD is independent of species richness of the pair of assemblages and has relatively higher statistical power than other common metrics (Ives and Helmus 2010).

As PD and MNTD are both correlated with species richness (Miller et al. 2017), null models that retain species composition while randomly shuffling tips of the phylogeny are commonly used to standardize phylogenetic diversity results. Despite the fact that MPD is independent of species richness, its variance changes relative to species richness (Miller et al. 2017). Therefore, null models are also frequently applied to MPD. Using the null model, standardized effect size (SES) for each metric can be calculated as $SES = \frac{X_{obs} - mean(X_{null})}{sd(X_{null})}$, where X_{obs} is the observed value, and X_{null} are the n values calculated based on null models. Recently, analytic solutions for the SES of phylogenetic alpha diversity metrics were developed (Tsirogiannis and Sandel 2016). The analytic solutions eliminate the need for computationally expensive simulations used to calculate SES values, especially for studies in high-diversity systems. In our simulations, because all sites have the same species richness, we expected that the SES values based on the analytic solutions would have the identical results as the observed phylogenetic diversity values for the statistical analyses we conducted (correlation and linear mixed models, see the Statistical analyses section below). Our simulations confirmed this expectation (Appendix Fig. A3-A6). No analytic solutions for the SES of Unif, MNTD_beta, and PCD are available. However, the pairwise beta diversity metrics share the same core formula with their corresponding alpha diversity metrics. We thus expect that the results based on SES of these beta diversity metrics will be the same as those based on the observed diversity values in our simulations. Given the similarity in results between raw and standardized phylogenetic alpha diversity measures and the large computational burden of calculating SES for phylogenetic beta diversity metrics, we did not include the results for SES in this study.

Statistical analyses

We have two primary goals. First, we want to test the correlation between phylogenetic diversity values calculated from purpose-built phylogenies and those calculated from synthesis phylogenies. Second, we want to investigate whether phylogenetic diversity calculated from synthesis phylogenies over- or under-estimates phylogenetic diversity when compared to purpose-built phylogenies. For the first goal, we calculated the average Spearman's rank-based measure of the correlation between phylogenetic diversity values from all phylogenies across the 1000 simulations. We used rank-based correlation because it is the relative phylogenetic diversity, not the absolute one, that we are interested in. For the second goal, we used Linear Mixed Models (LMMs) with phylogenetic diversity values from the purpose-built phylogeny as the response variable, the phylogenetic diversity values from one of the synthesis phylogenies as the predictor, and the simulation dataset as the random term. We scaled the diversity values to have mean zero and standard deviation one before fitting the models. We also forced the regression line through the origin. If the slope of the regression line is significantly different from zero, then phylogenetic diversity based on purpose-built phylogenies and synthesis phylogenies is highly correlated. Furthermore, if the slope is higher/lower than one, then the phylogenetic diversity based on the synthesis phylogenies under-/over-estimates phylogenetic diversity. For pairwise beta diversity, because of the large number of samples across all 1000 simulations ($\binom{30}{2} \times 1000 = 435{,}000$), we randomly selected 100 simulations on which to conduct LMMs. In addition, for pairwise beta diversity, because one site can be compared with all other sites, the beta diversity values are not independent. To account for this, we included datasets, `site1` within each dataset (the first site in the site pair), and `site2` within each site (the other site in the site pair) as random terms in the LMMs (cf. Li and Waller 2017). The workflow of this study is outlined in Fig. 1. All analyses were conducted with R v3.4.3 (R Core Team 2017).

