Chapter 19
Metrics and Models of Community Phylogenetics

William D. Pearse, Andy Purvis, Jeannine Cavender-Bares and Matthew R. Helmus

Abstract Community phylogenetics combines ideas from community ecology and evolutionary biology, using species phylogeny to explore the processes underlying ecological community assembly. Here, we describe the development of the field's comparative methods and their roots in conservation biology, biodiversity quantification, and macroevolution. Next, we review the multitude of community phylogenetic structure metrics and place each into one of four classes: *shape*, *evenness*, *dispersion*, and *dissimilarity*. Shape metrics examine the structure of an assemblage phylogeny, while evenness metrics incorporate species abundances. Dispersion metrics examine assemblages given a phylogeny of species that could occupy those assemblages (the source pool), while dissimilarity metrics compare phylogenetic structure between assemblages. We then examine how metrics perform in simulated communities that vary in their phylogenetic structure. We provide an example of model-based approaches and argue that they are a promising area of future research in community phylogenetics. Code to reproduce all these analyses is available in the Online Practical Material (http://www.mpcm-evolution.org). We conclude by discussing future research directions for the field as a whole.

W. D. Pearse (✉) · J. Cavender-Bares
Department Ecology, Evolution, and Behavior, University of Minnesota,
1987 Upper Buford Circle, Saint Paul, MN 55108, USA
e-mail: will.pearse@gmail.com

J. Cavender-Bares
e-mail: cavender@umn.edu

A. Purvis
Department of Life Sciences, Natural History Museum,
Cromwell Road, London SW7 5BD, UK
e-mail: andy.purvis@nhm.ac.uk

M. R. Helmus
Amsterdam Global Change Institute, Department of Animal Ecology,
Vrije Universiteit, 1081 HV Amsterdam, The Netherlands
e-mail: mrhelmus@gmail.com

19.1 Overview

Community phylogenetics seeks to explore the ecological and evolutionary factors that underlie the assembly of communities and how species interactions influence evolutionary and ecosystem processes. The field represents a (re-)integration of community ecology and evolution, in the hope that historical species interactions and environmental conditions reflected in phylogeny can inform us about present-day ecology. However, rapid advances in computational tools, phylogenetic inference methods, DNA databases, and metrics mean the scope of community phylogenetics is constantly expanding and developing.

This chapter should provide the reader with an entry point to begin critically conducting their own community phylogenetic analysis. To this end, the Online Practical Material (http://www.mpcm-evolution.org) contains annotated R (R Core Team 2014) code with which the reader can repeat all the analyses and simulations presented in this chapter. We begin by describing the development of community phylogenetics and follow by outlining a framework to understand community phylogenetic metrics. We then examine the performance of several metrics in a simulated data set and give a brief introduction to the field of community phylogenetic modelling. We conclude the chapter by discussing caveats and future directions for the field.

19.2 Historical Overview of the Metrics of Community Phylogenetics

It was Darwin 1859 who first hypothesised a relationship between species' taxonomic proximity and competitive interactions, arguing that congeners use the same resources and so competition should be strongest among them. While Darwin was interested in how this increasing competition would affect natural selection, later scientists (Jaccard 1901; Elton 1946) would ask how the number of congeners present in a community reflected the biogeographic and ecological processes structuring it. Despite controversies over the sensitivity of such approaches to species richness (Järvinen 1982), the idea that ecological processes could be detected in the evolutionary relationships among species in ecological communities took hold.

Conservation biologists were quick to recognise the utility of phylogeny as a way to quantify species uniqueness and thus aid conservation prioritisation. Vane-Wright et al. (1991) first argued to prioritise more basal evolutionary lineages (acknowledged by May (1990) who published first), and Altschul and Lipman (1990) suggested incorporating time-calibrated phylogenies and Felsenstein's comparative method (1985). Soon after, Faith (1992) coined the phylogenetic diversity (PD) metric as the summed phylogenetic branch length connecting all species in a set to rank areas for preservation. Later metrics partitioned the

phylogenetic diversity of clades among their species to facilitate species-based conservation of phylogenetic diversity (Pavoine et al. 2005; Redding and Mooers 2006; Isaac et al. 2007).

In parallel, the taxonomy-based metrics developed in conservation biology (May 1990; Vane-Wright et al. 1991) were adapted to understand community assembly in degraded ecosystems. Warwick and Clarke (1995) counted the mean number of taxonomic ranks separating community members to derive the *Δ* family of metrics (some members were independently derived by Izsáki and Papp 1995), which were later extended to estimate taxonomic similarity among communities (Izsáki and Price 2001). Although not the first modern study of ecological taxonomic structure (cf., e.g. Douglas and Matthews 1992), Warwick and Clarke (1995) provided the first clear example of how habitat filtering can change the taxonomic (and so phylogenetic) composition of ecological communities.

These antecedents provided the basis for papers (Webb 2000; Webb et al. 2002) that developed a framework and set of hypotheses for the use of phylogenetics in mainstream ecology and mark the beginning of modern community phylogenetics. Webb (2000) developed the Net Relatedness Index (NRI) and the Nearest Taxon Index (NTI) to measure the phylogenetic structure of a tropical forest plot. NRI and NTI examine whether the relatedness of species to one another in a community differs from what would be expected under random assembly from a list of potential species (the source pool). Most community phylogenetic studies assume close relatives are ecologically similar (niche conservatism; reviewed in Wiens et al. 2010). Under this assumption, communities whose species are more closely related than under random assembly (underdispersed, or clustered, communities) reflect habitat-filtered assembly, while communities of unexpectedly distantly related species (overdispersed communities) indicate the influence of competitive exclusion.