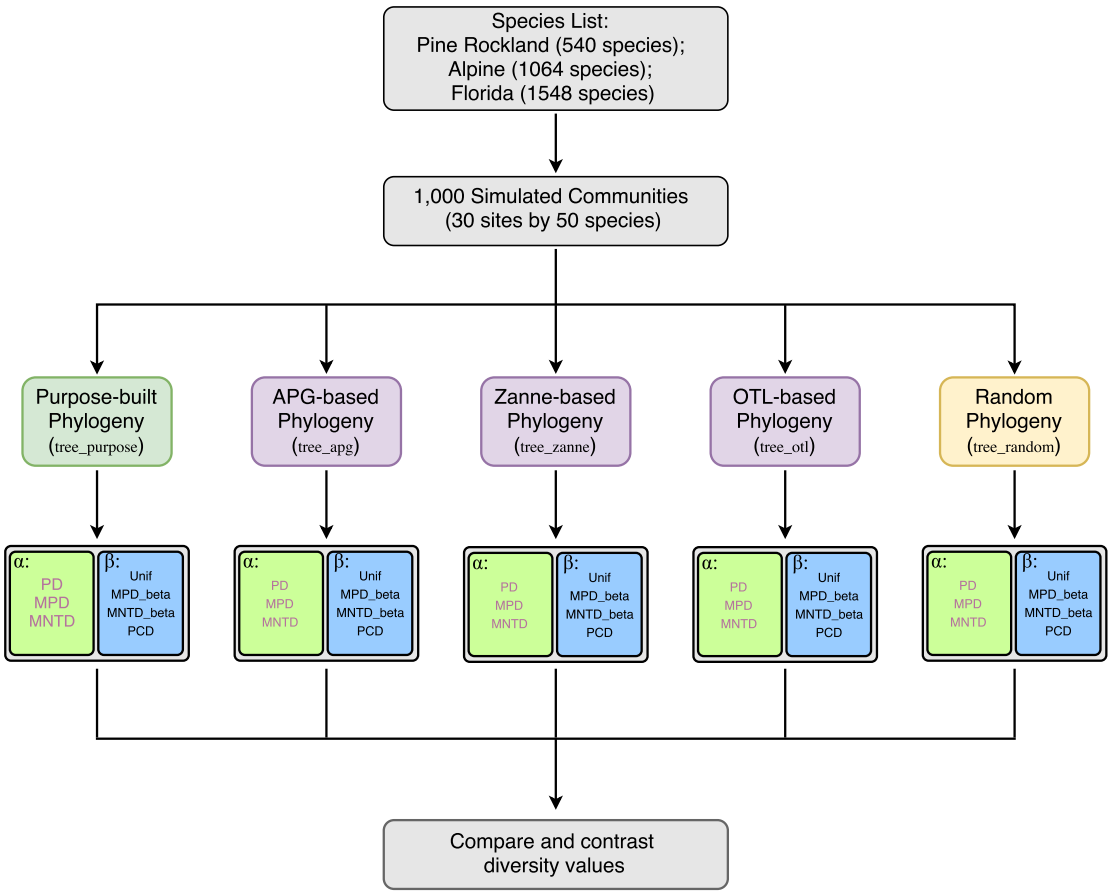

Figure 1: Workflow to assess effects of commonly used synthesis phylogenies on phylogenetic diversity estimations. Abbreviations: APG, Angiosperm Phylogeny Group; OTL, Open Tree of Life; PD, Faith's Phylogenetic diversity; MPD, Mean pairwise distance; MNTD, Mean nearest taxon distance; Unif, Unifraction; PCD, Phylogenetic community dissimilarity.

Results

Alpha diversity

Phylogenetic alpha diversity (PD, MPD, and MNTD) values calculated with different phylogenies (`tree_purpose`, `tree_apg`, `tree_zanne`, and `tree_otl`) were highly correlated. The median Spearman's correlation of the 1000 simulations was larger than 0.63 across all comparisons ($p < 0.05$ for all simulations and comparisons; Fig. 2). In most cases, the median Spearman's correlation was larger than 0.85, especially for PD and MPD. Therefore, PD and MPD were more robust to varying the source of the phylogeny than MNTD. Across all comparisons, diversity values based on `tree_otl` showed the highest correlations with those based on `tree_purpose`, with an average correlation across all comparisons of 0.902. As expected, diversity values based on the random phylogeny `tree_random` were not correlated with diversity values based on other phylogenies, with median Spearman's correlations close to zero (Fig. 2).

The slopes of linear mixed models (LMM) were all less than one (Table 1), suggesting that diversity values based on synthesis phylogenies generally over-estimated the diversity values based on the purpose-built phylogenies. The PD metrics based on the Open Tree of Life phylogeny (`tree_otl`) had estimates closest to those calculated from the purpose-built phylogenies (Table 1).

Beta diversity

The phylogenetic beta diversity results (Unfi, MPD_beta, MNTD_beta, and PCD) show a similar pattern to the alpha diversity results. Beta diversity of community pairs based on different phylogenies was also highly correlated, with the median Spearman's correlation from the 1000 simulations greater than 0.69 across all comparisons (Fig. 3). Overall, phylogenetic beta diversity is more sensitive to the source of the phylogeny than alpha diversity. MPD_beta is the most robust beta diversity metric to the source of the phylogeny, followed by MNTD_beta, Unif, and PCD.

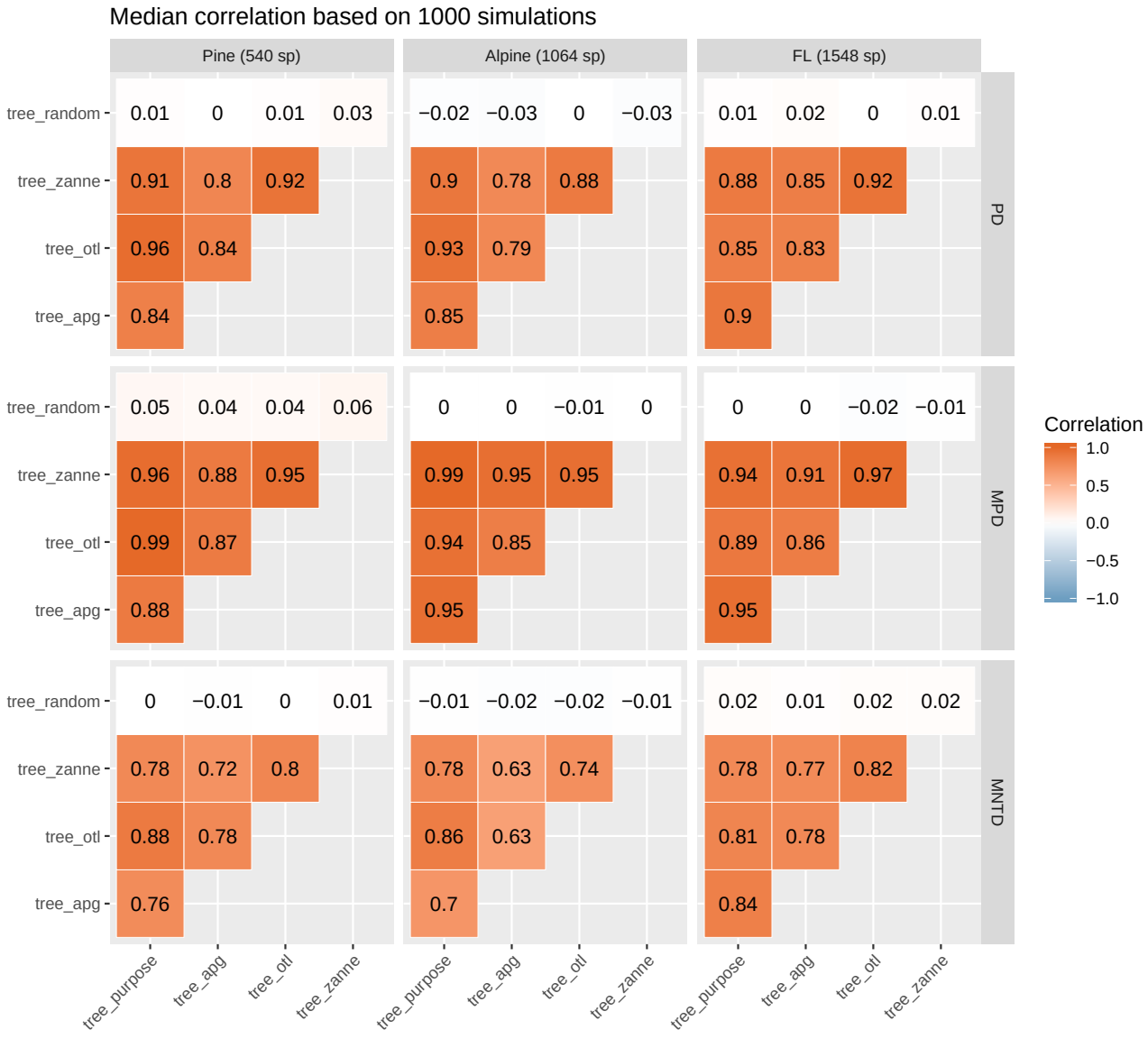

Figure 2: Median correlations of phylogenetic alpha diversity values based on different phylogenies.

Table 1: Slopes based on linear mixed models (LMMs). Within the model, the response variable is the phylogenetic alpha diversity values based on the purpose-built phylogeny; the predictor is the phylogenetic alpha diversity values based on one of the synthesis phylogenies (`tree_apg`, `tree_zanne`, `tree_otl`, and `tree_random`). Therefore, slopes less than one indicate overestimations. Numbers within parentheses are the 95% confidence intervals for the slopes.