The assumption of niche conservatism has subsequently been scrutinised, and inferring ecological process purely on the basis of phylogenetic pattern is now treated with scepticism. For example, in one of the early empirical tests of the Webb et al. (2002) framework, Cavender-Bares et al. (2004) demonstrated that niche convergence (rather than conservatism) among distantly related oak lineages caused overdispersion in hyper-diverse oak forest communities. Subsequently, Valiente-Banuet and Verdú (2007) found that facilitation among distantly related species could lead to overdispersion, and Mayfield and Levine (2010) argued that competition may lead to phylogenetic clustering. The ecological and evolutionary mechanisms that produce phylogenetic community structure vary and depend on where in the tree of life one is looking (phylogenetic and biogeographic scale) and the modes of trait evolution at work (Cavender-Bares et al. 2006). Kraft et al. (2007) demonstrated that when known ecological and evolutionary processes are simulated, the anticipated community phylogenetic patterns are reliably recovered, but Kembel (2009) has shown that dispersal can mask such patterns. Development of model-based methods (see Sect. 19.5) offers hope of explicitly testing mechanistic hypotheses about how evolutionary and ecological processes interact, reconciling many objections about inferring process from phylogenetic pattern.
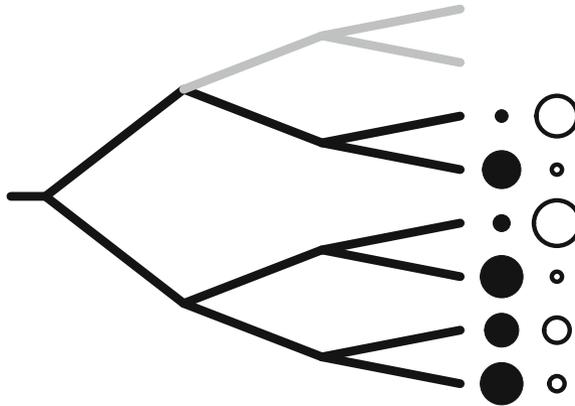
Fig. 19.1 Overview of phylogenetic *shape*, *evenness*, *dispersion*, and *dissimilarity* metrics, as described in Sect. 19.3. *Shape* metrics measure only the observed assemblage phylogeny—the parts of the phylogeny in *black*. *Evenness* metrics measure how evenly species' abundances are distributed across the assemblage phylogeny; the abundances of species in two communities are represented by the size of filled and open *circles* on the figure. *Dispersion* metrics examine whether the observed members of an assemblage are a random subset of the species pool (*grey* and *black* parts of the phylogeny). *Dissimilarity* metrics quantify phylogenetic similarity between observed assemblages. The two assemblages in this figure contain the same species, and so their phylogenetic dissimilarity is null unless abundances are taken into account

19.3 A Systematic Classification of Community Phylogenetic Metrics

While many have reviewed issues in community phylogenetics (e.g. Emerson and Gillespie 2008; Graham and Fine 2008; Cavender-Bares et al. 2009; Vamosi et al. 2009; Mouquet et al. 2012; Swenson 2013), our focus here is specifically on methodology. Pavoine and Bonsall (2011) are of note in that they emphasise analogues between phylogenetic and functional trait diversity and define six major classes of diversity metric. Three of these classes (all shape measures in our classification) are of particular interest here: *multivariate richness* (the sum of a phylogeny's branches; essentially *PD*), *regularity* (the balance of a tree; see Sect. 19.3.1), and *divergence* (the mean distance among species). Vellend et al. (2011) chose a very different scheme, classifying phylogenetic diversity metrics as 'type I' or 'type II' depending on whether they begin by measuring the phylogenetic distinctiveness of species (I) or examine subsets of a regional phylogeny (II).

We propose four classes of community phylogenetic structure with names chosen to reflect existing community ecological literature: shape, evenness, dispersion, and dissimilarity. A graphical overview of these measures is given in Fig. 19.1, and more than 40 metrics are placed into the scheme in the online supplementary materials. *Shape* metrics describe an assemblage phylogeny's topology, branch lengths, size, how closely related its species are, and many

predate community phylogenetics. *Evenness* metrics reflect how species' abundances are distributed throughout a phylogeny, and many are extensions of existing metrics of species diversity. *Dispersion* metrics ask whether an assemblage phylogeny differs from what would be expected, under a given null model, from a source pool phylogeny of potential and actual members of that assemblage. Finally, *dissimilarity* metrics quantify differences in the phylogenetic composition of species occupancy and abundance between assemblages.

19.3.1 Shape

Shape metrics assess the structure of a phylogeny alone and can be calculated with only a list of species and their phylogeny (reviewed in Mooers and Heard 1997). Many predate community phylogenetics itself and were intended for use in macroevolutionary studies. One of the more well known is Colless' Index (I_C), which measures phylogenetic balance as the extent to which nodes in a phylogeny define subgroups of equal size (Colless 1982). An unbalanced assemblage phylogeny indicates that particular clades dominate that assemblage, perhaps because they display key traits that adapt them to that environment. The γ statistic (Pybus and Harvey 2000) was originally intended to detect decreases in the rate of diversification through time; in a community phylogenetic context, this is consistent with an assemblage containing species that are relatively unrelated to one another. Phylogenetic species richness (PSV; Helmus et al. 2007) measures whether the distribution of species across the phylogeny differs from expectation under a Brownian null model and is analogous to the mean phylogenetic distance (MPD) among species on a phylogeny. We caution that an assemblage phylogeny is affected by processes operating outside the assemblage (see Heard and Cox 2007); shape measures sensitive to symmetry at different phylogenetic depths (see Agapow and Purvis 2002) may be useful tools when exploring these issues.

19.3.2 Evenness

Measures of evenness ask whether species abundances are biased towards any particular clade(s) throughout the phylogeny. Many are extensions of existing measures of ecological diversity or shape measures; for instance, the Imbalance in Abundance of higher Clades (IAC; Cadotte et al. 2010) metric is essentially an abundance-weighted form of I_C. Classical measures of the phylogenetic signal of species' traits (e.g. Pagel's λ; 1999) are evenness metrics when calculated using species' abundances, although in most cases statistical transformation of abundances (e.g. taking their logarithm) is advised. Often shape and evenness metrics are calculated for individual sites (α shape/evenness) and across a landscape (γ) to measure β shape or evenness (see Graham and Fine 2008). We intend to use the

term *evenness* analogously to its use in other fields of ecology (see Magurran 2004; Pavoine and Bonsall 2011), but the reader should note that Kraft et al. (2007) and others use the term 'phylogenetic evenness' to indicate communities that contain more distantly related species than expected from null models. We suggest the use of the term 'phylogenetic overdispersion' for this case in reference to the statistical definition of overdispersion (see Sect. 19.3.3 below).