index	dataset	tree_apg	tree_zanne	tree_otl	tree_random
PD	Pine (540 sp)	0.843 (0.837, 0.849)	0.917 (0.913, 0.922)	0.971 (0.969, 0.974)	-0.001 (-0.013, 0.01)
PD	Alpine (1064 sp)	0.854 (0.848, 0.86)	0.915 (0.91, 0.919)	0.937 (0.933, 0.941)	-0.022 (-0.034, -0.01)
PD	FL (1548 sp)	0.92 (0.916, 0.924)	0.891 (0.886, 0.896)	0.871 (0.865, 0.876)	0.006 (-0.005, 0.018)
MPD	Pine (540 sp)	0.891 (0.885, 0.896)	0.972 (0.969, 0.974)	0.996 (0.995, 0.997)	0.047 (0.036, 0.059)
MPD	Alpine (1064 sp)	0.957 (0.954, 0.96)	0.997 (0.997, 0.998)	0.941 (0.937, 0.945)	0.004 (-0.008, 0.015)
MPD	FL (1548 sp)	0.962 (0.958, 0.965)	0.95 (0.946, 0.953)	0.895 (0.889, 0.9)	-0.002 (-0.014, 0.009)
MNTD	Pine (540 sp)	0.78 (0.773, 0.788)	0.787 (0.78, 0.794)	0.897 (0.892, 0.902)	0.006 (-0.006, 0.017)
MNTD	Alpine (1064 sp)	0.713 (0.705, 0.721)	0.794 (0.787, 0.801)	0.874 (0.869, 0.88)	-0.016 (-0.028, -0.004)
MNTD	FL (1548 sp)	0.856 (0.85, 0.862)	0.797 (0.79, 0.804)	0.831 (0.824, 0.837)	0.03 (0.018, 0.041)

Again, PD metrics based on `tree_otl` showed the highest correlation with metrics based on the purpose-built tree, followed by `tree_zanne` and `tree_apg`. Beta diversity values based on `tree_random` did not correlate with values based on any other phylogeny.

The slopes of LMMs were generally less than one (Table 2), suggesting over-estimates of beta diversity from the synthesis-based phylogenies compared with the purpose-built phylogenies. However, slopes for MPD_beta values based on `tree_otl` were all greater than one, suggesting that beta PD metrics were under-estimated when compared to those calculated from the purpose-built trees. Metrics based on `tree_zanne` for the flora of Florida dataset were also under-estimated (Table 2). For the other beta diversity metrics (i.e., Unif, MNTD_beta, and PCD), `tree_otl` generally gave results closer to those based on the purpose-built trees than did the other synthesis-based phylogenies.

Discussion

We examined how different phylogenies, purpose-built and synthesis-based, influenced phylogenetic alpha and beta diversity measures commonly used in community phylogenetic

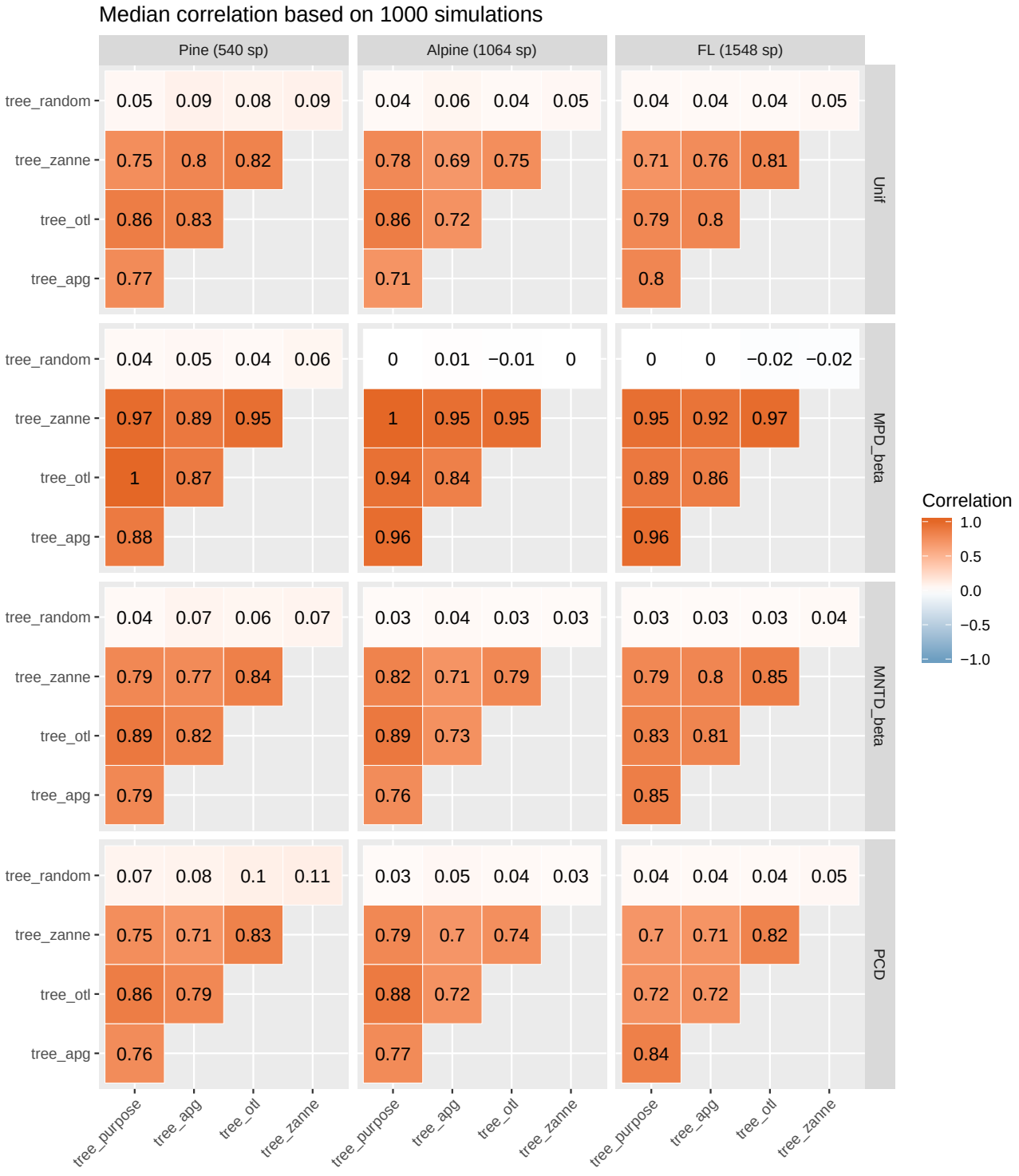

Figure 3: Median correlations of phylogenetic beta diversity values based on different phylogenies.

Table 2: Slopes based on linear mixed models (LMMs). Within the model, the response variable is the phylogenetic beta diversity values based on the purpose-built phylogeny; the predictor is the phylogenetic beta diversity values based on one of the synthesis phylogenies (`tree_apg`, `tree_zanne`, `tree_otl`, and `tree_random`). Therefore, slopes less than one indicate overestimations, and slopes greater than one are underestimates. Numbers within parentheses are the 95% confidence intervals for the slopes.

index	dataset	tree_apg	tree_zanne	tree_otl	tree_random
Unif	Pine (540 sp)	0.823 (0.816, 0.83)	0.791 (0.785, 0.797)	0.87 (0.866, 0.875)	0.054 (0.04, 0.067)
Unif	Alpine (1064 sp)	0.806 (0.798, 0.815)	0.87 (0.863, 0.876)	0.895 (0.89, 0.9)	0.063 (0.052, 0.074)
Unif	FL (1548 sp)	0.871 (0.865, 0.877)	0.792 (0.785, 0.8)	0.815 (0.809, 0.822)	0.065 (0.052, 0.078)
MPD_beta	Pine (540 sp)	0.343 (0.336, 0.35)	0.967 (0.959, 0.975)	1.249 (1.234, 1.263)	0.011 (0.004, 0.017)
MPD_beta	Alpine (1064 sp)	0.801 (0.794, 0.807)	0.977 (0.975, 0.98)	1.12 (1.104, 1.137)	0.002 (-0.002, 0.006)
MPD_beta	FL (1548 sp)	0.788 (0.78, 0.797)	1.347 (1.331, 1.363)	1.815 (1.787, 1.842)	-0.003 (-0.007, 0.001)
MNTD_beta	Pine (540 sp)	0.855 (0.846, 0.864)	0.854 (0.846, 0.862)	0.931 (0.926, 0.937)	0.049 (0.038, 0.061)
MNTD_beta	Alpine (1064 sp)	0.891 (0.883, 0.9)	0.948 (0.941, 0.955)	0.94 (0.935, 0.945)	0.062 (0.051, 0.073)
MNTD_beta	FL (1548 sp)	0.788 (0.782, 0.793)	0.759 (0.752, 0.765)	0.752 (0.746, 0.758)	0.035 (0.024, 0.045)
PCD	Pine (540 sp)	0.855 (0.847, 0.864)	0.836 (0.827, 0.846)	0.876 (0.869, 0.884)	0.083 (0.07, 0.095)
PCD	Alpine (1064 sp)	0.824 (0.816, 0.832)	0.909 (0.9, 0.918)	0.905 (0.898, 0.911)	0.076 (0.065, 0.088)
PCD	FL (1548 sp)	0.805 (0.798, 0.812)	0.755 (0.747, 0.763)	0.727 (0.718, 0.735)	0.052 (0.039, 0.066)