19.3.3 Dispersion

Metrics of phylogenetic dispersion describe whether an observed assemblage is a phylogenetically biased subset of the species that could coexist in that assemblage (the source pool). Bias can reflect community assembly or survival of an extinction episode, and most metrics focus on whether individuals or species are more or less related to one another (under- or overdispersed, respectively) than under a null expectation. They differ from shape and evenness measures, upon which they are often based, in that they require a null expectation; their value is contingent not just upon the observed assemblage but also a null expectation derived from random assembly of same-sized assemblages from a more inclusive source pool. NRI and NTI are the best known: the Net Relatedness Index (NRI) compares the phylogenetic distance among all members of a community, while the Nearest Taxon Index (NTI) examines only distances among nearest relatives. The first definition of NRI and NTI (Webb 2000) counted nodal distance between species, and the second (Webb et al. 2002) used phylogenetic branch lengths. Kembel (2009) defined standard effect sizes of MPD and the mean nearest taxon distance (SES_{MPD} and SES_{MNTD}), which are the negations of NRI and NTI, respectively. Pearse et al. (2013) showed that the randomisations that control for phylogenetic structure in NRI and NTI make the measures test statistics, and so their absolute values can be misleading. They found that D (Fritz and Purvis 2010), which is based upon independent contrasts (Felsenstein 1985) and a Brownian null distribution, can be more sensitive than NRI.

19.3.4 Dissimilarity

Measures of dissimilarity explicitly examine differences in assemblages' compositions, and many have analogues with classical ecological measures (e.g. *PhyloSor* and Sørensen's Index; Bryant et al. 2008). Unlike standard dissimilarity metrics, phylogenetic dissimilarity metrics differentiate among communities with no shared species. The metric phylogenetic community dissimilarity (PCD; Ives and Helmus 2010), for example, partitions dissimilarity into compositional (the proportion of shared species) and phylogenetic (the relatedness of unshared species) components, but the most widely used measure—especially by microbial

ecologists—is *UniFrac* (Lozupone and Knight 2005). *UniFrac* measures the amount of phylogenetic branch length unique to each community, essentially asking how much PD is unique to each community. These measures are distinct from measures of co-evolution (see Chap. 20) in comparing assemblages across a common phylogeny.

19.4 Quantitative Classification of Community Phylogenetic Metrics

Our classification of metrics into four groups is based on how the metrics are calculated and what the metrics attempt to measure. Here, we ask whether the members of our groupings give similar results in common data sets.

Given a particular number of species, n, in a species pool, there is a finite number of unique community compositions that we label the feasible set of community compositions for n species (Haegeman and Loreau 2008; Locey and White 2013). The feasible set can be calculated for any n (though it is often approximated for large n), and given a phylogeny, the distribution of the feasible values of any phylogenetic metric can be derived. Figure 19.2a provides the feasible distribution of PSV for a fully balanced phylogeny of 8 species with equal branch lengths (as in Fig. 19.1). Using this same tree, we simulated 6,000 communities, half structured by phylogenetic attraction where closely related species were more likely to be found together and half by repulsion (the converse). For an ultrametric phylogeny of n species with covariance matrix $\mathbf{V}$, we defined attraction as the Cholesky decomposition of $\mathbf{V}$ and repulsion as the decomposition of $\mathbf{V}^{-1}$, referring to either as the matrix $\mathbf{L}$ below. The probability of species s residing in a simulated community was a stochastic process as defined by

$$p_s = \frac{e^{c\mathbf{LR}}}{1 + e^{c\mathbf{LR}}} \tag{19.1}$$

where c was a scalar (fixed at 10) that determined the strength of attraction/ repulsion and R an $n \times 1$ matrix of normally distributed random numbers centred at 0. The PSV distributions of these two simulated communities (Fig. 19.2b, c) differ markedly from the feasible set (Fig. 19.2a) and from each other, suggesting that metric distributions for empirical and feasible sets of communities can be compared to detect processes that cause phylogenetic attraction and repulsion. To group metrics calculated on these three data sets, we obtained the values of 27 metrics across the simulated communities and hierarchically clustered the metrics (*R* function *hclust*, complete linkage method) based on their standardised (centred to have a mean of zero and standard deviation of 1) Euclidian distances. Our methods were chosen to permit direct comparison with a similar study by Cadotte et al. (2010).
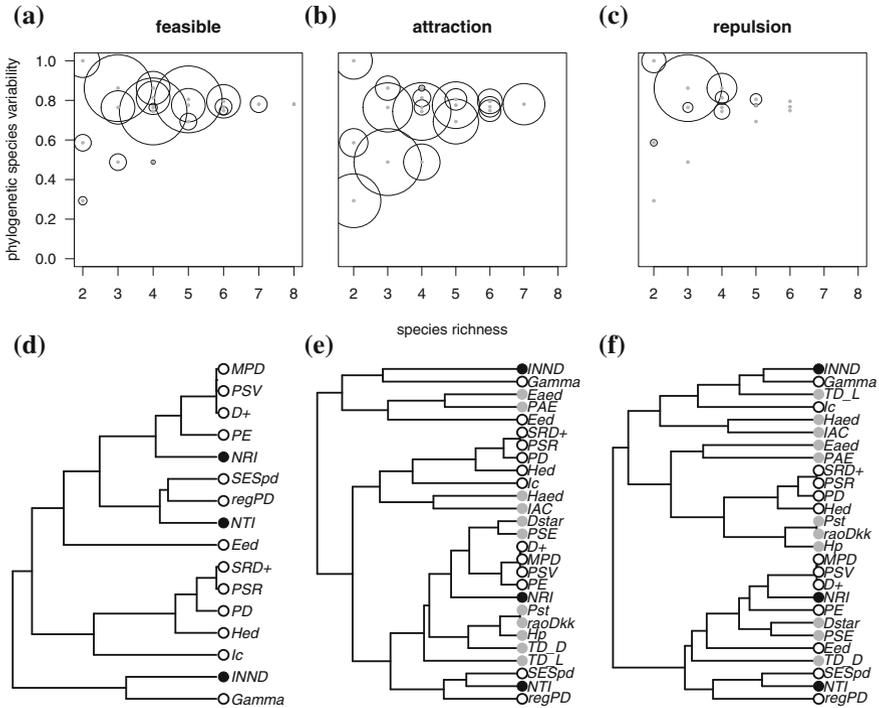
Fig. 19.2 Distributions and clustering dendrograms of community phylogenetic metrics calculated for feasible communities and communities simulated under models of phylogenetic attraction or repulsion. In (**a**), (**b**), and (**c**), the size of the circles (centres marked in *grey*) represent the numbers of unique species compositions that give each PSV value. **a** The feasible distribution of PSV for an eight species, balanced phylogeny. **b** The PSV distribution for communities simulated with attraction is generally lower than the feasible distribution and much lower than the communities simulated under repulsion (**c**). Below each distribution (**d**–**f**) are dendrograms based on a hierarchical clustering of the values of 27 community phylogenetic metrics calculated for each data set. The number of metrics differs among (**d**) and (**e**, **f**) because (**d**) uses the feasible set of species which is defined only for species presence/absence; in (**e**) and (**f**), we simulated abundances and thus incorporated evenness metrics. The *white*, *black*, and *grey* circles indicate shape, dispersion, and evenness metrics, respectively. See the Online Electronic Material for all metric names and abbreviations