analyses. We found two main results. First, the synthesis phylogenies generally over-estimated phylogenetic diversity compared with purpose-built phylogenies. This is not surprising because synthesis phylogenies generally have higher proportions of polytomies than purpose-built ones, which, in turn, leads to larger distances between species within these polytomies. This result agrees with Boyle and Adamowicz (2015) and Qian and Zhang (2016) but contradicts Swenson (2009), who found that phylogenies with more polytomies under-estimated phylogenetic diversity. Second, phylogenetic diversity values calculated from synthesis trees were highly correlated with those based on purpose-built phylogenies, even if they were over-estimated. These results hold for both alpha and beta diversity and for phylogenies with different numbers of tips. While our study focuses on plants, we expect that our results will generalize to any taxonomic group. Therefore, phylogenies derived from synthesis trees can provide similar results to purpose-built phylogenies while saving effort and time when quantifying and comparing phylogenetic diversity of communities.

One main reason for this conclusion is that, as ecologists and conservation biologists, we mostly care about the relative diversity among communities instead of their absolute diversity. For

example, for a set of communities within one region, we may be interested in which communities have the highest/lowest phylogenetic diversity. The absolute phylogenetic diversity of each community does not mean much without comparing it to other communities. Because phylogenetic values based on different phylogenies are highly correlated with each other, the information available for community phylogenetic questions does not differ much between approaches. Even though such synthesis phylogenies may over-estimate absolute phylogenetic diversity for communities, the relative phylogenetic diversity among communities will be similar to those calculated from typically better resolved but less accessible phylogenies. Based on the information provided by relative values of phylogenetic diversity, the potential improved resolution of purpose-built trees for calculating the absolute PD may not be worth the effort for community phylogenetic questions.

Our finding that phylogenetic diversity metrics are relatively insensitive to the phylogenies from which they are derived has been supported by other recent studies. For example, using simulated fully bifurcating and gradually unresolved phylogenies, Swenson (2009) found that phylogenetic diversity measures are generally robust to the uncertainty of the phylogenies, especially if the uncertainty is concentrated in recent nodes of the phylogeny. Using multiple posterior phylogenies of bats, Patrick and Stevens (2014) rearranged branches across these phylogenies and also found that phylogenetic diversity measures are robust to the phylogenies from which they are calculated. More recently, Cadotte (2015) transformed a phylogeny with different evolution models and found that phylogenetic diversity measures are insensitive to the branch lengths of the phylogeny; getting the topology right is more important when calculating phylogenetic diversity. Qian and Zhang (2016) found similar phylogenetic diversity values of the angiosperm tree flora of North America based on phylogenies derived from Zanne et al. (2014) and Phylomatic (Webb and Donoghue 2005). These studies, however, only focused on alpha diversity. Our study extends the literature by also examining the effects of phylogenies on beta diversity. We found the same pattern for beta diversity and alpha diversity. Taken together, a general pattern emerges: community phylogenetic alpha and beta diversity metrics are robust to reasonably good modern

phylogenies.

Why are phylogenetic diversity values from purpose-built and synthesis phylogenies highly correlated? There are two possible reasons. First, both purpose-built and synthesis phylogenies likely share a similar systematic backbone and empirical resources such as genes, taxonomies, and expert knowledge. This guarantees that phylogenetic diversity based on these phylogenies will not be dramatically different. Second, phylogenetic diversity metrics aggregate (by summing or averaging) all information into one value for each site, which could help buffer most uncertainty and further mask most of the differences between different phylogenies.

Our results should encourage ecologists to increasingly include phylogenetic analyses in community ecology studies given the growing accessibility of synthesis phylogenies and the robustness of phylogenetic diversity measures based on them. However, our results should not discourage the construction of purpose-built phylogenies, which are clearly valuable for many ecological and evolutionary questions. This is especially the case for purpose-built trees constructed from local DNA samples. The sequencing of species in a given community can yield data for species that have never been sequenced before. These new sequences can then be incorporated into synthesis trees, improving their resolution for future research. Direct sequencing of samples collected for a community is also important when the community contains un-described (Pons et al. 2006) or cryptic species (Hebert et al. 2004). Furthermore, for many taxonomic groups, synthesis trees are not available or are far too poorly sampled, and constructing purpose-built trees is the only approach possible for community phylogenetic analyses.