The classification of Sect. 19.3 does not perfectly map onto groupings in Fig. 19.2d, e, f; the clustering of the metrics was inconsistent between the attraction and repulsion simulations. This suggests that a single quantitative classification of metrics is unlikely since the metric correlations depend on the underlying data set, but the systematic classification of Sect. 19.3 provides definitive categories for all the metrics.

19.5 Statistical Models of Community Phylogenetic Structure

While community phylogenetic metrics will continue to be developed, explicit statistical models are the next methodological frontier. Models explain phylogenetic structure across a number of assemblages simultaneously, maximising statistical power, and can incorporate phylogenetic, environmental, trait, and other information. While not models in the statistical sense of fitting probability distributions, the first model-based approach stemmed from randomisation methods developed to infer meta-community processes (Pillar and Duarte 2010). There are a number of related approaches (Leibold et al. 2010; Pavoine and Bonsall 2011; Peres-Neto et al. 2012), most involving the comparison of site-by-species matrices with matrices of environmental and species trait data.

Fitting a statistical model to community data allows for estimates of covariate effects and their errors, prediction of community composition, and model comparison using test statistics. Phylogenetic generalised linear mixed models (PGLMMs; Ives and Helmus 2011) were the first statistical models to be developed for community phylogenetics. Here, we illustrate the simplest PGLMM that predicts community composition on the basis of phylogeny alone. Fitting more complex models is useful and possible, but comes with a greater computational complexity and risk of fitting models too complex to be parameterised from the data at hand.

For n species distributed across m sites, the probability of any species being found at a site is logistically modelled as follows:

$$\mu_i = \text{logit}^{-1}(\alpha_{\text{spp}} + b_i + c_{\text{site}}) \tag{19.2}$$

where i indexes a particular spp at a particular *site*, α_{spp} is a categorical fixed effect that accounts for variation in species prevalence across communities, and b_i is a Gaussian distributed random effect with mean 0 that accounts for phylogeny. The covariance matrix of b_i is the Kronecker product of $\mathbf{I_m}$ and $\sigma^2_{\text{spp}}\mathbf{V_{spp}}$, where $\mathbf{I_m}$ is the $m \times m$ identity matrix and $\sigma^2_{\text{spp}}\mathbf{V_{spp}}$ an estimated scalar multiplied by the $n \times n$ phylogenetic covariance matrix. The resulting covariance matrix of b_i is block diagonal with $\sigma^2_{\text{spp}}\mathbf{V_{spp}}$ repeated as the blocks and zeroes elsewhere. Lastly, c_{site} is similar to b but with ones as the blocks in its covariance matrix. Including α_{spp} and c_{site} separates differences in species prevalence across communities or community size from phylogenetic effects.

We fit the PGLMM in Eq. 19.2 to the two simulated data sets from Sect. 19.4 (depicted in Fig. 19.2b and c). Note that the estimated scalar σ^2_{spp} gives the strength of the phylogenetic attraction, not repulsion, and so we expected σ^2_{spp} to only be significant for the communities simulated with attraction. Indeed, we identified significant phylogenetic pattern only in the attraction-simulated communities (attraction σ^2_{spp}: 0.87, 0.78–0.95 95 % CI, repulsion σ^2_{spp}: 0.0, 0.0–0.1 95 % CI). To test for repulsion, we altered the covariance matrix of b_i by replacing

$\sigma_{spp}\mathbf{V_{spp}}$ with $\sigma_{spp}\mathbf{V_{spp}^{-1}}$ (i.e. we replaced b_i with d_i from model III of Ives and Helmus 2011) and fit this new model to both data sets (note the similarity with Sect. 19.4 model). This second PGLMM only detected phylogenetic pattern in the communities simulated under the repulsion model (attraction σ_{spp}^2: 0.0, 0.0–0.10 95 % CI, repulsion σ_{spp}^2: 0.96, 0.87–1.06 95 % CI). The nature of this PGLMM and its performance detecting phylogenetic dispersion did not change, as many metrics did in Sect. 19.4. However, these models were calculated across only the first 50 of the communities simulated in Sect. 19.4 due to computational limitations, although new algorithms may overcome this (Ho and Ané 2014).

19.6 Future Developments

Studies often use trait data to justify investigators' assumption of niche conservatism and thus map phylogenetic pattern onto ecological process (which is contentious at best; Cavender-Bares et al. 2009). Yet if phylogeny is only a proxy for species traits, it is unclear why a phylogenetic 'middleman' (Swenson 2013) is needed when the trait data themselves are available. If we are to claim that a perfect phylogeny reflects species' niches better than trait data (Srivastava et al. 2012) or that phylogeny is a useful proxy for difficult to obtain functional trait data (Mace et al. 2003), then we must directly compare the explanatory power of traits and phylogeny. Yet phylogenetic signal in a trait does not mean phylogenetic and trait data are in perfect agreement; even if they were (not), measurement error may falsely indicate (dis)agreement. Recent developments, such as the *traitgram* approach (Cadotte et al. 2013), allow the explanatory power of phylogeny and traits to be partitioned and interactions between the two explored. An alternative is to contrast the evolution of species' traits with their present-day ecology; Cavender-Bares et al. (2006) found that traits critical to habitat filtering (such as plant height) were convergent, while traits associated with local competition (such as leaf habit) were conserved in oak trees. Silvertown et al. (2006; also see Ackerly et al. 2006) went a step further, categorising traits as α, β, or γ depending on their order of evolutionary divergence and relating these evolutionary dynamics to the likelihood of species coexisting in the present.