Conclusion

Community phylogenetics is rapidly becoming an important component of community ecology, macroecology, and biodiversity conservation (Webb et al. 2002, Vamosi et al. 2009). For calculations and comparisons of phylogenetic diversity of communities, an important question

arises: can we derive phylogenies from already-available synthesis trees, or should we generate our own purpose-built phylogenies? Our results suggest that phylogenies derived from common synthesis trees over-estimate phylogenetic diversity metrics when compared to purpose-built trees, but values of phylogenetic diversity are highly correlated with purpose-built metrics. Particularly, the Open Tree of Life, which includes all major phylogenetic groups (e.g. plants, birds, fishes, mammals, fungi, Archaea, Bacteria, etc.), produced the most similar values of phylogenetic diversity when compared to metrics derived from purpose-built trees. Furthermore, a recently updated Open Tree of Life phylogeny for seed plants has branch lengths calculated based on molecular data (Smith and Brown 2018). With new data and studies continuously being integrated into synthesis trees such as the Open Tree of Life, these resources are poised to continue to improve rapidly. As a result, for comparing phylogenetic diversity among communities, we recommend taking advantage of recent well-developed products such as the Open Tree of Life.

Data Accessibility

All phylogenies and R code used will be uploaded to figshare upon acceptance.

Acknowledgments

This study was supported by NSF grants ABI-458034 to BB, DEB-1442280 and DBI-1458640 to PSS and DES, EF-1115210 and DBI-1547229 to PSS, and EF-1550838 (supported HEM).

References

Antonelli, A. et al. 2017. Toward a self-updating platform for estimating rates of speciation and migration, ages, and relationships of taxa. - Systematic biology 66: 152–166.

APG III 2009. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG iii. - Botanical Journal of the Linnean Society 161: 105–121.

APG IV et al. 2016. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG iv. - Botanical Journal of the Linnean Society 181: 1–20.

Baum, D. A. and Smith, S. D. 2012. Tree thinking: An introduction to phylogenetic biology. - Roberts; Co., Greenwood Village, CO.

Bininda-Emonds, O. R. et al. 2007. The delayed rise of present-day mammals. - Nature 446: 507.

Boyle, E. E. and Adamowicz, S. J. 2015. Community phylogenetics: Assessing tree reconstruction methods and the utility of dna barcodes. - PloS one 10: e0126662.

Cadotte, M. W. 2015. Phylogenetic diversity–ecosystem function relationships are insensitive to phylogenetic edge lengths. - Functional Ecology 29: 718–723.

Cardillo, M. 2011. Phylogenetic structure of mammal assemblages at large geographical scales: Linking phylogenetic community ecology with macroecology. - Philosophical Transactions of the Royal Society B: Biological Sciences 366: 2545–2553.

Cavender-Bares, J. et al. 2006. Phylogenetic structure of floridian plant communities depends on taxonomic and spatial scale. - Ecology 87: S109–S122.

Drummond, A. J. and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. - BMC evolutionary biology 7: 214.

Eastman, J. M. et al. 2013. Congruification: Support for time scaling large phylogenetic trees. -

412 Methods in Ecology and Evolution 4: 688–691.

413 Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput.
414 - Nucleic acids research 32: 1792–1797.

415 Faith, D. P. 1992. Conservation evaluation and phylogenetic diversity. - Biological conservation 61:
416 1–10.

417 Forest, F. et al. 2007. Preserving the evolutionary potential of floras in biodiversity hotspots. -
418 Nature 445: 757–760.

419 Fritz, S. A. et al. 2009. Geographical variation in predictors of mammalian extinction risk: Big is
420 bad, but only in the tropics. - Ecology letters 12: 538–549.

421 Hebert, P. D. et al. 2004. Ten species in one: DNA barcoding reveals cryptic species in the
422 neotropical skipper butterfly astraptes fulgerator. - Proceedings of the National Academy of
423 Sciences of the United States of America 101: 14812–14817.

424 Helmus, M. R. et al. 2007. Separating the determinants of phylogenetic community structure. -
425 Ecology letters 10: 917–925.

426 Hinchliff, C. E. et al. 2015. Synthesis of phylogeny and taxonomy into a comprehensive tree of life.
427 - Proceedings of the National Academy of Sciences 112: 12764–12769.

428 Ives, A. R. and Helmus, M. R. 2010. Phylogenetic metrics of community similarity. - The American
429 Naturalist 176: E128–E142.

430 Jetz, W. et al. 2012. The global diversity of birds in space and time. - Nature 491: 444–448.

431 Kumar, S. et al. 2017. TimeTree: A resource for timelines, timetrees, and divergence times. -
432 Molecular Biology and Evolution 34: 1812–1819.

433 Lanfear, R. et al. 2012. PartitionFinder: Combined selection of partitioning schemes and

substitution models for phylogenetic analyses. - Molecular biology and evolution 29: 1695–1701.