Community phylogenetics provides an excellent framework within which to examine the 'problem and promise of scale dependency' (Swenson et al. 2006), and spatial and taxonomic scaling continues to draw interest (e.g. Cavender-Bares et al. 2006; Kembel and Hubbell 2006). There is evidence of variation among clades in phylogenetic structure even within well-defined groups (e.g. Parra et al. 2010) and tentative evidence of links between clade age and phylogenetic dispersion (Pearse et al. 2013). Variation among clades is to be expected; under a Brownian model of trait evolution (which most metrics assume or, like PSV, are derived under; see also Peres-Neto et al. 2012), phylogeny is a poorer predictor of similarity for distantly related species. Advances in the modelling of species' trait evolution (reviewed in see also Cooper et al. 2010; Chap. 14) have provided us

with more sophisticated models of trait evolution, which should generate different expectations for ecological dissimilarity and so present-day phylogenetic structure. Indeed, the mode of speciation in a clade could affect its community phylogenetic structure today. To give a simplified example, species brought back into secondary contact may be unlikely to coexist due to shared environmental tolerance (or gene flow; see Fig. 5 in Cavender-Bares et al. 2009), while descendents of rapid adaptive radiations might be sufficiently dissimilar to coexist.

19.7 Conclusion

There will never be one perfect definition of an ecological assemblage, and so there will never be one perfect way of describing one. It is no surprise that some express misgivings about the incursion of phylogenetic structure into ecology; initial attempts to incorporate phylogenies into comparative analysis were met with criticism, and many feared that implicit assumptions of the approach were ignored (e.g. Westoby et al. 1995). Such initial scepticism is healthy—there is always a danger that a new framework will be applied simply because it can be, without any critical evaluation of its implications. The incorporation of phylogenetic structure into ecology is not without its pitfalls, but a little over a decade since Webb et al. (2002) outlined their research paradigm we have a remarkably mature suite of metrics and methods. Looking forward, ecologists and evolutionary biologists are moving beyond describing phylogenetic structure, and instead, testing detailed hypotheses about how that structure came to be. Species' evolutionary history was shaped by their ecology, and it seems natural to see what the shape of species' evolutionary past can reveal about their ecology today.

References

Ackerly DD, Schwilk DW, Webb CO (2006) Niche evolution and adaptive radiation: testing the order of trait divergence. Ecology 87:S50–S61

Agapow P-M, Purvis A (2002) Power of eight tree shape statistics to detect nonrandom diversification: a comparison by simulation of two models of cladogenesis. Syst Biol 51(6):866–872

Altschul SF, Lipman DJ (1990) Equal animals. Nature 348:493–494

Bryant JA et al (2008) Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. Proc Natl Acad Sci 105(S1):11505–11511

Cadotte M, Albert CH, Walker SC (2013) The ecology of differences: assessing community assembly with trait and evolutionary distances. Ecol Lett 16(10):1234–1244

Cadotte MW et al (2010) Phylogenetic diversity metrics for ecological communities: integrating species richness, abundance and evolutionary history. Ecol Lett 13(1):96–105

Cavender-Bares J, Keen A, Miles B (2006) Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. Ecology 87(7):S109–S122

Cavender-Bares J et al (2004) Phylogenetic overdispersion in Floridian oak communities. Am Nat 163(6):823–843

Cavender-Bares J et al (2009) The merging of community ecology and phylogenetic biology. Ecol Lett 12:693–715

Colless DH (1982) Review of phylogenetics: the theory and practice of phylogenetic systematics. Syst Zool 31(1):100–104

Cooper N, Jetz W, Freckleton RP (2010) Phylogenetic comparative approaches for studying niche conservatism. J Evol Biol 23(12):2529–2539

Darwin C (1859) On the origin of species. John Murray, London

Douglas ME, Matthews WJ (1992) Does morphology predict ecology? Hypothesis testing within a freshwater stream fish assemblage. Oikos 65(2):213–224

Elton C (1946) Competition and the structure of ecological communities. J Anim Ecol 15(1):54–68

Emerson BC, Gillespie RG (2008) Phylogenetic analysis of community assembly and structure over space and time. Trends Ecol Evol 23(11):619–630

Faith DP (1992) Conservation evaluation and phylogenetic diversity. Biol Conserv 61(1):1–10

Felsenstein J (1985) Phylogenies and the comparative method. Am Nat 125:1–15

Fritz SA, Purvis A (2010) Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. Conserv Biol 24(4):1042–1051

Graham CH, Fine PVA (2008) Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. Ecol Lett 11(12):1265–1277

Haegeman B, Loreau M (2008) Limitations of entropy maximization in ecology. Oikos 117:1700–1710

Heard SB, Cox GH (2007) The shapes of phylogenetic trees of clades, faunas, and local assemblages: exploring spatial pattern in differential diversification. Am Nat 169(5):E107–E118

Helmus MR et al (2007) Phylogenetic measures of biodiversity. Am Nat 169(3):E68–E83

Ho LST, Ane C (2014). A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. Syst Biol 63:397–408

Isaac NJB et al (2007) Mammals on the EDGE: conservation priorities based on threat and phylogeny. PLoS ONE 2(3):e296

Ives AR, Helmus MR (2010) Phylogenetic metrics of community similarity. Am Nat 176(5):E128–E142

Ives AR, Helmus MR (2011) Generalized linear mixed models for phylogenetic analyses of community structure. Ecol Monogr 81(3):511–525

Izsak C, Price ARG (2001) Measuring β-diversity using a taxonomic similarity index, and its relation to spatial scale. Mar Ecol Prog Ser 215:69–77

Izsáki J, Papp L (1995) Application of the quadratic entropy indices for diversity studies of drosophilid assemblages. Environ Ecol Stat 2(3):213–224

Jaccard P (1901) Etude comparative de la distribution florale dans une portion des Alpes et des Jura. Bull de la Soc Vaudoise des Sci Nat 37:547–579

Järvinen O (1982) Species-to-genus ratios in biogeography: a historical note. J Biogeogr 9(4):363–370

Kembel SW (2009) Disentangling niche and neutral influences on community assembly: assessing the performance of community phylogenetic structure tests. Ecol Lett 12(9):949–960

Kembel SW, Hubbell SP (2006) The phylogenetic structure of a neotropical forest tree community. Ecology 87(7):S86–S99

Kraft NJB et al (2007) Trait evolution, community assembly, and the phylogenetic structure of ecological communities. Am Nat 170(2):271–283

Leibold MA, Economo EP, Peres-Neto P (2010) Metacommunity phylogenetics: separating the roles of environmental filters and historical biogeography. Ecol Lett 13(10):1290–1299

Locey KJ, White EP (2013) How species richness and total abundance constrain the distribution of abundance. Ecol Lett 16(9):1177–1185

Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol 71(12):8228–8235

Mace GM, Gittleman JL, Purvis A (2003) Preserving the tree of life. Science 300(5626):1707–1709

Magurran AE (2004). Measuring biological diversity. Oxford University Press, Oxford

May RM (1990) Taxonomy as destiny. Nature 347:129–130

Mayfield MM, Levine JM (2010) Opposing effects of competitive exclusion on the phylogenetic structure of communities. Ecol Lett 13(9):1085–1093

Mooers AØ, Heard SB (1997) Inferring evolutionary process from phylogenetic tree shape. Q Rev Biol 72(1):31–54

Mouquet N et al (2012) Ecophylogenetics: advances and perspectives. Biol Rev 87(4):769–785

Pagel M (1999) Inferring the historical patterns of biological evolution. Nature 401(6756):877–884

Parra JL, McGuire JA, Graham CH (2010) Incorporating clade identity in analyses of phylogenetic community structure: an example with hummingbirds. Am Nat 176(5):573–587

Pavoine S, Bonsall MB (2011) Measuring biodiversity to explain community assembly: a unified approach. Biol Rev 86(4):792–812

Pavoine S, Ollier S, Dufour A-B (2005) Is the originality of a species measurable? Ecol Lett 8(6):579–586

Pearse WD, Jones A, Purvis A (2013) Barro Colorado Island's phylogenetic assemblage structure across fine spatial scales and among clades of different ages. Ecology 94(12):2861–2872

Peres-Neto PR, Leibold MA, Dray S (2012) Assessing the effects of spatial contingency and environmental filtering on metacommunity phylogenetics. Ecology 93:S14–S30

Pillar VD, Duarte LS (2010) A framework for metacommunity analysis of phylogenetic structure. Ecol Lett 13(5):587–596

Pybus OG, Harvey PH (2000) Testing macro-evolutionary models using incomplete molecular phylogenies. Proc Roy Soc B Biol Sci 267(1459):2267–2272

R Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Redding DW, Mooers AØ (2006) Incorporating evolutionary measures into conservation prioritization. Conserv Biol 20(6):1670–1678

Silvertown J et al (2006) Phylogeny and the hierarchical organization of plant diversity. Ecology 87(7):S39–S166

Srivastava DS et al (2012) Phylogenetic diversity and the functioning of ecosystems. Ecol Lett 15(7):637–648

Swenson NG (2013) The assembly of tropical tree communities—the advances and shortcomings of phylogenetic and functional trait analyses. Ecography 36(3):264–276

Swenson NG et al (2006) The problem and promise of scale dependency in community phylogenetics. Ecology 87(10):2418–2424

Valiente-Banuet A, Verdú M (2007) Facilitation can increase the phylogenetic diversity of plant communities. Ecol Lett 10(11):1029–1036

Vamosi S et al (2009) Emerging patterns in the comparative analysis of phylogenetic community structure. Mol Ecol 18(4):572–592

Vane-Wright RI, Humphries CJ, Williams PH (1991) What to protect?—systematics and the agony of choice. Biol Conserv 55(3):235–254

Vellend M et al. (2011) "Measuring phylogenetic biodiversity". In: Magurran AE, McGill BJ Biological Diversity, Oxford University Press, Oxford Chap. 14

Warwick RM, Clarke KR (1995) New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. Mar Ecol Prog Ser 129(1):301–305

Webb CO (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. Am Nat 156(2):145–155

Webb CO et al (2002) Phylogenies and community ecology. Annu Rev Ecol Syst 33(1):475–505

Westoby M, Leishman MR, Lord JM (1995) On misinterpreting the 'phylogenetic correction'. J Ecol 83(3):531–534

Wiens JJ et al (2010) Niche conservatism as an emerging principle in ecology and conservation biology. Ecol Lett 13(10):1310–1324

Measure	Reference	Description
Colless' Index (I_C)	Colless (1982)	Comparison of the number of tips either 'side' of each node with the number in a perfectly balanced tree.
PD (Phylogenetic Distance)	Faith (1992)	Sum of the branch lengths in a phylogeny.
MPD (Mean Phylogenetic Distance)	Unknown	The mean of PD; some use the median phylogenetic distance (e.g., Maitner et al., 2012).
Δ (taxonomic diversity index)	Warwick & Clarke (1995)	Extension of the Shannon index, with weighting according to the absolute taxonomic distance between individuals.
Δ^+ (taxonomic distinctiveness index; TDI_D+	Warwick & Clarke (1995)	Essentially Δ divided by the value of Δ given the simplest possible taxonomic structure (all species in the same genus).
γ	Pybus & Harvey (2000)	γ values < 0 suggest internal nodes are closer to the root than expected. The converse cannot be tested with γ.
LTT (Lineages Through Time) Plot	Martin (2002)	Number of lineages through time in a phylogeny. Not a metric (rather a plot), but Martin (2002) suggest departure from a curve generated under a Yule null model indicates overdispersion or clustering.
PSV (Phylogenetic Species Variability)	Helmus et al. (2007)	Compares branch lengths of species with the variance expected under the Brownian motion of a neutral trait across a phylogeny. Claims specifics of the model are independent of Brownian model.
PSR (Phylogenetic Species Richness)	Helmus et al. (2007)	Product of PSV and phylogeny size.
Corrected PD (regPD)	Vamosi & Vamosi (2007)	Residuals of regression of PD on species richness. Intended to control for the confound of species richness and phylogenetic diversity.
E_{ED}	Cadotte et al. (2010a)	Equitability of an entropic measure of diversity of species' evolutionary distinctiveness.
Phylogenetic Eigenvector Methods (PE)	Diniz-Filho et al. (2011)	Sum of the axes retained in a principal coordinate analysis of a phylogenetic distance matrix is related to phylogenetic imbalance; in our analyses, we used the first two axes. See Freckleton et al. (2011) for a discussion of issues with the approach in comparative analysis.
EPD (Expected Phylogenetic Diversity)	O'Dwyer et al. (2012)	Sum, across all potential sizes of clade, of the product of a clade having at least one species present in a community and the sum of all branch lengths in all clades of that size. Since analytical null expectations can be generated, this facilitates the examination of extremely large phylogenies without time-consuming null randomisations.