Lavergne, S. et al. 2010. Biodiversity and climate change: Integrating evolutionary and ecological responses of species and communities. - Annual review of ecology, evolution, and systematics 41: 321–350.

Li, D. and Waller, D. M. 2017. Fire exclusion and climate change interact to affect long-term changes in the functional composition of plant communities. - Diversity and Distributions 23: 496–506.

Li, D. et al. 2017. Can functional traits account for phylogenetic signal in community composition? - New Phytologist 214: 607–618.

Lozupone, C. and Knight, R. 2005. UniFrac: A new phylogenetic method for comparing microbial communities. - Applied and environmental microbiology 71: 8228–8235.

Maherali, H. and Klironomos, J. N. 2007. Influence of phylogeny on fungal community assembly and ecosystem functioning. - science 316: 1746–1748.

Marx, H. E. et al. 2017. Riders in the sky (islands): Using a mega-phylogenetic approach to understand plant species distribution and coexistence at the altitudinal limits of angiosperm plant life. - Journal of biogeography 44: 2618–2630.

Miller, E. T. et al. 2017. Phylogenetic community structure metrics and null models: A review with new methods and software. - Ecography 40: 461–477.

Paradis, E. et al. 2004. APE: Analyses of phylogenetics and evolution in R language. - Bioinformatics 20: 289–290.

Park, D. S. et al. 2018. Taxon sampling effects on the quantification and comparison of community phylogenetic diversity. - Molecular ecology in press.

Patrick, L. E. and Stevens, R. D. 2014. Investigating sensitivity of phylogenetic community structure metrics using north american desert bats. - Journal of Mammalogy 95: 1240–1253.

457 Pearse, W. D. and Purvis, A. 2013. PhyloGenerator: An automated phylogeny generation tool for
458 ecologists. - Methods in Ecology and Evolution 4: 692–698.

459 Pons, J. et al. 2006. Sequence-based species delimitation for the dna taxonomy of undescribed
460 insects. - Systematic biology 55: 595–609.

461 Qian, H. and Zhang, J. 2016. Are phylogenies derived from family-level supertrees robust for
462 studies on macroecological patterns along environmental gradients? - Journal of systematics and
463 evolution 54: 29–36.

464 Rabosky, D. L. et al. 2013. Rates of speciation and morphological evolution are correlated across
465 the largest vertebrate radiation. - Nature communications in press.

466 R Core Team 2017. R: A language and environment for statistical computing. - R Foundation for
467 Statistical Computing.

468 Ronquist, F. and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed
469 models. - Bioinformatics 19: 1572–1574.

470 Smith, S. A. and Brown, J. W. 2018. Constructing a broadly inclusive seed plant phylogeny. -
471 American journal of botany 105: 302–314.

472 Smith, S. A. et al. 2009. Mega-phylogeny approach for comparative biology: An alternative to
473 supertree and supermatrix approaches. - BMC evolutionary biology 9: 37.

474 Smith, M. A. et al. 2014. Diversity and phylogenetic community structure of ants along a costa
475 rican elevational gradient. - Ecography 37: 720–731.

476 Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large
477 phylogenies. - Bioinformatics 30: 1312–1313.

478 Swenson, N. G. 2009. Phylogenetic resolution and quantifying the phylogenetic diversity and
479 dispersion of communities. - PloS one 4: e4390.

Trotta, L. et al. 2018. Community phylogeny of the globally critically imperiled pine rockland ecosystem. - American journal of botany in press.

Tsirogiannis, C. and Sandel, B. 2016. PhyloMeasures: A package for computing phylogenetic biodiversity measures and their statistical moments. - Ecography 39: 709–714.

Tucker, C. M. et al. 2017. A guide to phylogenetic metrics for conservation, community ecology and macroecology. - Biological Reviews 92: 698–715.

Vamosi, S. et al. 2009. Emerging patterns in the comparative analysis of phylogenetic community structure. - Molecular ecology 18: 572–592.

Webb, C. O. and Donoghue, M. J. 2005. Phylomatic: Tree assembly for applied phylogenetics. - Molecular Ecology Resources 5: 181–183.

Webb, C. O. et al. 2002. Phylogenies and community ecology. - Annual review of ecology and systematics 33: 475–505.

Webb, C. O. et al. 2008. Phylocom: Software for the analysis of phylogenetic community structure and trait evolution. - Bioinformatics 24: 2098–2100.

Wikström, N. et al. 2001. Evolution of the angiosperms: Calibrating the family tree. - Proceedings of the Royal Society of London B: Biological Sciences 268: 2211–2220.

Zanne, A. E. et al. 2014. Three keys to the radiation of angiosperms into freezing environments. - Nature 506: 89–92.