Table 1 Phylogenetic *shape* measures, in (approximate) order of publication. The abbreviations used in the chapter's Fig. 2 are given in the 'Measure' column.

Measure	Reference	Description
F_{ST}	Martin (2002)	Standard population genetics method, showing how differentiated samples are. Re-purposed by Martin (2002) for use in concert with the *P test*, to study the phylogenetic scale of clustering.
P_{ST}	Hardy & Senterre (2007)	An extension of Simpson's index to incorporate phylogenetic relatedness. Assesses how much of a site's diversity is explained by turnover within sites. Part of the framework mentioned in Π_{ST} below.
Π_{ST}	Hardy & Senterre (2007)	Part of a family of measures partitioning and examining hierarchical partitioning of evenness (and shape). Many other terms are defined, but they are (acknowledged to be) redefinitions of measures described elsewhere in these tables; interested readers should consult Hardy & Senterre (2007) for more details.
PSE (Phylogenetic Species Evenness)	Helmus et al. (2007)	As *PSV*, but with polytomies with branch lengths of zero and sizes equal to the frequency of members of the community added to the phylogeny.
Phylogenetic Entropy (Hp)	Allen et al. (2009)	The product of two species co-occurring and their natural logarithm of their phylogenetic relatedness. Generalisation of the Shannon Index.
$H_\alpha, H_\beta, H_\gamma$ $PD_\alpha, PD_\beta, PD_\gamma$	Mouchet & Mouillot (2011) Cadotte et al. (2010b)	An extension of the phylogenetic entropy concept of Allen et al. (2009). PD_α is the *MPD* within sites, PD_γ is the sum of the branch lengths connecting all species across sites, and PD_β is $PD_\gamma PD_\alpha$. Uses the Lande (1996) partitioning of diversity; note that PD_α is constructed from *MPD*, not *PD*.
PAE (Phylogenetic Abundance Evenness)	Cadotte et al. (2010a)	A scaled measure of the product of species frequency and each species' terminal branch length.
IAC	Cadotte et al. (2010a)	Relative per-node imbalance in the distribution of individuals. Can be thought of as an abundance-weighted form of I_C; ignores branch lengths.
E_{AED}	Cadotte et al. (2010a)	Equitability of an abundance-weighted entropic measure of diversity of species' evolutionary distinctiveness. Similar to E_{ED}.

Table 2 Phylogenetic *evenness* measures, in (approximate) order of publication. The abbreviations used in the chapter's Fig. 2 are given in the 'Measure' column.

Measure	Reference	Description
NRI (Net Relatedness Index)	Webb (2000); Webb *et al.* (2002)	Null randomisations' mean *MPD* subtracted from an observed assemblage's *MPD*, scaled by the standard deviation of those null randomisations. The 2000 definition is based on nodal distance.
NTI (Nearest Taxon Index)	Webb (2000); Webb *et al.* (2002)	Null randomisations' mean smallest between-species phylogenetic distance *MPD* subtracted from an observed assemblage's mean least phylogenetic distance, scaled by the standard deviation of those null randomisations. The 2000 definition is based on nodal distance.
SES$_{MPD}$ (Standardised Effect Size of Mean Phylogenetic Distance)	Kembel (2009)	The negation of *NRI* (Webb *et al.*, 2002).
SES$_{MNTD}$ (Standardised Effect Size of Mean Phylogenetic Distance)	Kembel (2009)	The negation of *NTI* (Webb *et al.*, 2002).
SES$_{PD}$ (Standardised Effect Size of Phylogenetic Diversity)	Proche *et al.* (2006)	Null randomisations' mean *PD* subtracted from an observed assemblage's *PD*, scaled by the standard deviation of those null randomisations.
Summed number of state changes	Chazdon *et al.* (2003)	Reconstructed number of state changes in community presence across the phylogeny; fewer changes indicate greater phylogenetic conservatism (and so phylogenetic clustering). Many methods for ancestral state reconstruction exist, this is the first example we found of it being used in a community phylogenetic context.
D	Fritz & Purvis (2010)	Sum of differences in independent contrasts of species' presence/absence, scaled according to random and Brownian expectation.
INND/MIPD (Inverse Nearest Neighbour Distance/Mean of Inverse Pairwise Distances)	Ness *et al.* (2011)	As *NRI/NTI* (Webb *et al.*, 2002), but using the *inverse* of phylogenetic distance. Assumes a non-linear relationship between ecological and phylogenetic distance.
Standardised *PD* (*stdPD*)	Mi *et al.* (2012)	Difference between observed *PD* of subsets of a sample and the total *PD* of that sample. Used to examine influence of rare species on *PD* by regressing *stdPD* against the rank abundance distribution.

Table 3 Phylogenetic *dispersion* measures. The abbreviations used in the chapter's Fig. 2 are given in the 'Measure' column.

Measure	Reference	Description
Quadratic Diversity (raoDkk)	Rao (1982)	Sum of the product of the proportions of sites where two species are present, and their phylogenetic relatedness. Generalisation of the Simpson Index.
P Test	Maddison & Slatkin (1991); Martin (2002)	Compares distribution of sequences unique to each community within the phylogeny and compares the number of evolutionary transitions required for this distribution with a null model. A parsimony-based approach adapted from the cladistic literature by Martin.
Θ^+	Clarke & Warwick (1998); Clarke et al. (2006)	As Γ^+, but each samples minimum distances are kept listed separately and then averaged. Related to the Bray-Curtis index.
Γ^+	Izsak (2001); Clarke et al. (2006)	Mean of the smallest taxonomic distances between species in each sample. Related to the Bray-Curtis index.
Δ_C (raoDkk)	Singleton (2001)	Compares curves of number of unique sequences within a sample against a cut-off for sequence dissimilarity. Designed for bacterial systems and rRNA.
UniFrac	Lozupone & Knight (2005)	The fraction of total branch length that two communities share. Can thus create a distance matrix for all communities to which one can apply multivariate statistics. Widely used by the microbiological community, but ignored by the wider ecological community.
PhyloSor	Bryant et al. (2008)	Fraction of branch length shared among communities. So-called because of its derivation from the Sørensen's index.
βNRI / βNTI	Webb et al. (2008)	Mean phylogenetic distance (or mean nearest-relative phylogenetic distance) between communities. Terminology from Fine & Kembel (2011), who called NRI αNTI and NTI αNTI and αNRI to emphasise the links between them.
PCD	Ives & Helmus (2010)	An extension of PSV that partitions dissimilarity into phylogenetic and non-phylogenetic components.

Table 4 Phylogenetic *dissimilarity* measures, in (approximate) order of publication.

References

Allen, B., M. Kon, & Y. Bar-Yam (2009). A new phylogenetic diversity measure generalizing the Shannon index and its application to phyllostomid bats. *The American Naturalist* 174.2, pp. 236–243.

Bryant, J. A. *et al.* (2008). Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences* 105.Supplement 1, pp. 11505–11511.

Cadotte, M. W. *et al.* (2010a). Phylogenetic diversity metrics for ecological communities: integrating species richness, abundance and evolutionary history. *Ecology Letters* 13.1, pp. 96–105.

Cadotte, M. W. *et al.* (2010b). Phylogenetic patterns differ for native and exotic plant communities across a richness gradient in Northern California. *Diversity and Distributions* 16.6, pp. 892–901.

Chazdon, R. L. *et al.* (2003). Community and Phylogenetic Structure of Reproductive Traits of Woody Species in Wet Tropical Forests. *Ecological Monographs* 73.3, pp. 331–348.

Clarke, K. R. & R. M. Warwick (1998). A Taxonomic Distinctness Index and its Statistical Properties. *Journal of Applied Ecology* 35.4, pp. 523–531.

Clarke, K. R., P. J. Somerfield, & M. G. Chapman (2006). On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray-Curtis coefficient for denuded assemblages. *Journal of Experimental Marine Biology and Ecology* 330.1, pp. 55–80.

Colless, D. H. (1982). Review of phylogenetics: the theory and practice of phylogenetic systematics. *Systematic Zoology* 31.1, pp. 100–104.

Diniz-Filho, J. A. F. *et al.* (2011). Eigenvector estimation of phylogenetic and functional diversity. *Functional Ecology* 25.4, pp. 735–744.

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological conservation* 61.1, pp. 1–10.

Fine, P. V. A. & S. W. Kembel (2011). Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. *Ecography* 34.4, pp. 552–565.

Freckleton, R. P., N. Cooper, & W. Jetz (2011). Comparative Methods as a Statistical Fix: The Dangers of Ignoring an Evolutionary Model. English. *The American Naturalist* 178.1, pp. E10–E17.

Fritz, S. A. & A. Purvis (2010). Selectivity in Mammalian Extinction Risk and Threat Types: a New Measure of Phylogenetic Signal Strength in Binary Traits. *Conservation Biology* 24.4, pp. 1042–1051.

Hardy, O. J. & B. Senterre (2007). Characterizing the phylogenetic structure of communities by an additive partitioning of phylogenetic diversity. *Journal of Ecology* 32.0, pp. 493–506.

Helmus, M. R. *et al.* (2007). Separating the determinants of phylogenetic community structure. *Ecology Letters* 10.10, pp. 917–925.

Ives, A. R. & M. R. Helmus (2010). Phylogenetic Metrics of Community Similarity. *The American Naturalist* 176.5, E128–E142.

Izsak, C (2001). Measuring α-diversity using a taxonomic similarity index, and its relation to spatial scale. *Marine ecology. Progress series* 215.

Kembel, S. W. (2009). Disentangling niche and neutral influences on community assembly: assessing the performance of community phylogenetic structure tests. *Ecology Letters* 12.9, pp. 949–60.

Lande, R. (1996). Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* 76.1, pp. 5–13.

Lozupone, C. & R. Knight (2005). UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology* 71.12, pp. 8228–8235.

Maddison, W. P. & M. Slatkin (1991). Null Models for the Number of Evolutionary Steps in a Character on a Phylogenetic Tree. *Evolution* 45.5, pp. 1184–1197.

Maitner, B. S. *et al.* (2012). Patterns of bird invasion are consistent with environmental filtering. *Ecography* 35.7, pp. 614–623.

Martin, A. P. (2002). Phylogenetic Approaches for Describing and Comparing the Diversity of Microbial Communities. *Applied and Environmental Microbiology* 68.8, pp. 3673–3682.

Mi, X. *et al.* (2012). The Contribution of Rare Species to Community Phylogenetic Diversity across a Global Network of Forest Plots. English. *The American Naturalist* 180.1, pp. E17–E30.

Mouchet, M. A. & D. Mouillot (2011). Decomposing phylogenetic entropy into α, β and γ components. *Biology Letters* 7.2, pp. 205–209.

Ness, J. H., E. J. Rollinson, & K. D. Whitney (2011). Phylogenetic distance can predict susceptibility to attack by natural enemies. *Oikos* 120, pp. 1327–1334.

O'Dwyer, J. P., S. W. Kembel, & J. L. Green (2012). Phylogenetic Diversity Theory Sheds Light on the Structure of Microbial Communities. *PLoS Comput Biol* 8.12, e1002832.

Proche, ., J. R. Wilson, & R. M. Cowling (2006). How much evolutionary history in a 1010Łm plot? *Proceedings of the Royal Society B: Biological Sciences* 273.1590, pp. 1143–1148.

Pybus, O. & P. Harvey (2000). Testing macro–evolutionary models using incomplete molecular phylogenies. *Proceedings of the Royal Society B: Biological Sciences* 267.1459, pp. 2267–2272.

Rao, C. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology* 21, pp. 24–43.

Singleton, D. R. (2001). Quantitative comparisons of 16S rRNA gene sequence libraries from environmental samples. *Applied and environmental microbiology* 67.9, pp. 4374–4376.

Vamosi, J. C. & S. M. Vamosi (2007). Body size, rarity, and phylogenetic community structure: insights from diving beetle assemblages of Alberta. *Diversity and Distributions* 13.1, pp. 1–10.

Warwick, R. M. & K. R. Clarke (1995). New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. *Marine Ecology Progress Series* 129.1.

Webb, C. O., C. H. Cannon, & S. J. Davies (2008). "Ecological organization, biogeography, and the phylogenetic structure of tropical forest tree communities". *Tropical Forest Community Ecology*. Ed. by W. P. Carson & S. A. Schnitzer. Wiley Blackwell. Chap. Ecological organization, biogeography, and the phylogenetic structure of tropical forest tree communities, pp. 79–98.

Webb, C. O. (2000). Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist* 156.2, pp. 145–155.

Webb, C. O. *et al.* (2002). Phylogenies and Community Ecology. *Annual Review of Ecology and Systematics* 33.1, pp. 475–505.